AMIA

INFORMATICS PROFESSIONALS. LEADING THE WAY.

2026 Annual Symposium

#AMIA2026



Target Trial Emulation & Real-World Data

Speaker: Nansu Zong
Institution: Mayo Clinic
Connect with me Zong.Nansu@mayo.edu

Disclosure of Relevant Financial Relationships

NA

Learning Objectives

At the conclusion of this presentation the learner will be able to:

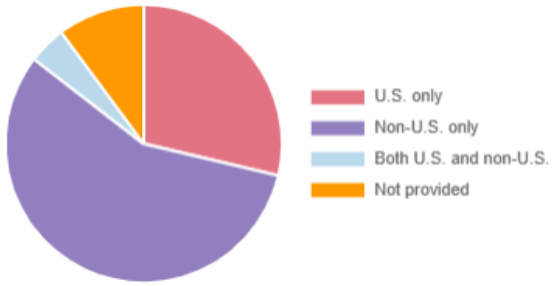
1. What is Clinical Trial Simulation
2. Key Research Topics

Content

- Patient Eligibility Matching
- LLM for Data Extraction
- Large-scale Simulation Study for Drug Repurposing
- Trial Optimization

Clinical Trials

Percentage of registered studies by location (as of 2026-02-25)
Total of 573,165 studies



Location	Number of Registered Studies and Percentage of Total (as of 2026-02-25)
U.S. only	164,296 (29%)
Non-U.S. only	325,254 (57%)
Both U.S. and non-U.S.	25,332 (4%)
Not provided	58,283 (10%)
Total	573,165 (100%)

Number of Registered Studies Over Time (as of 2026-02-25)



Source: https://ClinicalTrials.gov

Note:
The International Committee of Medical Journal Editors (ICMJE) began requiring trial registration as a condition of publication in September 2005
The registration requirements of Food and Drug Administration Amendments Act of 2007 (FDAAA) began and were implemented on ClinicalTrials.gov in December 2007

Study and Intervention Type (as of 2026-02-25)		Number of Registered Studies and Percentage of Total	Number of Studies With Posted Results and Percentage of Total***
Total		573,165	76,687
Interventional		437,507 (76%)	72,171 (94%)
Type of Intervention*	Drug or biologic	213,778	51,359
	Behavioral, other	169,552	17,929
	Surgical procedure	46,257	3,672
	Device**	60,323	10,980
Observational		133,660 (23%)	4,516 (6%)
Expanded Access		1,033	N/A

$4 million (Phase I) to over $100 million (Phase III) and total development often taking **7–10+ years**.
Median costs for a single trial are estimated around **$19 million to $48 million**

Clinical Trial Cost and Time Breakdown

- **Phase I (Safety)**: ~$3.4 million - $4 million, takes months to 1 year.
- **Phase II (Effectiveness)**: ~$8.6 million - $13 million, lasts up to 2 years.
- **Phase III (Confirmation)**: ~$20 million - $100M+, lasts 3-5 years.
- **Phase IV (Post-marketing)**: ~$20 million

https://clinicaltrials.gov/about-site/trends-charts

Clinical Trials



6-10 Participants

20-50 Participants

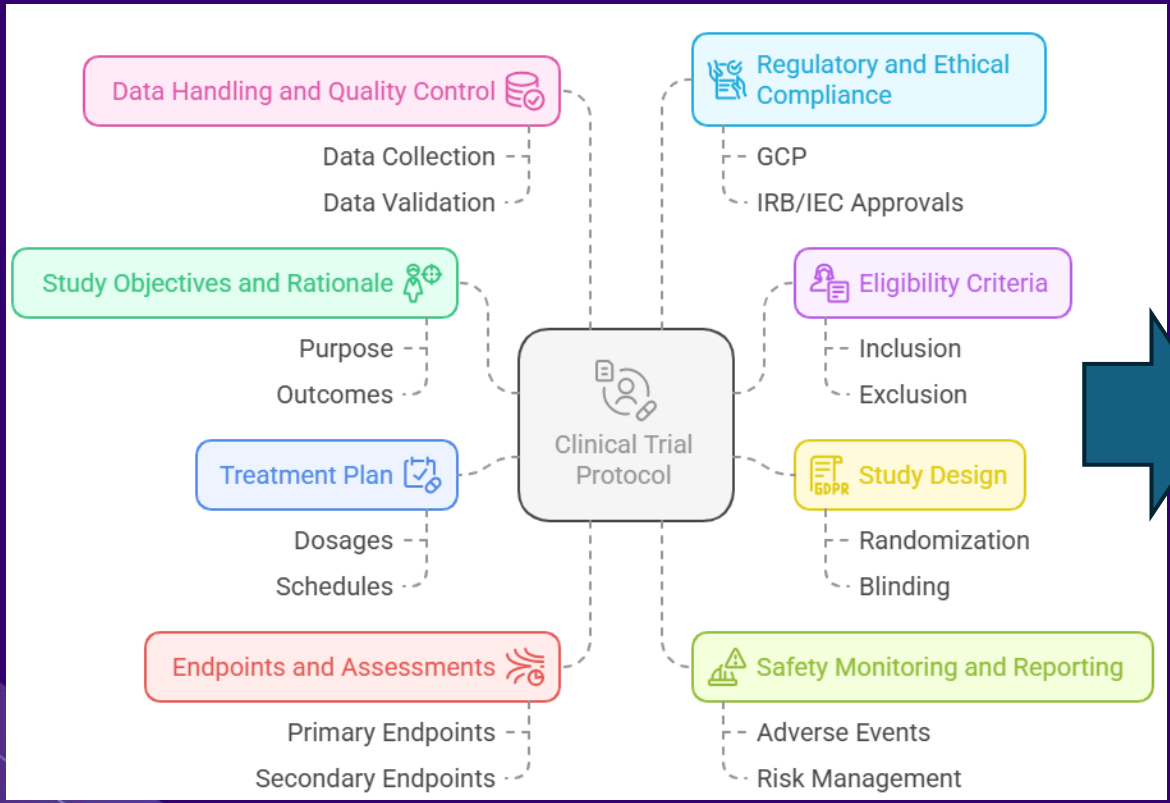
100-200 Participants

200+ Participants

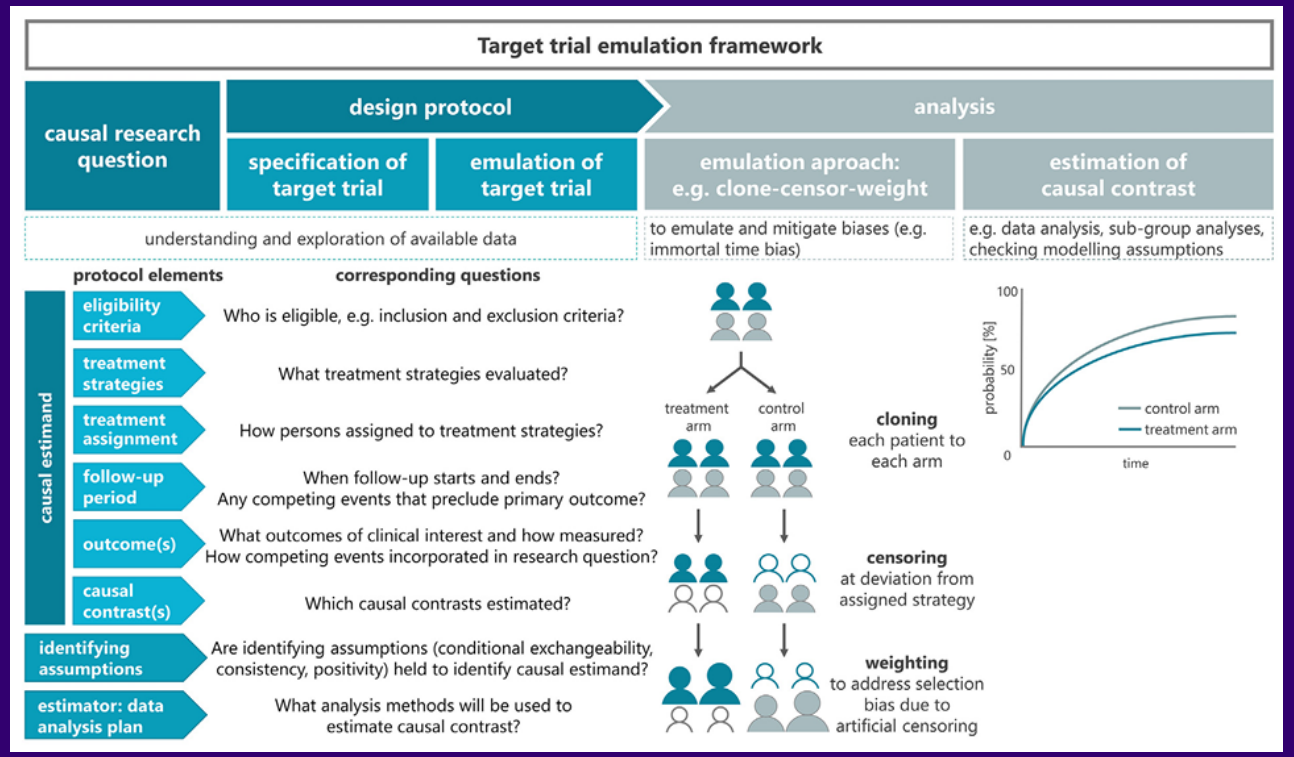
Phase	Core Focus	Category	Key Factors	Typical Design
I	Safety and Dosage	Safety and Tolerability	Toxicity, Maximum Tolerated Dose, Pharmacokinetics, Adverse Events	Single-Arm
II	Proof of Concept	Clinical Activity and Biomarkers	Biomarker Changes, Objective Response Rate, Surrogate Endpoints, Dosage Optimization	Single-Arm OR Randomized
III	Confirmatory Efficacy	Clinical Benefit and Risk-Benefit Ratio	Overall Survival, Progression-Free Survival, Mortality Rate, Quality of Life	Randomized Controlled
IV	Post-Marketing Surveillance	Long-term Safety and Real-World Use	Rare Adverse Reactions, Real-World Evidence, Long-term Survival, Pharmacoeconomics	

https://gene.vision/knowledge-base/clinical-trials/
https://lustgarten.org/clinical-trial-phases/

Clinical Trials and Emulation

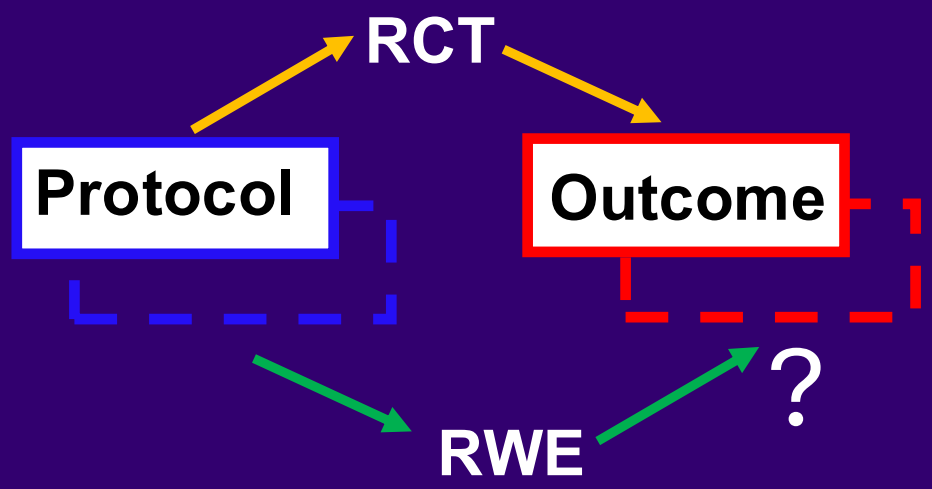


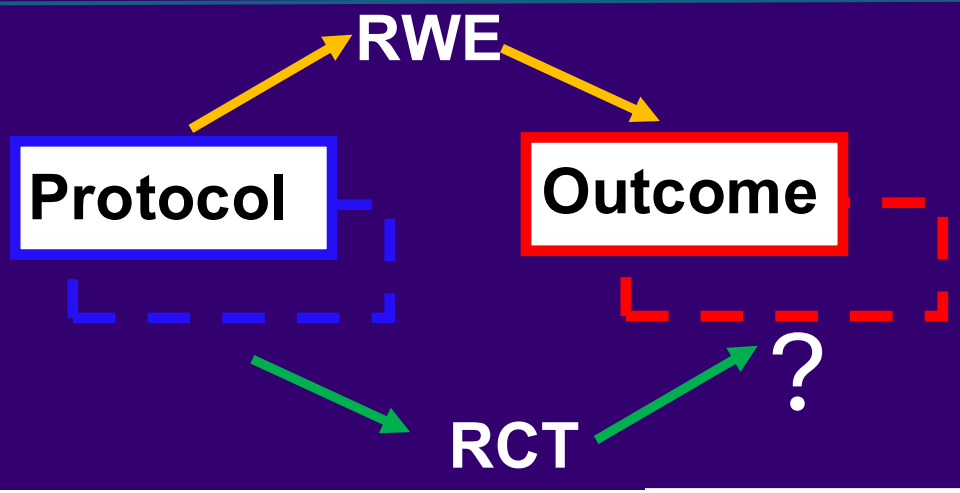
Clinical Trial Protocol: The Definitive Guide with Examples, https://ccrps.org/



Oksana Martinuka, Clinical Microbiology and Infection, 2025

Research Question-Prediction of Success of Clinical Trial





RWD and RWE-focused Demonstration Projects

Share | Post | LinkedIn | Email | Print

To promote shared learning and understanding with external stakeholders, FDA continues to oversee and support numerous demonstration (i.e., research) projects through multiple venues and programs including the broad agency announcements, funding opportunity announcements, CERSI grants, and inter-Agency agreements.

The table below lists ongoing and completed demonstration projects.

Use filters and search box to find resources

Advanced search (combine topic and status terms)

Topic | Status

Clear Filters

FDA funded 26 projects

npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01705-z

Advancing efficacy prediction for electronic health records based emulated trials in repurposing heart failure therapies

Check for updates

Nansu Zong[1,8], Shaika Chowdhury[1,8], Shibo Zhou[1,8], Sivaraman Rajaganapathy[1], Yue Yu[2], Liewei Wang[3], Qiying Dai[4], Pengyang Li[5], Xiaoke Liu[6], Suzette J. Bielinski[2], Jun Chen[2], Yongbin Chen[7] & James R. Cerhan[2]

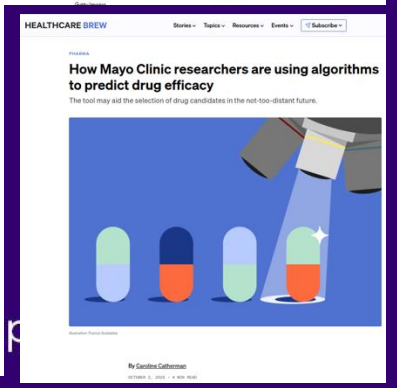
The complexities inherent in EHR data create discrepancies between real-world evidence and RCTs, posing substantial challenges in determining whether a treatment is likely to have a beneficial impact compared to the standard of care in RCTs. The objective of this study is to enhance the prediction of efficacy direction for repurposed drugs tested in RCTs for heart failure (HF). To achieve this, we propose an efficacy direction prediction framework that integrates drug-target predictions with EHR-based Emulation Trials (ET) to derive surrogate endpoints for prediction using HF prognostic markers. Our validation of the proposed novel drug-target prediction model against the BETA benchmark demonstrates superior performance, surpassing existing baseline algorithms. Furthermore, an evaluation of our framework in identifying 17 repurposed drugs—derived from 266 phase 3 HF RCTs—using data from 59,000 patients at the Mayo Clinic highlights its remarkable predictive accuracy.



'Virtual clinical trials' may predict success of heart failure drugs

Marla Broadfoot, Ph.D.

September 16, 2025



HEALTHCARE BREW

How Mayo Clinic researchers are using algorithms to predict drug efficacy

The tool may aid the selection of drug candidates in the not-too-distant future.

npj | health systems

Explore content · About the journal · Publish with us

nature › npj health systems › articles › article

npj Health Systems has APC waivers available that can be allocated upon acceptance on an ad-hc
Chen.

Article | Open access | Published: 01 April 2025

Synoptic reporting by summarizing cancer pathology reports using large language models

Sivaraman Rajaganapathy, Shaika Chowdhury, Xiaodi Li, Vincent Buchner, Zhe He, Rui Zhang, Xiaoqian Jiang, Ping Yang, James R. Cerhan & Nansu Zong

npj Health Systems 2, Article number: 11 (2025) | Cite this article

medRxiv

THE PREPRINT SERVER FOR HEALTH SCIENCES

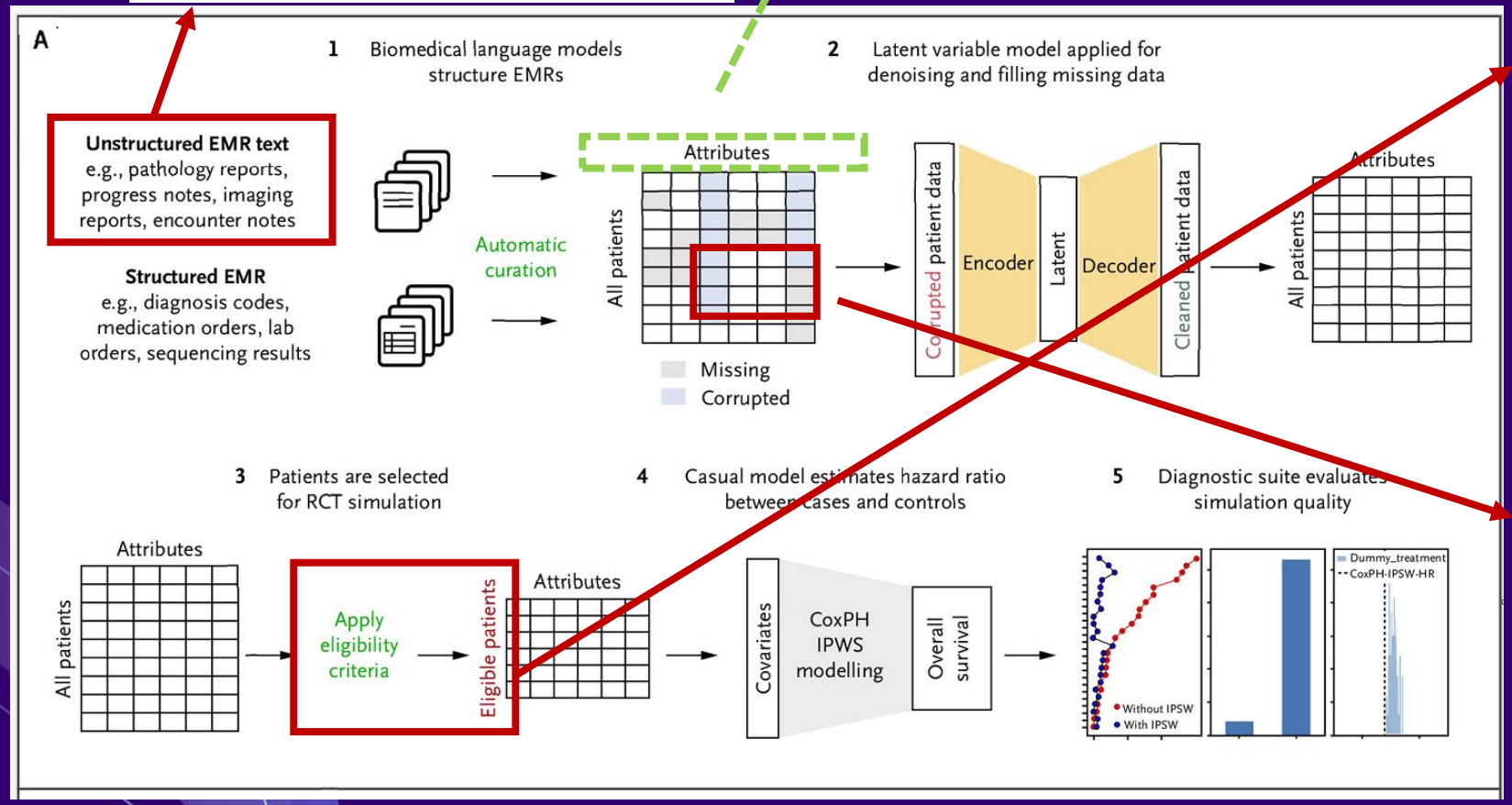
Accepted by AMIA 2026

Follow this p

AERO:An AI Agent for Adaptive Eligibility Refinement and Optimization of Clinical Trial Criteria in Real-World Trial Emulation

Xiaodi Li, Jose James, Patricia Pellikka, Nansu Zong

doi: https://doi.org/10.64898/2026.04.30.26352142



arXiv

Computer Science > Computation and Language

[Submitted on 17 Mar 2025 (v1), last revised 24 Mar 2025 (this version, v3)]

LLM-Match: An Open-Sourced Patient Matching Model Based on Large Language Models and Retrieval-Augmented Generation

Xiaodi Li, Shaika Chowdhury, Chung Il Wi, Maria Vassilaki, Xiaoke Liu, Terence T Sio, Owen Garrick, Young J Juhn, James R Cerhan, Cui Tao, Nansu Zong

Accepted by IEEE ICHI 2026

npj | health systems

Explore content · About the journal · Publish with us

nature › npj health systems › articles › article

npj Health Systems has APC waivers available that can be allocated upon acceptance on an ad-hoc
Chen.

Article | Open access | Published: 27 May 2026

From "negative" trial to positive clinical impact: mitigating eligibility criteria–induced temporal selection bias in emulated clinical trials

Xiaodi Li, Sivaraman Rajaganapathy, Xinyue Hu, MunHwan Lee, Jingna Feng, Jianfu Li, Yue Yu, Phil Fiero, Soulmaz Boroumand, Richard Larsen, Jun Chen, Pengyang Li, Jose James, Xiaoke Liu, Cui Tao & Nansu Zong

González, J., Ueno, R., Wong, C., Gero, Z., Bagga, J., Chien, I., Oravkin, E., Kiciman, E., Nori, A., Weerasinghe, R. and Leidner, R.S., 2025. TRIALSCOPE—A framework for clinical trial simulation from real-world data. NEJM AI, 2(10), p.AIoa2400859.

Challenge 1: Patient-Eligibility Matching

Inclusion Criteria

- Cardiac EF <=35% by radionuclide ventriculography, left ventriculography or quantitive echocardiographic measurement or an echocardiographic Wall Motion Index of <=1.2, within three months of enrollment. The patient's clinical cardiac state at enrollment should be similar to their state at the time of the qualifying echocardiogram. The qualifying left ventricular function measurement must be obtained at least three months after an MI, coronary bypass grafting, PTCA, and at least one month after pacemaker insertion. Patients scheduled for mitral valve repair should have qualifying echo after surgery.
- Modified Rankin score <=4.
- Patient must be taking ACE inhibitors. If intolerant of ACE inhibitor, patient must be on angiotensin II receptor blockers or hydralazine and nitrates.
- Patient is able to follow an outpatient protocol (requiring monthly blood tests and clinic visits every four months for the duration of the study) and is available by telephone.
- Patient understands the purpose and requirements of the study, can make him/herself understood, and has provided informed consent.
- Patients with recent stroke or TIA within twelve (12) months will be eligible to be included in the recent stroke (RS) subgroup.
- Chronic CHF patients (NYHA I * IV) admitted to the hospital can be randomized prior to discharge if the patient is stable, taking oral medications for 24 hours and ambulatory at the time of discharge. Stable New York Heart Association Class IV patients will be eligible for randomization.

Exclusion Criteria

- The presence of any of the following unequivocal cardiac sources of embolism: chronic or paroxysmal AF, mechanical valve, endocarditis, intracardiac mobile or pedunculated thrombus, and valvular vegetation.
- Cyanotic congenital heart disease, Eisenmenger's syndrome.
- Decompensated heart failure.
- Cardiac surgery, angioplasty, or MI within the past 3 months prior to randomization.
- A contraindication to the use of either warfarin or aspirin, e.g. active peptic ulcer disease, active bleeding diathesis, platelets <100,000*, hematocrit <30, INR >1.3 (if not on warfarin), clotting factor abnormality that increases the risk of bleeding, alcohol or substance abuse, severe gait instability, cerebral hemorrhage, systemic hemorrhage within the past year, severe liver impairment (AST >3x normal*, cirrhosis), any condition requiring regular use of non-steroidal anti-inflammatory agents, allergy to aspirin or warfarin, uncontrolled severe hypertension (systolic pressure >180 mm Hg or diastolic pressure > 110 mm Hg), positive stool guaiac not attributable to hemorrhoids, creatinine >3.0*. *on most recent test done within 30 days prior to randomization
- Patient needs continuing therapy with intravenous heparin or low molecular weight heparin or a specific antiplatelet agent.
- Dementia or psychiatric or physical problem that prevents the patient from following an outpatient program reliably.
- Comorbid conditions that may limit survival to less than five years.
- Pregnancy, or female of childbearing potential who is not sterilized or is not using a medically accepted form of contraception* (see procedure manual). *A pregnancy test is required for all women of childbearing age.
- Enrollment in another study that would conflict with WARCEF.
- Hospitalization for new diagnosis of onset CHF within the past one month or carotid endarterectomy or pacemaker insertion within the past one month prior to randomization .
- Person under 18 years of age.

Searching method to find answers

Trial	NCT02008357
Drug	Solanezumab
Inclusion Criteria	Searching method to find answers
Has a Mini-Mental State Examination (MMSE) score at screening of 25 to 30	NLP in the cognitive test
Has a global Clinical Dementia Rating (CDR) scale score at screening of 0	NLP in the cognitive test
Has a Logical Memory (LM) II score at screening of 6 to 18	NLP in the cognitive test
Has a florbetapir positron emission tomography (PET) scan that shows evidence of brain amyloid pathology at screening	NLP in searching brain amyloid pathology in PET reports
Exclusion Criteria:	
Is receiving a prescription acetylcholinesterase inhibitor (AChEI) and/or memantine at screening or baseline	NLP in medication names
Has current serious or unstable illness including cardiovascular, hepatic, renal, gastroenterologic, respiratory, endocrinologic, neurologic, psychiatric, immunologic, or hematologic disease or other conditions that, in the investigator's opinion, could interfere with the analyses of safety and efficacy in this study	Hospitalization in the previous 6 months
Has had a history within the last 5 years of a serious infectious disease affecting the brain (including neurosyphilis, meningitis, or encephalitis) or head trauma resulting in protracted loss of consciousness	ICD 9/10 codes in diagnosis
Has had a history within the last 5 years of a primary or recurrent malignant disease with the exception of any in situ cancer that was appropriately treated and is being appropriately monitored, such as resected cutaneous squamous cell carcinoma in situ or in situ prostate cancer with normal prostate-specific antigen post-treatment	ICD 9/10 codes in diagnosis
Has a known history of human immunodeficiency virus (HIV), clinically significant multiple or severe drug allergies, or severe post-treatment hypersensitivity reactions (including, but not limited to, erythema multiforme major, linear immunoglobulin A dermatosis, toxic epidermal necrolysis, or exfoliative dermatitis)	ICD 9/10 codes in diagnosis
Is clinically judged by the investigator to be at serious risk for suicide	NLP and ICD 9/10 codes in diagnosis
Has a history within the past 2 years of major depression or bipolar disorder as defined by the most current version of the Diagnostic and Statistical Manual of Mental Disorders (DSM)	NLP and ICD 9/10 codes in diagnosis
Has a history within the past 5 years of chronic alcohol or drug abuse/dependence as defined by the most current version of the DSM	ICD 9/10 codes in diagnosis

Searching & Matching



Eligibility Criteria

Text Processing
(understand Question)

Pre-processing rules
(where to find)



Generalizable, Adaptable

EHR





Medical notes

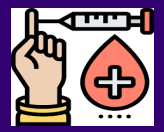




Comorbidities



Demographics

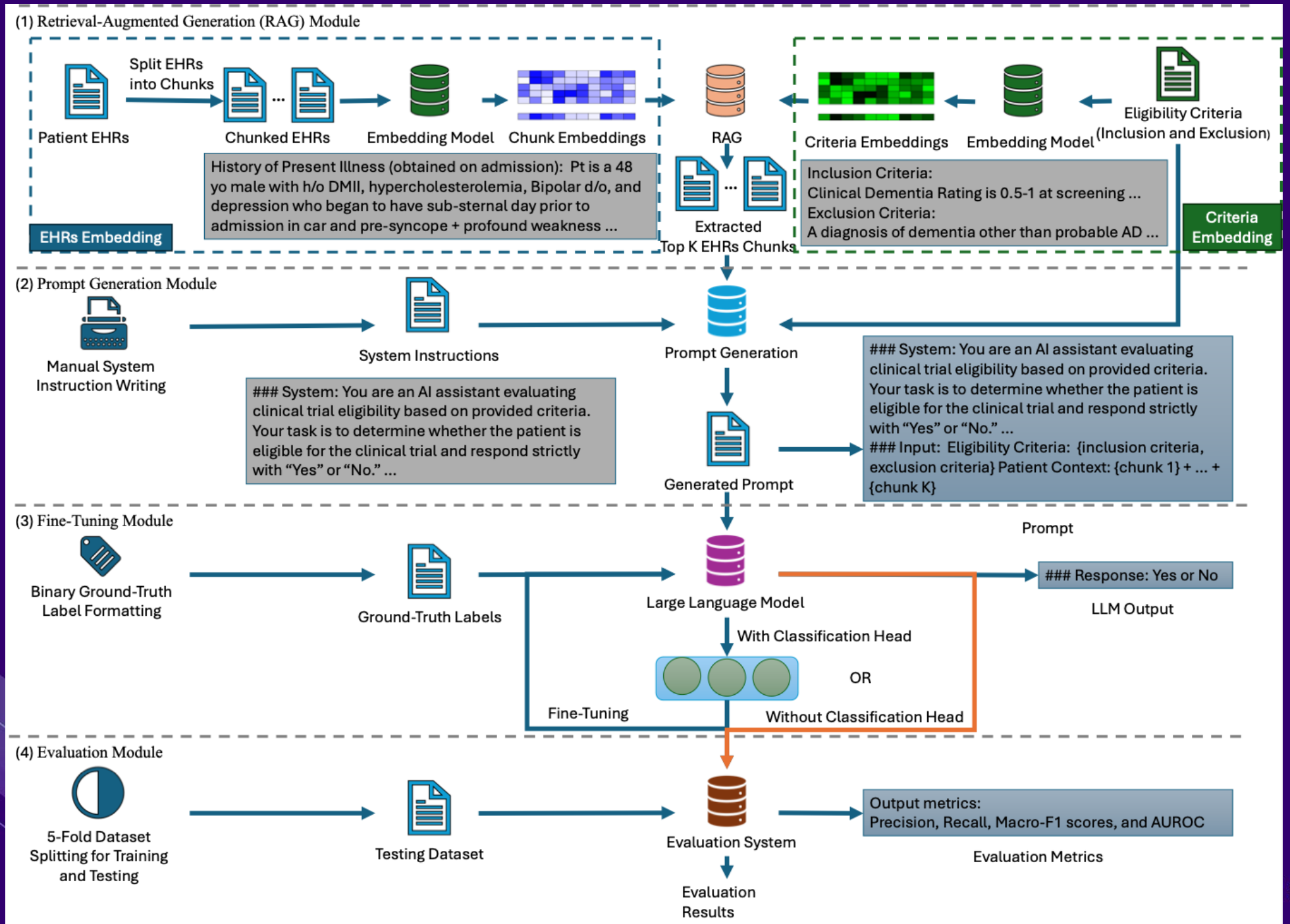


Lab test

Searching (how to find)



Scoring
(how to evaluate partial matching)





Lead by:
Dr. Xiaodi Li
IEEE ICHI 2027

TABLE II. DATASETS STATISTICS.

Dataset	# Patients	# Trials	Patient-Trial Pairs	Method Compared
N2C2	288	13	3744	Zero-Shot [1]
SIGIR	58	2991	3141	TrialGPT [2]
TREC 2021	75	13142	15437	
TREC 2022	50	11604	12682	

[1] M. Wornow, A. Lozano, D. Dash, J. Jindal, K. W. Mahaffey, and N. H. Shah, "Zero-shot clinical trial patient matching with llms," NEJM AI, vol. 2, no. 1, pp. AIcs2400360, 2025.

[2] Q. Jin, Z. Wang, C. S. Floudas, F. Chen, C. Gong, D. Bracken-Clarke et al., "Matching patients to clinical trials with large language models," Nature communications, vol. 15, no. 1, pp. 9074, 2024.

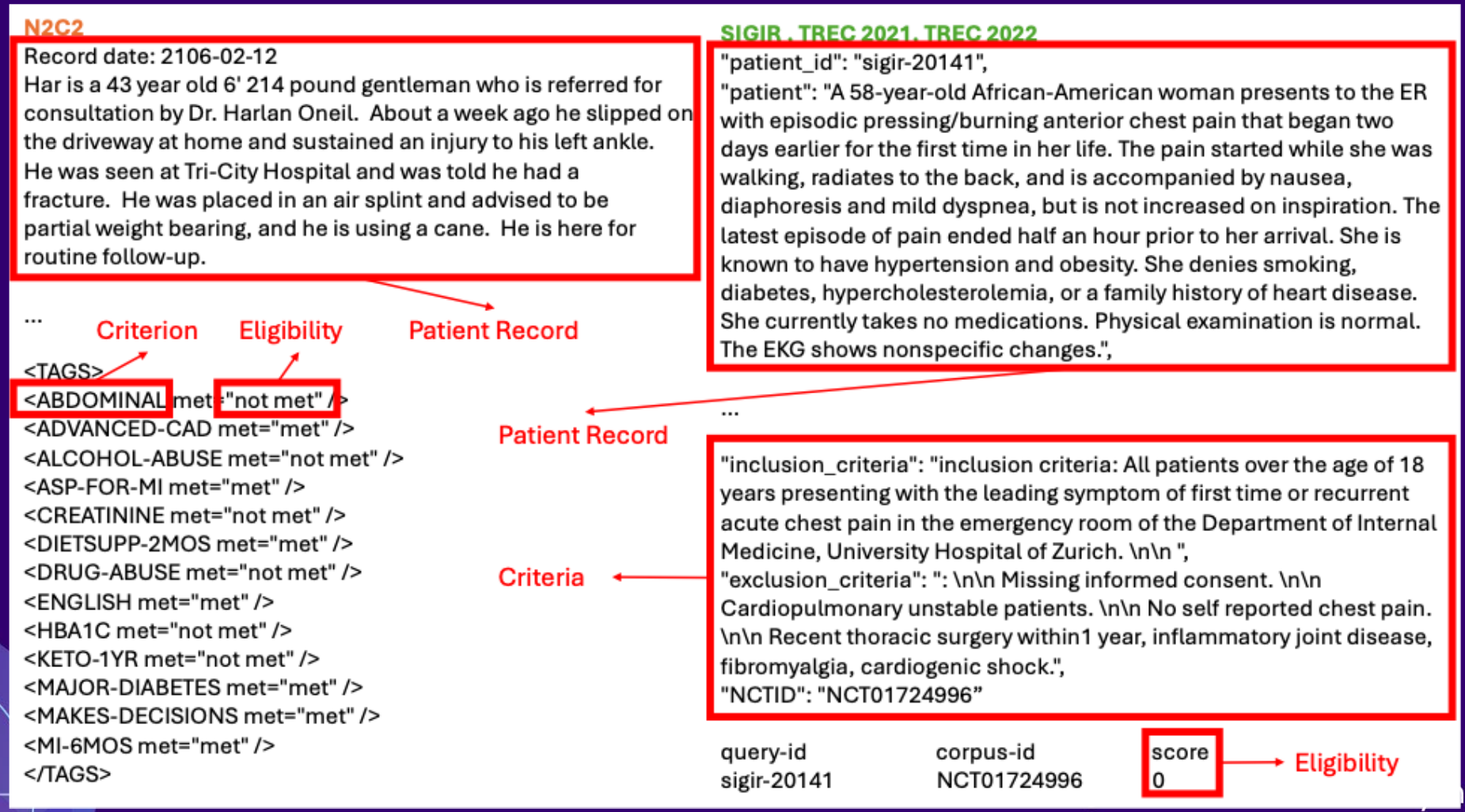


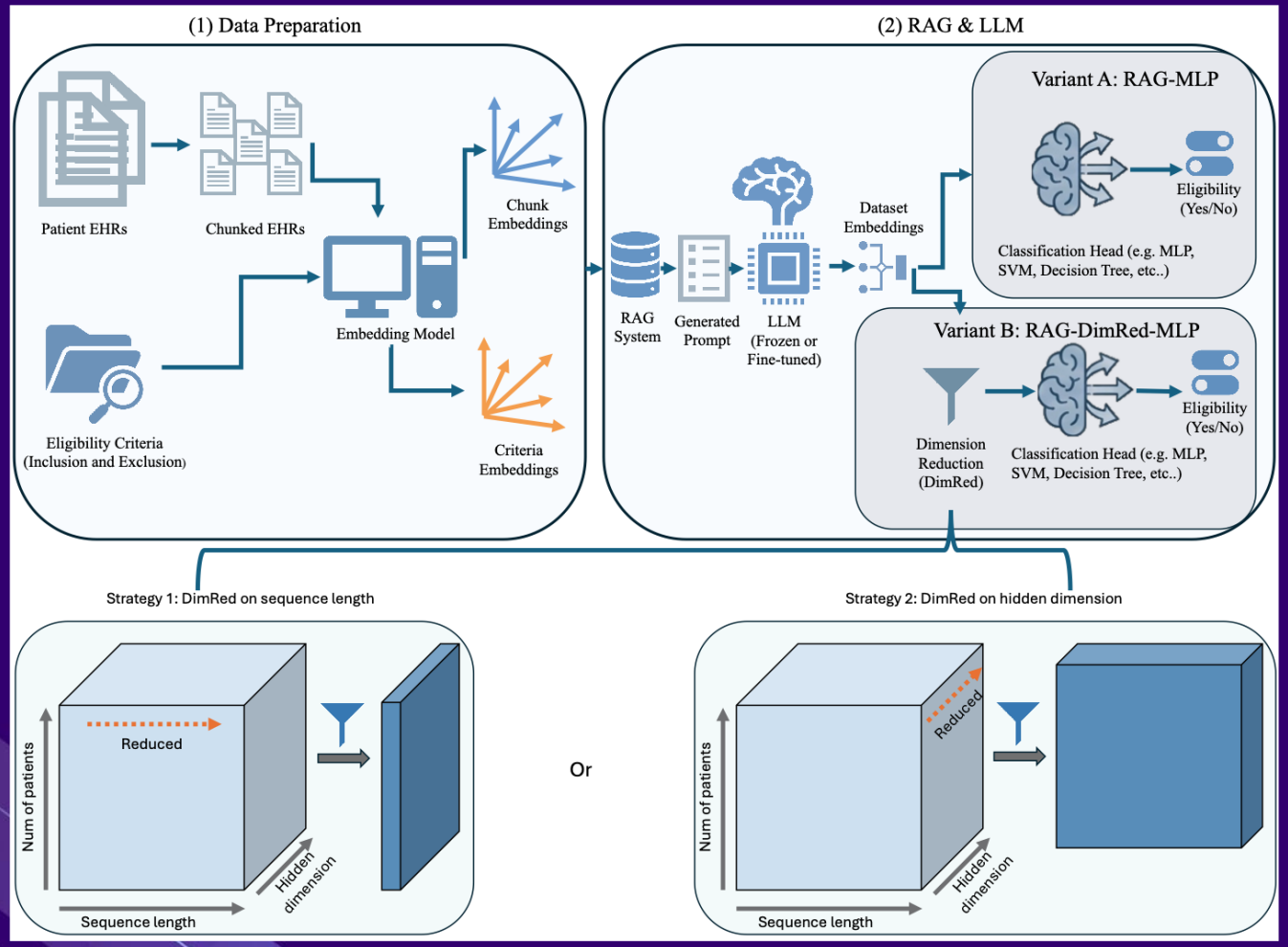
TABLE III. EXPERIMENTAL RESULTS ON THE N2C2 2018 COHORT SELECTION CHALLENGE COMPARED WITH ZERO-SHOT. STANDARD DEVIATIONS ARE REPORTED ACROSS FIVE RUNS.

Model	Precision	Recall	Overall Macro-F1
Zero-Shot (Llama-2-70b)*	0.82 ± 0.0255	0.41 ± 0.0255	0.46 ± 0.0102
Zero-Shot (Mixtral-8x7B)*	0.72 ± 0.0179	0.83 ± 0.0179	0.64 ± 0.0204
Zero-Shot (Llama-3-70b)*	0.76 ± 0.0179	0.88 ± 0.0153	0.68 ± 0.0179
Zero-Shot (Qwen-2-72b)*	0.80 ± 0.0179	0.94 ± 0.0102	0.74 ± 0.0230
Zero-Shot (GPT-3.5)*	0.74 ± 0.0179	0.80 ± 0.0179	0.59 ± 0.0230
Zero-Shot (GPT-4)*	0.91 ± 0.0128	0.92 ± 0.0128	0.81 ± 0.0179
LLM-Match (Llama-3-8B-Instruct)	0.85 ± 0.0162	0.80 ± 0.0147	0.85 ± 0.0044
LLM-Match with classification head (Llama-3-8B-Instruct)	0.84 ± 0.0073	0.83 ± 0.0079	**0.86** ± 0.0035
LLM-Match (Falcon3-10B-Instruct)	0.83 ± 0.0092	0.82 ± 0.0057	**0.86** ± 0.0038
LLM-Match with classification head (Falcon3-10B-Instruct)	0.83 ± 0.0117	0.83 ± 0.0138	0.85 ± 0.0068
LLM-Match (DeepSeek-R1-Distill-Llama-8B)	0.82 ± 0.0202	0.73 ± 0.0571	0.81 ± 0.0212
LLM-Match with classification head (DeepSeek-R1-Distill-Llama-8B)	0.84 ± 0.0124	0.82 ± 0.0114	**0.86** ± 0.0070

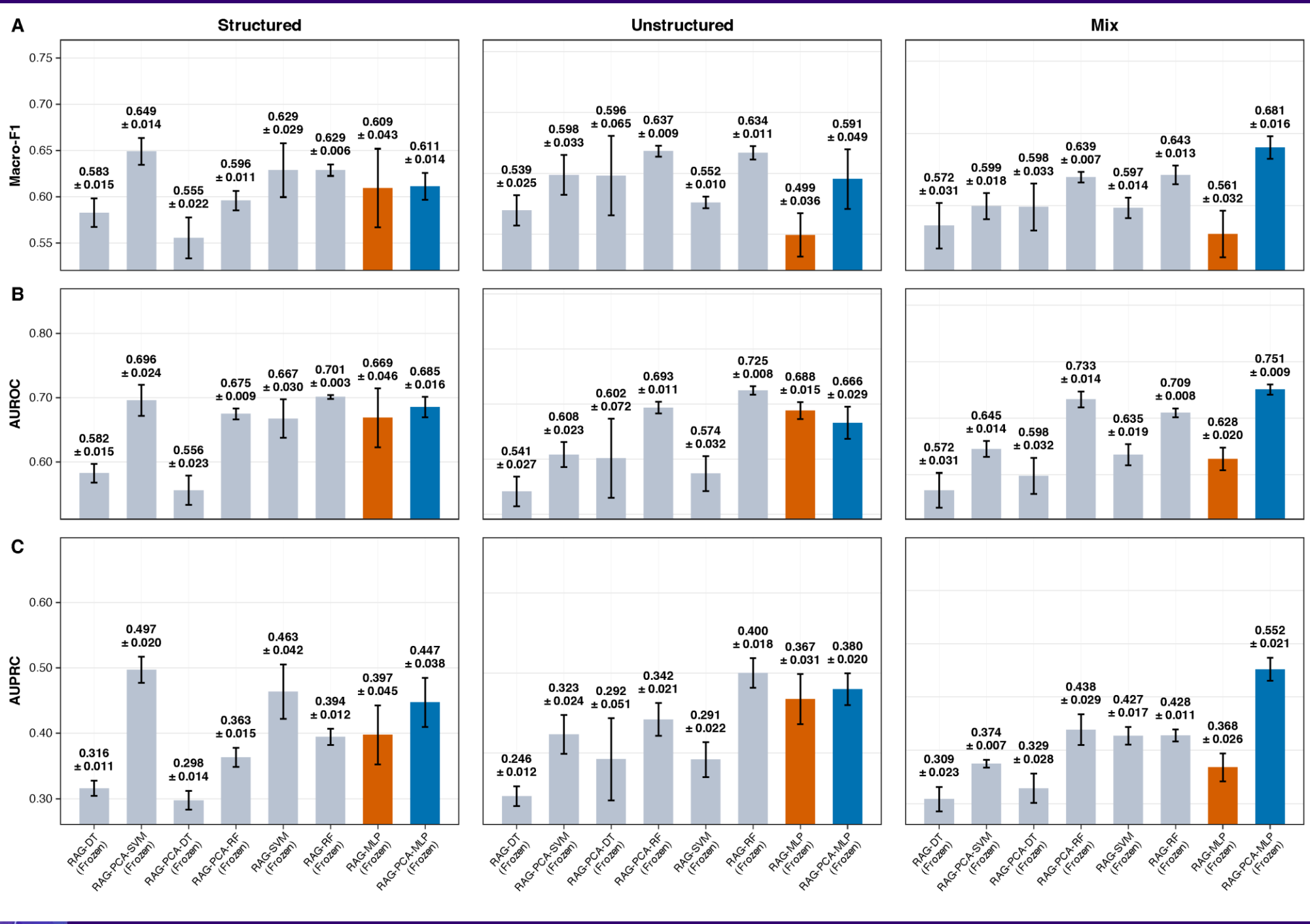
TABLE IV. EXPERIMENTAL RESULTS ON THE SIGIR, TREC 2021, AND TREC 2022 DATASETS (AVERAGE RESULTS ACROSS THE THREE DATASETS) COMPARED WITH TRIALGPT. STANDARD DEVIATIONS ARE REPORTED ACROSS ALL THREE DATASETS.

Model	AUROC
SciFive (encoder-decoder)*	0.5895
BioBERT (dual-encoder) *	0.5952
PubMedBERT (dual-encoder) *	0.5976
SapBERT (dual-encoder) *	0.5933
BioLinkBERT (cross-encoder) *	0.6176
Trial-LLAMA 70B §	0.6528
TrialGPT-Ranking (GPT-3.5) *	0.6582
TrialGPT-Ranking (GPT-4) *	0.7979
LLM-Match (Llama-3-8B-Instruct)	0.7736 ± 0.0145
LLM-Match with classification head (Llama-3-8B-Instruct)	0.8010 ± 0.0157
LLM-Match (Falcon3-10B-Instruct)	0.7766 ± 0.0326
LLM-Match with classification head (Falcon3-10B-Instruct)	0.7908 ± 0.0190
LLM-Match (DeepSeek-R1-Distill-Llama-8B)	0.7605 ± 0.0372
LLM-Match with classification head (DeepSeek-R1-Distill-Llama-8B)	**0.8157** ± 0.0004

1. Performance on Real Data, 2. More Portable and Feasible



Drugs	Solanezumab		Lecanemab		RVT-101		Insulin		CAD106, CNP520	
NCT number	NCT02008357		NCT04468659		NCT02669433		NCT01767909		NCT02565511	
Screened patients	N=114		N=158		N=23		N=30		N=39	
Inclusion Criteria:	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N
Demographic	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Cognitive test	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
MRI/PET marker	0.0%	36.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Family history	0.0%	0.0%	55.0%	38.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gene test	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Medication	0.0%	0.0%	0.0%	0.0%	66.7%	91.3%	0.0%	0.0%	0.0%	0.0%
Exclusion Criteria:										
Diagnosis	37.3%	96.5%	37.9%	91.8%	52.6%	82.6%	53.3%	100.0%	26.9%	66.7%
Medication	0.9%	98.2%	0.0%	0.0%	0.0%	0.0%	40.0%	100.0%	0.0%	100.0%
Surgical	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Hospitalization	0.9%	100.0%	2.5%	100.0%	0.0%	100.0%	0.0%	100.0%	0.0%	100.0%
Psychiatric/psychological	5.5%	96.5%	9.0%	91.8%	0.0%	0.0%	13.3%	100.0%	0.0%	0.0%
Life habit	0.0%	96.5%	0.7%	91.8%	0.0%	0.0%	0.0%	0.0%	0.0%	66.7%
Allergy/Hypersensitivity	0.0%	0.0%	0.0%	53.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%


A
Structured
Unstructured
Mix
Macro-F1
0.75
0.70
0.65
0.60
0.55
0.583 ± 0.015
0.649 ± 0.014
0.555 ± 0.022
0.596 ± 0.011
0.629 ± 0.029
0.629 ± 0.006
0.609 ± 0.043
0.611 ± 0.014
0.539 ± 0.025
0.598 ± 0.033
0.596 ± 0.065
0.637 ± 0.009
0.552 ± 0.010
0.634 ± 0.011
0.499 ± 0.036
0.591 ± 0.049
0.572 ± 0.031
0.599 ± 0.018
0.598 ± 0.033
0.639 ± 0.007
0.597 ± 0.014
0.643 ± 0.013
0.561 ± 0.032
0.681 ± 0.016
B
AUROC
0.80
0.70
0.60
0.582 ± 0.015
0.696 ± 0.024
0.556 ± 0.023
0.675 ± 0.009
0.667 ± 0.030
0.701 ± 0.003
0.669 ± 0.046
0.685 ± 0.016
0.541 ± 0.027
0.608 ± 0.023
0.602 ± 0.072
0.693 ± 0.011
0.574 ± 0.032
0.725 ± 0.008
0.688 ± 0.015
0.666 ± 0.029
0.572 ± 0.031
0.645 ± 0.014
0.598 ± 0.032
0.733 ± 0.014
0.635 ± 0.019
0.709 ± 0.008
0.628 ± 0.020
0.751 ± 0.009
C
AUPRC
0.60
0.50
0.40
0.30
0.316 ± 0.011
0.497 ± 0.020
0.298 ± 0.014
0.363 ± 0.015
0.463 ± 0.042
0.394 ± 0.012
0.397 ± 0.045
0.447 ± 0.038
0.246 ± 0.012
0.323 ± 0.024
0.292 ± 0.051
0.342 ± 0.021
0.291 ± 0.022
0.400 ± 0.018
0.367 ± 0.031
0.380 ± 0.020
0.309 ± 0.023
0.374 ± 0.007
0.329 ± 0.028
0.438 ± 0.029
0.427 ± 0.017
0.428 ± 0.011
0.368 ± 0.026
0.552 ± 0.021
RAG-DT (Frozen)
RAG-PCA-SVM (Frozen)
RAG-PCA-DT (Frozen)
RAG-PCA-RF (Frozen)
RAG-SVM (Frozen)
RAG-RF (Frozen)
RAG-MLP (Frozen)
RAG-PCA-MLP (Frozen)

Challenge 2: Data Extraction

Pembrolizumab/Vibostolimab (MK-7684A) or Atezolizumab in Combination With Chemotherapy in First Line Treatment of Extensive-Stage Small Cell Lung Cancer (MK-7684A-008, KEYVIBE-008)

ClinicalTrials.gov ID NCT05224141

Sponsor Merck Sharp & Dohme LLC

Information provided by Merck Sharp & Dohme LLC (Responsible Party)

Last Update Posted 2024-05-08

Pathology

Diagnosis

Medication

Inclusion Criteria:

- Has histologically or cytologically confirmed diagnosis of ES-SCLC in need of first-line therapy
- Has ES-SCLC defined as Stage IV (T any, N any, M1a/b/c) by the American Joint Committee on Cancer, Eighth Edition or T3-T4 due to multiple lung nodules that are too extensive or have tumor/nodal volume that is too large to be encompassed in a tolerable radiation plan

Exclusion Criteria:

- Is considered a poor medical risk due to a serious, uncontrolled medical disorder or non-malignant systemic disease
- Has received prior treatment for Small Cell Lung Cancer (SCLC)
- Is expected to require any other form of antineoplastic therapy for SCLC while on study

Can SQL solve the problem?

Not available

You Need Better Tool

Table A.1. Unsearchable criteria (white) in five trials. Pos= total identified patients with a value for all the questions answered, Idf= total Identified patients with this information source.

Drugs	Solanezumab		Lecanemab		RVT-101		Insulin		CAD106, CNP520	
NCT number	NCT02008357		NCT04468659		NCT02669433		NCT01767909		NCT02565511	
Screened patients	N=114		N=158		N=23		N=30		N=39	
Inclusion Criteria:	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N	Pos/Idf	Idf/N
Demographic	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Cognitive test	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
MRI/PET marker	0.0%	36.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Family history	0.0%	0.0%	55.0%	38.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Gene test	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Medication	0.0%	0.0%	0.0%	0.0%	66.7%	91.3%	0.0%	0.0%	0.0%	0.0%
Exclusion Criteria:										
Diagnosis	37.3%	96.5%	37.9%	91.8%	52.6%	82.6%	53.3%	100.0%	26.9%	66.7%
Medication	0.9%	98.2%	0.0%	0.0%	0.0%	0.0%	40.0%	100.0%	0.0%	100.0%
Surgical	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Hospitalization	0.9%	100.0%	2.5%	100.0%	0.0%	100.0%	0.0%	100.0%	0.0%	100.0%
Psychiatric/psychological	5.5%	96.5%	9.0%	91.8%	0.0%	0.0%	13.3%	100.0%	0.0%	0.0%
Life habit	0.0%	96.5%	0.7%	91.8%	0.0%	0.0%	0.0%	0.0%	0.0%	66.7%
Allergy/Hypersensitivity	0.0%	0.0%	0.0%	53.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Synoptic Report vs Pathology report

COLLEGE of AMERICAN PATHOLOGISTS

Definition of Synoptic Reporting

Synoptic reporting in surgical pathology is a style of reporting that has advantages for a variety of users of surgical pathology reports.[1-3] For pathologists, synoptic reporting can improve the completeness, accuracy, and ease of creating the report.[4-12] For clinicians, synoptic reports can make data extraction from the report both more rapid and more accurate.[13-15] For researchers and cancer registrars, synoptic reporting also ensures that these data elements are amenable to scalable data capture, interoperability, and exchange, enabling the creation of structured data sets to facilitate research.

In order to help pathologists achieve these goals, the CAP has developed a list of specific features that define *synoptic* report formatting for accreditation compliance. These include:

1. All required data elements outlined on the currently applicable surgical case summary from the cancer protocol that are included in the report must be displayed in synoptic format
   - Synoptic reporting is defined by the data element followed by its answer (response), e.g., "Tumor size: 5.5 cm." Outline format without the paired "data element: response" format is not considered synoptic.
   - The data element does not have to be identical (i.e., verbatim) to that listed in the CAP protocol and may be rephrased (e.g., for conciseness) as long as the intended meaning remains clear.
   - Multiple related elements can be combined into a single data entry, as long as the individual responses can be distinguished by the reader and as long as the intended meaning remains clear. Examples include but are not limited to:
     - Anatomic site or specimen, laterality, and procedure
     - Pathology Staging Tumor Node Metastasis (pTNM) staging elements
     - Negative margins, as long as all negative margins are specifically enumerated where applicable
     - Tumor type and grade
     - All parts of grade (e.g. “Gleason grade: 3+4 = 7 (Group 3)”)
     - Breast tubule formation, nuclear pleomorphism, and mitotic rate
     - All portions of an ancillary study result (e.g. “Estrogen receptor: Positive, 100% of cells, strong”)
     - Positive cores/total cores
     - Positive lymph nodes/total lymph nodes
     - Size (when giving more than one dimension)

COLLEGE of AMERICAN PATHOLOGISTS

Member Resources | Advocacy | Laboratory Improvement | Education

Home > Protocols and Guidelines > Cancer Protocol Templates

Cancer Protocol Templates

Laboratories across the country and around the world rely on the CAP to help them stay current with the latest evidence-based practices and maintain quality and proficiency among the entire pathology team.

The CAP Cancer Protocols provide guidelines for collecting the essential data elements for complete reporting of malignant tumors and optimal patient care. These protocols incorporate the latest standards to help pathologists and laboratories keep abreast of the advances and updates in cancer reporting, including:

- The World Health Organization Classification of Tumors (Blue Books)
- The American Joint Committee on Cancer Staging Manual

Synoptic reporting using the **CAP electronic Cancer Protocols (eCP)** also helps laboratories and hospitals comply with requirements for organizations like the American College of Surgeons Commission on Cancer, the CAP Laboratory Accreditation Program, and more.

The CAP Biomarker Reporting Protocols are intended to provide reporting guidance for commonly ordered biomarkers and are not required for accreditation purposes.

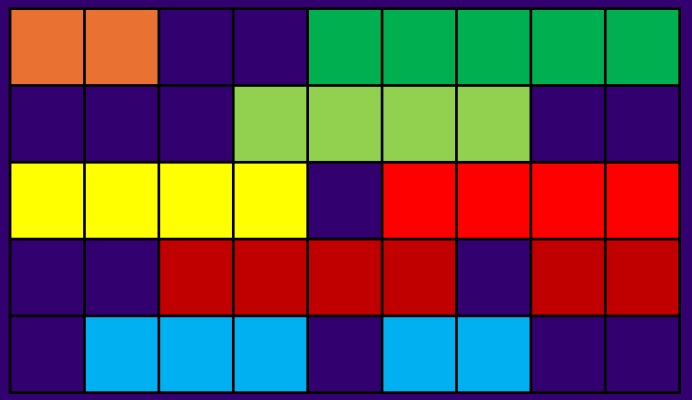
```
"name" : "DIAGNOSIS",
"body" : "A.  Rectum, wall, mass, transanal excision:  Invasive moderately
  differentiated adenocarcinoma, forming a 1.4 x 1.2 x 0.7 cm mass in
  the rectum.  The surgical resection margins are negative for tumor."


"name" : "GROSS DESCRIPTION",
"body" : "A.  Received fresh labeled \"rectal wall mass\" is a 3.7 x 2.3 x 1.4
 cm transanal excision.  The specimen is inked, and the margins are
 submitted perpendicularly.  There is a 1.4 x 1.1 x 0.7 cm tan,
 indurated lesion which is 0.2 cm from the deep surgical margin.
 Representative sections submitted.  Grossed by DJS.
 BLOCK SUMMARY: \r\n  Part A:  Rectal wall mass
 1 Mass 1
 2 Mass 2
 3 Mass 3"
```

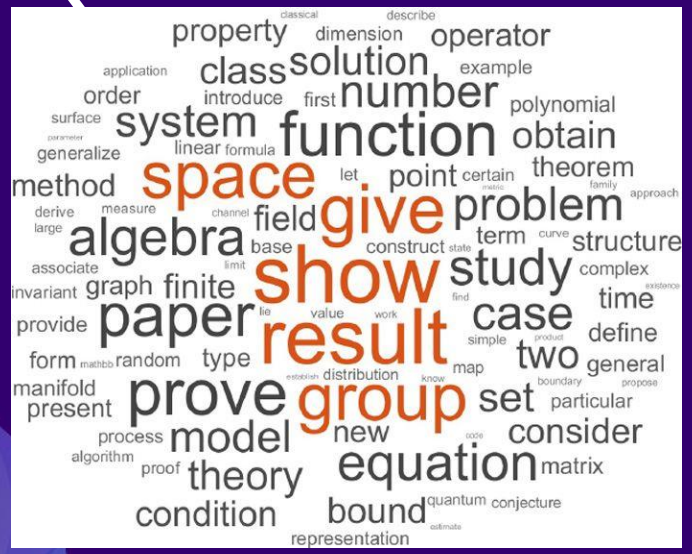
```
"name" : "SYNOPTIC REPORT",
"body" : "Specimen: Rectum.
Procedure: Transanal disk excision (local excision).
Tumor Site: Rectum.  \r\n  Tumor Size: 1.4 x 1.2 x 0.7 cm.
Macroscopic Tumor Perforation: Not identified.
Histologic Type: Adenocarcinoma.
Histologic Grade: Low-grade.
Microscopic Tumor Extension: Tumor invades muscularis propria.
Surgical Margins: All margins negative for tumor, 0.2 cm from: Radial.
Treatment Effect: Not applicable.
Lymphovascular Invasion: Not identified.
erineural Invasion: Not identified.
Tumor Deposits: Not applicable.
Primary tumor:  pT2
Regional lymph nodes: pNX
Number examined (total): 0
Number involved (total): 0
Distant Metastasis: Not applicable.
```

Can classification/Named Entity Recognition (NER) solve the problem?





TX
T0
T1
T2
T3
T4



Token (Terms)	Label
Lower anterior resection	Procedure
Invasive lobular carcinoma	Histologic Type

Problems





Text Processing



Pre-processing rules

Classification

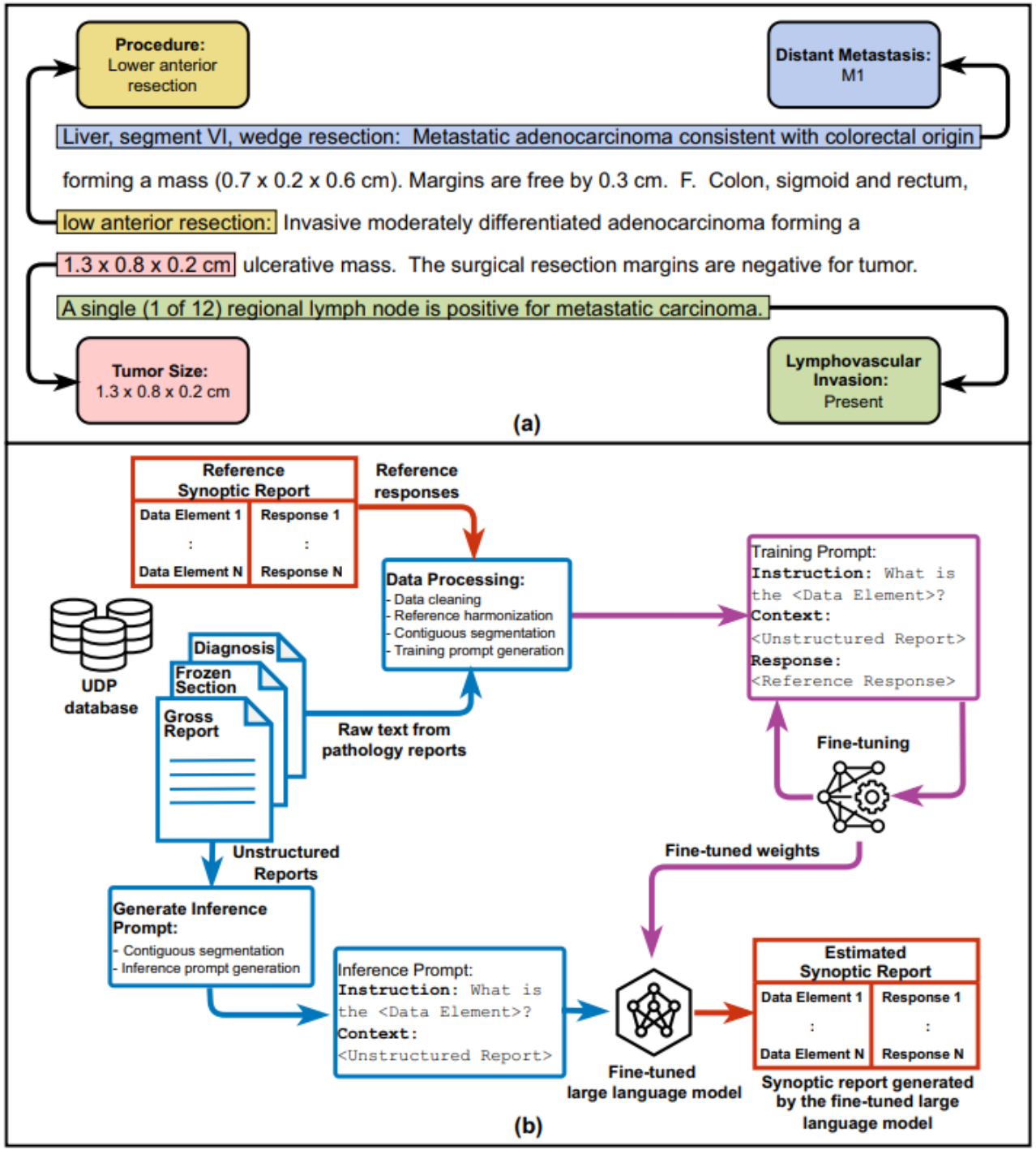
NER

Automated, Generalizable



Post--processing rules





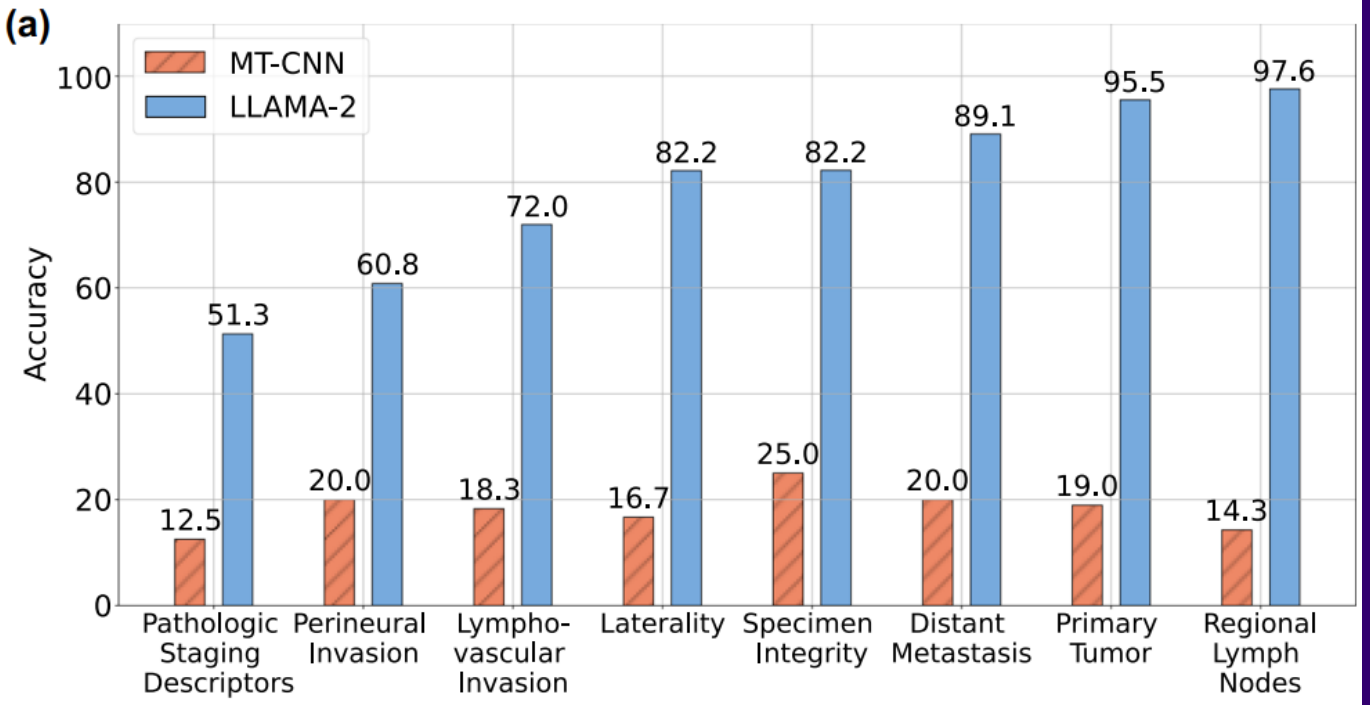


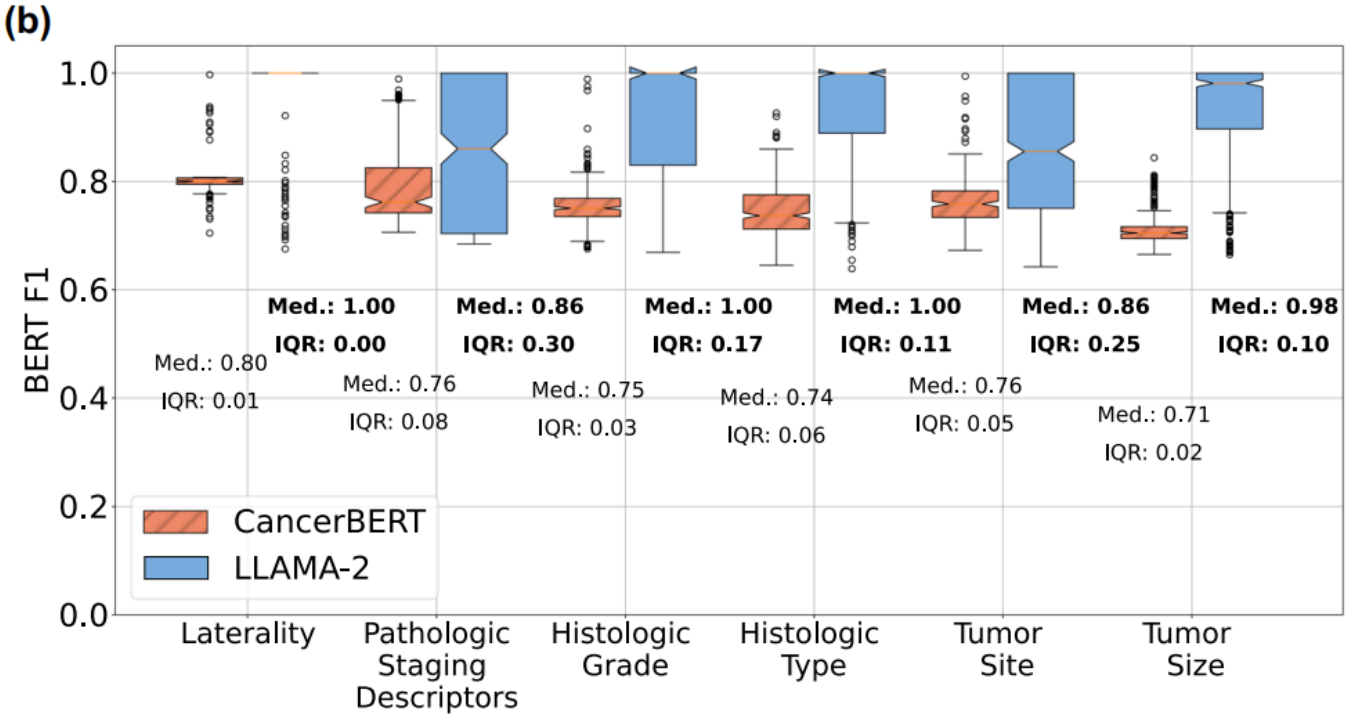
Lead by: Dr. Sivaraman Rajaganapathy
NPJ Health System, 2025

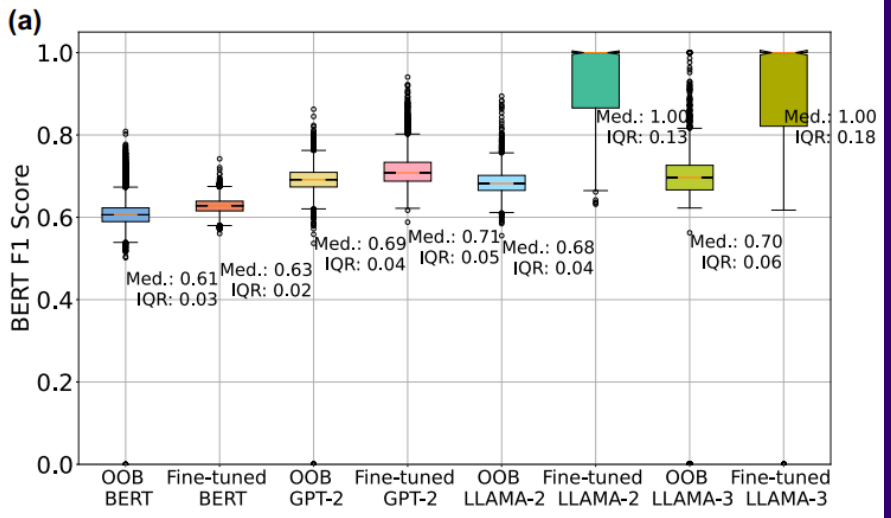
Table 1 | Summary of the 7774 Mayo Clinic cancer pathology reports dataset

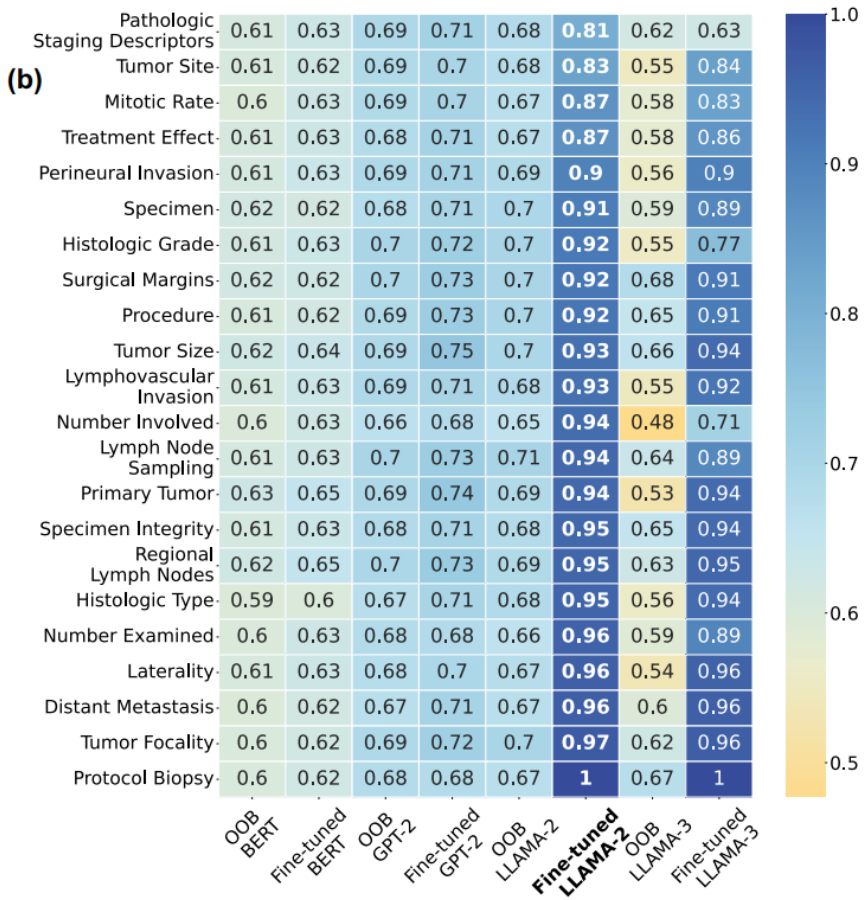
Data Element	% Non-missing	No. of Patients		Age in Years Mean [Min-Max]	Type of Cancer (%)							
		M	F		Breast	Lung	Pancreas	Skin	Digestive Organs	Female Reproductive Organs	Male Reproductive Organs	Other
Protocol Biopsy	25	1145	772	53.7 [19–85]	1	0	0	18	1	0	2	76
Procedure	70	2736	2682	60.6 [19–97]	13	6	3	2	13	8	16	38
Specimen	30	1096	1216	63.3 [21–97]	0	14	7	1	29	19	1	29
Primary Tumor	59	2401	2224	61.6 [19–97]	15	7	4	2	15	9	19	29
Specimen Integrity	17	530	801	63.3 [19–95]	0	23	0	5	4	30	3	34
Surgical Margins	44	1514	1898	62.1 [19–97]	20	10	5	3	20	7	2	33
Laterality	32	935	1545	58.0 [19–91]	29	13	0	1	1	0	1	54
Histologic Type	68	2674	2614	60.5 [19–97]	14	7	2	2	13	8	17	37
Histologic Grade	63	2525	2344	61.0 [19–97]	13	7	3	1	14	9	17	36
Mitotic Rate	13	209	793	59.3 [19–95]	59	1	4	7	3	0	0	26
Pathologic Staging Descriptors	30	1059	1295	60.9 [19–93]	25	6	5	2	14	8	15	25
Tumor Focality	27	873	1249	60.1 [19–92]	29	15	2	1	5	1	2	44
Tumor Site	43	1623	1698	60.5 [19–97]	7	10	5	3	19	8	1	47
Tumor Size	49	1584	2195	61.5 [19–97]	17	9	5	3	19	12	2	35
Lymphovascular Invasion	53	2226	1859	62.3 [19–97]	15	8	4	3	16	8	20	25
Regional Lymph Nodes	59	2376	2222	61.6 [19–97]	15	7	4	2	15	9	18	29
Lymph Node Sampling	30	917	1378	60.4 [19–93]	31	0	0	1	1	18	35	13
Number Examined	44	2270	1115	61.7 [19–97]	0	10	5	1	20	4	25	36
Number Involved	37	1957	924	62.1 [19–97]	0	10	6	1	22	3	26	32
Distant Metastasis	50	1940	1914	61.4 [19–97]	17	8	4	3	18	4	12	33
Perineural Invasion	16	764	462	63.1 [21–97]	0	2	13	4	47	0	1	32
Treatment Effect	29	1505	784	62.3 [19–95]	9	13	5	1	23	1	34	14

We show the top 22 most frequently reported data elements. The data elements listed do not occur in every report—the percent non-missing indicates the percentage of reports containing the corresponding data element.


(a)
MT-CNN
LLAMA-2
Accuracy
0
20
40
60
80
100
12.5
51.3
20.0
60.8
18.3
72.0
16.7
82.2
25.0
82.2
20.0
89.1
19.0
95.5
14.3
97.6
Pathologic Staging Descriptors
Perineural Invasion
Lympho-vascular Invasion
Laterality
Specimen Integrity
Distant Metastasis
Primary Tumor
Regional Lymph Nodes

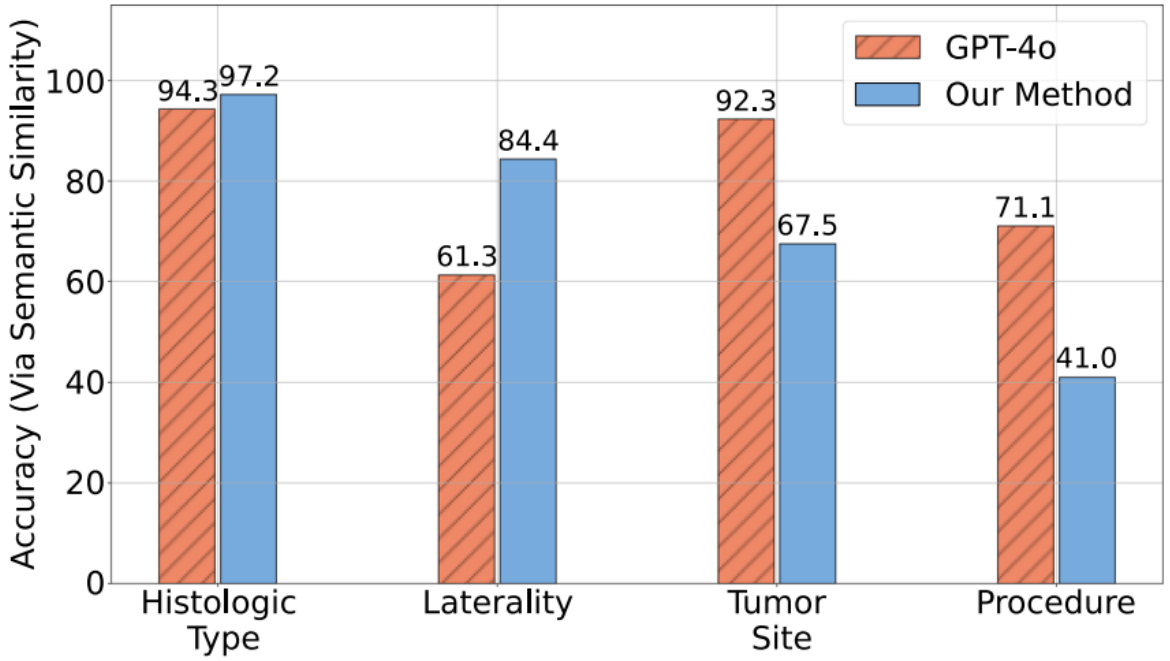

(b)
BERT F1
0.0
0.2
0.4
0.6
0.8
1.0
CancerBERT
LLAMA-2
Med.: 0.80
IQR: 0.01
Med.: 1.00
IQR: 0.00
Med.: 0.76
IQR: 0.08
Med.: 0.86
IQR: 0.30
Med.: 0.75
IQR: 0.03
Med.: 1.00
IQR: 0.17
Med.: 0.74
IQR: 0.06
Med.: 1.00
IQR: 0.11
Med.: 0.76
IQR: 0.05
Med.: 0.86
IQR: 0.25
Med.: 0.71
IQR: 0.02
Med.: 0.98
IQR: 0.10
Laterality
Pathologic Staging Descriptors
Histologic Grade
Histologic Type
Tumor Site
Tumor Size

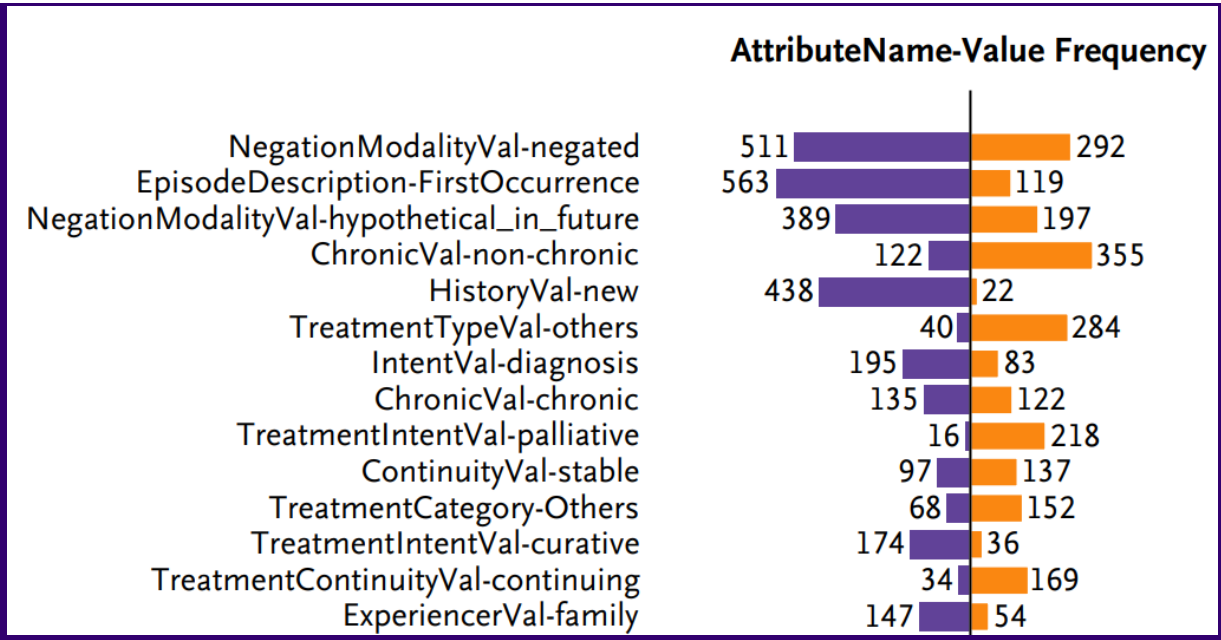

(a)
BERT F1 Score
0.0
0.2
0.4
0.6
0.8
1.0
Med.: 0.61
IQR: 0.03
Med.: 0.63
IQR: 0.02
Med.: 0.69
IQR: 0.04
Med.: 0.71
IQR: 0.05
Med.: 0.68
IQR: 0.04
Med.: 1.00
IQR: 0.13
Med.: 0.70
IQR: 0.06
Med.: 1.00
IQR: 0.18
OOB BERT
Fine-tuned BERT
OOB GPT-2
Fine-tuned GPT-2
OOB LLAMA-2
Fine-tuned LLAMA-2
OOB LLAMA-3
Fine-tuned LLAMA-3


(b)
Pathologic Staging Descriptors 0.61 0.63 0.69 0.71 0.68 0.81 0.62 0.63
Tumor Site 0.61 0.62 0.69 0.7 0.68 0.83 0.55 0.84
Mitotic Rate 0.6 0.63 0.69 0.7 0.67 0.87 0.58 0.83
Treatment Effect 0.61 0.63 0.68 0.71 0.67 0.87 0.58 0.86
Perineural Invasion 0.61 0.63 0.69 0.71 0.69 0.9 0.56 0.9
Specimen 0.62 0.62 0.68 0.71 0.7 0.91 0.59 0.89
Histologic Grade 0.61 0.63 0.7 0.72 0.7 0.92 0.55 0.77
Surgical Margins 0.62 0.62 0.7 0.73 0.7 0.92 0.68 0.91
Procedure 0.61 0.62 0.69 0.73 0.7 0.92 0.65 0.91
Tumor Size 0.62 0.64 0.69 0.75 0.7 0.93 0.66 0.94
Lymphovascular Invasion 0.61 0.63 0.69 0.71 0.68 0.93 0.55 0.92
Number Involved 0.6 0.63 0.66 0.68 0.65 0.94 0.48 0.71
Lymph Node Sampling 0.61 0.63 0.7 0.73 0.71 0.94 0.64 0.89
Primary Tumor 0.63 0.65 0.69 0.74 0.69 0.94 0.53 0.94
Specimen Integrity 0.61 0.63 0.68 0.71 0.68 0.95 0.65 0.94
Regional Lymph Nodes 0.62 0.65 0.7 0.73 0.69 0.95 0.63 0.95
Histologic Type 0.59 0.6 0.67 0.71 0.68 0.95 0.56 0.94
Number Examined 0.6 0.63 0.68 0.68 0.66 0.96 0.59 0.89
Laterality 0.61 0.63 0.68 0.7 0.67 0.96 0.54 0.96
Distant Metastasis 0.6 0.62 0.67 0.71 0.67 0.96 0.6 0.96
Tumor Focality 0.6 0.62 0.69 0.72 0.7 0.97 0.62 0.96
Protocol Biopsy 0.6 0.62 0.68 0.68 0.67 1 0.67 1
OOB BERT
Fine-tuned BERT
OOB GPT-2
Fine-tuned GPT-2
OOB LLAMA-2
Fine-tuned LLAMA-2
OOB LLAMA-3
Fine-tuned LLAMA-3
1.0
0.9
0.8
0.7
0.6
0.5

org

Fig. 6 | Benchmarking with external validation data. Performance of the OpenAI's GPT-4o and the fine-tuned LLAMA-2 7B model (best-performing model) on an external dataset in Sushil et al. [29]. We test the ability of the two models to extract 4 data elements which correspond to the labels provided in the external dataset. The accuracy reported here is obtained by manual comparison of the labels provided in ref. 29 against the model responses. *Note: Fine-tuned LLAMA-2 has only been fine-tuned on our data, it has not seen examples from the external data.*





Sushil, M. et al. CORAL: expert-curated medical oncology reports to advance language model inference. NEJM AI;1. https://doi.org/10. 1056/Aldbp2300110 (2024).

Research Question-Prediction of Success of Clinical Trial

npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01705-z

Advancing efficacy prediction for electronic health records based emulated trials in repurposing heart failure therapies

Check for updates

Nansu Zong[1,8] ✉, Shaika Chowdhury[1,8], Shibo Zhou[1,8], Sivaraman Rajaganapathy[1], Yue Yu[2], Liewei Wang[3], Qiying Dai[4], Pengyang Li[5], Xiaoke Liu[6], Suzette J. Bielinski[2], Jun Chen[2], Yongbin Chen[7] & James R. Cerhan[2]

The complexities inherent in EHR data create discrepancies between real-world evidence and RCTs, posing substantial challenges in determining whether a treatment is likely to have a beneficial impact compared to the standard of care in RCTs. The objective of this study is to enhance the prediction of efficacy direction for repurposed drugs tested in RCTs for heart failure (HF). To achieve this, we propose an efficacy direction prediction framework that integrates drug-target predictions with EHR-based Emulation Trials (ET) to derive surrogate endpoints for prediction using HF prognostic markers. Our validation of the proposed novel drug-target prediction model against the BETA benchmark demonstrates superior performance, surpassing existing baseline algorithms. Furthermore, an evaluation of our framework in identifying 17 repurposed drugs—derived from 266 phase 3 HF RCTs—using data from 59,000 patients at the Mayo Clinic highlights its remarkable predictive accuracy.

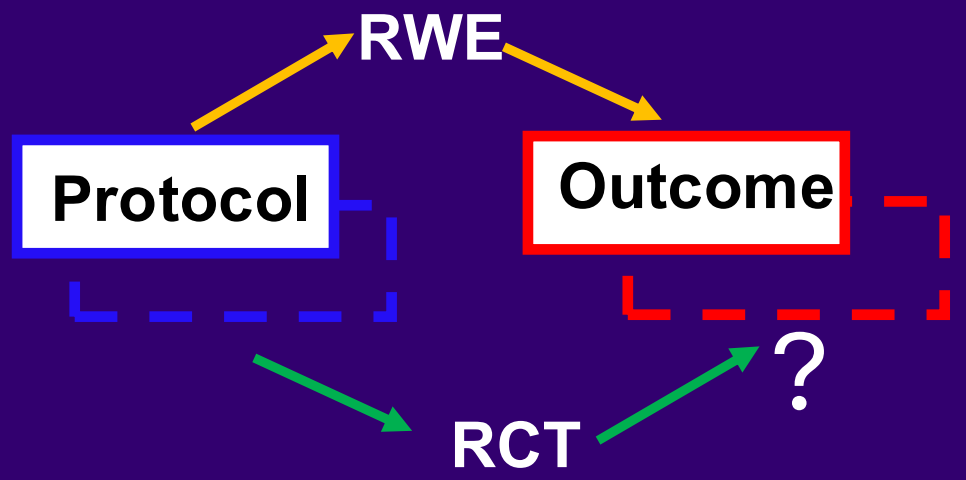


Table 1. Patient Characteristics in the Randomized, Controlled Trials (RCTs) Versus the Corresponding Real-World Evidence (RWE) Emulations

Characteristics	LEADER	DECLARE-TIMI58	EMPA-REG OUTCOME	CANVAS	CARME-LINA	TECOS	SAVOR-TIMI 53	CARO-LINA	TRITON-TIMI 38	PLATO
Age, y, mean±SD										
RCT	64.3±7.2	64.0±6.8	63.1±8.7	63.3±8.3	65.9±9.1	65.5±8.0	65.1±8.6	64.1±9.5	61*	62*
RWE	67.7±6.0	62.6±5.9	61.9±9.6	65.3±6.6	72.3±9.0	72.3±8.6	68.9±7.8	70.4±7.7	56.9±10.0	65.4±10.8
Female, %										
RCT	35.7	37.4	28.6	35.8	37.1	29.3	33.1	40.0	26.0	28.4
RWE	53.5	42.3	40.5	46.3	53.3	47.4	46.8	52.3	21.1	32.5
Smoking, %										
RCT	12.1	14.5	13.2	17.8	10.2	11.4	—	19.7	38.0	35.9
RWE	10.2	7.8	9.4	8.4	11.5	15.7	7.8	6.0	36.7	37.4
History of MI, %										
RCT	30.7	—	46.6	—	—	42.6	37.8	—	18.0	20.5
RWE	13.0	9.3	10.3	8.8	21.3	29.9	11.2	9.0	17.2	23.7
Hypertension, %										
RCT	90.0	89.4	95.0†	90.0	91.0	≥78.8†	81.8	90.1	64.0	65.4
RWE	97.1	88.3	90.9	96.6	96.4	94.0	88.7	92.7	51.1	70.2
CHF, %										
RCT	17.9	10.1	10.1	14.4	26.8	18.0	12.8	4.5	—	5.6
RWE	20.9	11.4	11.8	12.6	33.2	39.3	17.6	12.1	5.0	11.9
Diabetes, %										
RCT	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	23.0	25.0
RWE	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	28.1	47.4
Previous CVD, %										
RCT	81.4	40.6	75.6	65.6	58.1	74.0	78.6	41.9	—	—
RWE	63.0	51.4	45.9	45.4	88.1	100.0	56.3	54.4	37.2	54.7

Table 2. Study Sizes and Event Rates

		RCT						RWE					
		Exposure			Comparator			Exposure			Comparator		
Study	Outcome	Events	N	Rate*	Events	N	Rate*	Events	N	Rate*	Events	N	Rate*
LEADER	3P MACE	608	4668	3.4	694	4672	3.9	1352	84346	2.1	1955	84346	2.6
DECLARE-TIMI 58	HHF+cardiovascular death	417	8582	1.2	496	8578	1.5	242	24895	1.6	367	24895	2.4
EMPA-REG OUTCOME	3P MACE	490	4687	3.7	282	2333	4.4	416	51875	1.5	478	51875	1.9
CANVAS	3P MACE	564†	5795	2.7	496†	4347	3.2	772	76099	1.5	990	76099	1.9
CARMELINA	3P MACE	434	3494	5.8	420	3485	5.6	1540	50913	4.6	1826	50913	5.2
TECOS	3P MACE+angina	839	7257	4.1	851	7266	4.2	8106	174739	7.3	9692	174739	8.3
SAVOR-TIMI 53	3P MACE	613	8280	3.6	609	8212	3.6	1662	91064	2.4	2390	91064	3.1
CAROLINA	3P MACE	356	3023	2.1	362	3010	2.1	373	24131	2.7	458	24131	3.0
TRITON-TIMI 38	3P MACE	643	6813	7.9	781	6795	9.7	718	21932	3.8	960	21932	3.9
PLATO	3P MACE	864	9333	9.8	1014	9291	11.7	649	13980	8.0	858	13980	7.1

Franklin, J.M., Patorno, E., Desai, R.J., Glynn, R.J., Martin, D., Quinto, K., Pawar, A., Bessette, L.G., Lee, H., Garry, E.M. and Gautam, N., 2021. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative. *Circulation*, *143*(10), pp.1002-1013.

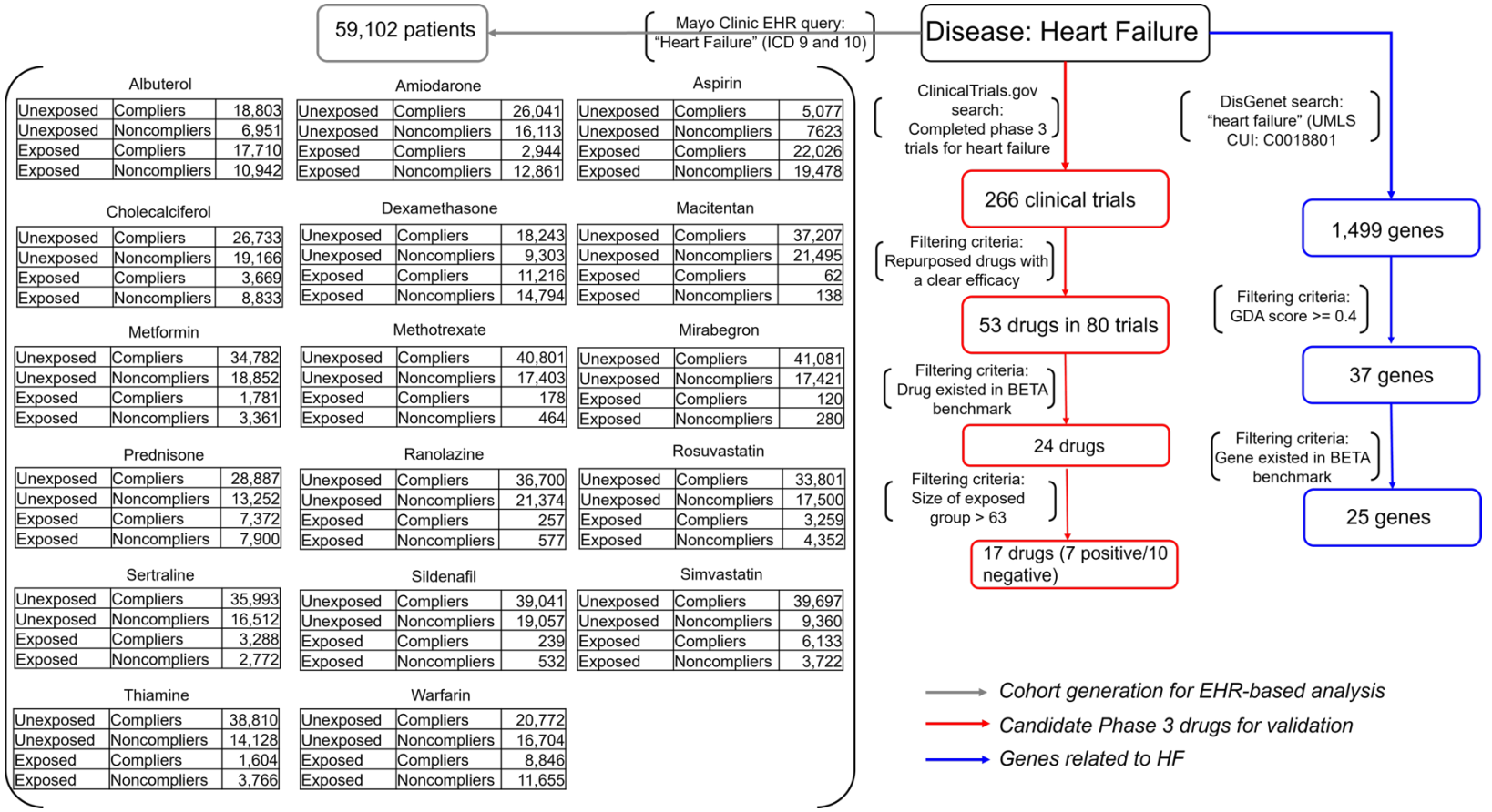
Table 1. Effect Estimates and Agreement Metrics

Study No.	Trial name	Effect estimates (95% CI)			Agreement			
		RCT	Database studyAdjusted[a,b]	Database study crude[a,b]	Standardized difference[c]	Statistical significance	Estimate	Standardized difference
1	LEADER	0.87 (0.78 to 0.97)	0.82 (0.76 to 0.87)	0.57 (0.54 to 0.61)	0.90	SA	EA	SD
2	DECLARE-TIMI58	0.83 (0.73 to 0.95)	0.69 (0.59 to 0.81)	0.47 (0.41 to 0.53)	1.76	SA		SD
3	EMPA-REG	0.86 (0.74 to 0.99)	0.83 (0.73 to 0.95)	0.63 (0.57 to 0.70)	0.35	SA	EA	SD
4	CANVAS	0.86 (0.75 to 0.97)	0.77 (0.70 to 0.85)	0.58 (0.54 to 0.62)	1.34	SA	EA	SD
5	CARMELINA	1.02 (0.89 to 1.17)	0.90 (0.84 to 0.96)	0.90 (0.86 to 0.95)	1.61	SAP	EA	SD
6	TECOS	0.98 (0.88 to 1.09)	0.89 (0.86 to 0.91)	0.81 (0.79 to 0.84)	1.71	SAP	EA	SD
7	SAVOR-TIMI	1.00 (0.89 to 1.12)	0.81 (0.76 to 0.86)	0.65 (0.62 to 0.69)	3.16	SAP		
8	LEAD-2	0 (−0.20 to 0.20)	0.05 (−0.11 to 0.22)	0.01 (−0.11 to 0.13)	−0.37	SA	EA	SD
9	TRITON-TIMI	0.81 (0.73 to 0.90)	0.88 (0.79 to 0.97)	0.70 (0.65 to 0.76)	−1.11	SA	EA	SD
10	PLATO	0.84 (0.77 to 0.92)	0.92 (0.83 to 1.02)	0.84 (0.78 to 0.91)	−1.31		EA	SD
11	ISAR-REACT 5	1.36 (1.09 to 1.70)	NA[d]	NA[d]	NA[d]	NA[d]		
12	ARISTOTLE	0.79 (0.66 to 0.95)	0.68 (0.61 to 0.76)	0.66 (0.62 to 0.71)	1.36	SA	EA	SD
13	RE-LY	0.66 (0.53 to 0.82)	0.73 (0.60 to 0.90)	0.67 (0.58 to 0.78)	−0.66	SA	EA	SD
14	ROCKET AF	0.79 (0.66 to 0.96)	0.70 (0.62 to 0.80)	0.76 (0.69 to 0.84)	1.00	SA	EA	SD
15	EINSTEIN DVT	0.68 (0.44 to 1.04)	0.75 (0.62 to 0.90)	0.85 (0.76 to 0.95)	−0.42	SAP	EA	SD
16	EINSTEIN PE	1.12 (0.75 to 1.68)	0.67 (0.55 to 0.80)	0.73 (0.64 to 0.83)	2.28	SAP		
17	RE-COVER II	1.08 (0.64 to 1.80)	1.15 (0.74 to 1.78)	1.48 (1.09 to 2.00)	−0.18	SA	EA	SD
18	AMPLIFY	0.84 (0.60 to 1.18)	0.81 (0.54 to 1.23)	0.64 (0.50 to 0.82)	0.13	SA	EA	SD
19	RECORD1	0.25 (0.14 to 0.47)	0.17 (0.10 to 0.29)	0.25 (0.18 to 0.34)	0.63	SA	EA	SD
20	TRANSCEND	0.92 (0.81 to 1.05)	0.88 (0.81 to 0.96)	0.80 (0.74 to 0.85)	0.55		EA	SD
21	ONTARGET	1.01 (0.94 to 1.09)	0.83 (0.77 to 0.90)	0.68 (0.64 to 0.72)	3.46	SAP		
22	HORIZON-PFT	0.59 (0.42 to 0.83)	0.72 (0.55 to 0.94)	1.08 (0.86 to 1.35)	−0.90	SA	EA	SD
23	VERO	0.44 (0.29 to 0.68)	NA[d]	NA[d]	NA[d]	NA[d]	NA[d]	NA[d]
24	DAPA-CKD	0.61 (0.51 to 0.72)	0.80 (0.52 to 1.26)	0.41 (0.29 to 0.58)	−1.10	SD		SD
25	PARADIGM-HF	0.80 (0.73 to 0.87)	1.02 (0.91 to 1.14)	0.95 (0.90 to 1.02)	−3.42			
26	P04334[e,f]	0.56 (0.44 to 0.72)	0.78 (0.62 to 0.97)	0.87 (0.76 to 0.99)	−1.95	SA	SD	SD
27	D5896	1.07 (0.70 to 1.65)	1.38 (0.90 to 2.13)	1.41 (1.00 to 1.98)	−0.81	SA	EA	SD
28	IMPACT[e,g]	0.85 (0.80 to 0.90)	1.13 (1.04 to 1.23)	1.22 (1.15 to 1.30)	−5.46			
29	POET-COPD	0.83 (0.77 to 0.90)	1.02 (0.93 to 1.12)	1.05 (0.99 to 1.12)	−3.27			
30	INSPIRE[h]	0.97 (0.84 to 1.12)	0.93 (0.90 to 0.96)	0.83 (0.81 to 0.85)	0.56	SA	EA	SD
31	CAROLINA[i]	0.98 (0.84 to 1.14)	0.91 (0.79 to 1.05)	0.92 (0.83-1.01)	0.70	SA	EA	SD
32	PRONOUNCE[i]	1.28 (0.59 to 2.79)	1.35 (0.94 to 1.93)	1.70 (1.30 to 2.21)	−0.12	SA	EA	SD

Wang, S.V., Schneeweiss, S., Franklin, J.M., Desai, R.J., Feldman, W., Garry, E.M., Glynn, R.J., Lin, K.J., Paik, J., Patorno, E. and Suissa, S., 2023. Emulation of randomized clinical trials with nonrandomized database analyses: results of 32 clinical trials. *Jama*, *329*(16), pp.1376-1385.

Classification	Trial Name	RCT HR (95% CI)	Emulation Adjusted HR (95% CI)	SD
Full success	DECLARE-TIMI58	0.83 (0.70–0.98)	0.81 (0.70–0.95)	1.76
Full success	CANVAS	0.86 (0.75–0.98)	0.83 (0.70–0.98)	0.66
Full success	LEADER	0.87 (0.78–0.97)	0.89 (0.80–0.98)	0.53
Full success	PATENT-1	0.77 (0.62–0.96)	0.74 (0.60–0.90)	1.31
Full success	ARISTOTLE	0.66 (0.59–0.75)	0.67 (0.61–0.74)	1.26
Full success	EINSTEIN DVT	0.68 (0.57–0.81)	0.69 (0.58–0.83)	0.83
Full success	EINSTEIN PE	0.89 (0.70–1.13) → Non-significant	0.90 (0.74–1.10) → Non-significant	1.12
Full success	RUXO MF	0.50 (0.37–0.66)	0.52 (0.38–0.71)	0.47
Full success	HORIZON-PFT	0.73 (0.58–0.91)	0.72 (0.58–0.89)	0.95
Full success	DAPA-CKD	0.61 (0.51–0.72)	0.64 (0.54–0.76)	0.43
Full success	PARADIGM-HF	0.80 (0.72–0.87)	0.82 (0.74–0.91)	0.83
Full success	CARMELINA	1.02 (0.89–1.17) → Non-significant	1.00 (0.88–1.14) → Non-significant	1.13
Full success	PLATO	0.84 (0.76–0.92)	0.81 (0.74–0.90)	1.31
Full success	RE-COVER	0.79 (0.63–1.00) → Beneficial	0.74 (0.60–0.92)	1.38
Full success	AMPLIFY	0.66 (0.52–0.83)	0.64 (0.50–0.82)	1.53
Full success	DAPA-HF	0.74 (0.65–0.85)	0.78 (0.69–0.89)	1.67
Partial success	SAVOR-TIMI	1.01 (0.94–1.09) → Non-significant	0.94 (0.87–1.01) → Non-significant	1.32
Partial success	ROCKET AF	0.88 (0.75–1.03) → Non-significant	0.85 (0.74–0.98) → Beneficial	1.26
Partial success	ENGAGE AF	0.89 (0.76–1.05) → Non-significant	0.85 (0.73–1.00) → Beneficial	1.46
Partial success	IMPACT	1.02 (0.89–1.17) → Non-significant	1.17 (1.02–1.34) → Harmful	1.78
Failure	EXAMINE	1.00 (0.83–1.20) → Non-significant	0.80 (0.67–0.97) → Beneficial	2.21
Failure	ISAR-REACT5	0.74 (0.58–0.95) → Beneficial	1.06 (0.79–1.42) → Non-significant	2.27
Failure	POSEIDON	0.80 (0.59–1.08) → Non-significant	1.03 (0.77–1.37) → Non-significant	2.02
Failure	IDEFICOPD	0.88 (0.70–1.11) → Non-significant	1.27 (1.03–1.56) → Harmful	3.56
Failure	INSPIRE	0.97 (0.84–1.12) → Non-significant	0.93 (0.90–0.96) → Beneficial	0.56

Repurposing 17 repurposed drugs for HF (selected from 266 phase 3 HF trials)



Albuterol

Unexposed	Compliers	18,803
Unexposed	Noncompliers	6,951
Exposed	Compliers	17,710
Exposed	Noncompliers	10,942

Amiodarone

Unexposed	Compliers	26,041
Unexposed	Noncompliers	16,113
Exposed	Compliers	2,944
Exposed	Noncompliers	12,861

Aspirin

Unexposed	Compliers	5,077
Unexposed	Noncompliers	7623
Exposed	Compliers	22,026
Exposed	Noncompliers	19,478

Cholecalciferol

Unexposed	Compliers	26,733
Unexposed	Noncompliers	19,166
Exposed	Compliers	3,669
Exposed	Noncompliers	8,833

Dexamethasone

Unexposed	Compliers	18,243
Unexposed	Noncompliers	9,303
Exposed	Compliers	11,216
Exposed	Noncompliers	14,794

Macitentan

Unexposed	Compliers	37,207
Unexposed	Noncompliers	21,495
Exposed	Compliers	62
Exposed	Noncompliers	138

Metformin

Unexposed	Compliers	34,782
Unexposed	Noncompliers	18,852
Exposed	Compliers	1,781
Exposed	Noncompliers	3,361

Methotrexate

Unexposed	Compliers	40,801
Unexposed	Noncompliers	17,403
Exposed	Compliers	178
Exposed	Noncompliers	464

Mirabegron

Unexposed	Compliers	41,081
Unexposed	Noncompliers	17,421
Exposed	Compliers	120
Exposed	Noncompliers	280

Prednisone

Unexposed	Compliers	28,887
Unexposed	Noncompliers	13,252
Exposed	Compliers	7,372
Exposed	Noncompliers	7,900

Ranolazine

Unexposed	Compliers	36,700
Unexposed	Noncompliers	21,374
Exposed	Compliers	257
Exposed	Noncompliers	577

Rosuvastatin

Unexposed	Compliers	33,801
Unexposed	Noncompliers	17,500
Exposed	Compliers	3,259
Exposed	Noncompliers	4,352

Sertraline

Unexposed	Compliers	35,993
Unexposed	Noncompliers	16,512
Exposed	Compliers	3,288
Exposed	Noncompliers	2,772

Sildenafil

Unexposed	Compliers	39,041
Unexposed	Noncompliers	19,057
Exposed	Compliers	239
Exposed	Noncompliers	532

Simvastatin

Unexposed	Compliers	39,697
Unexposed	Noncompliers	9,360
Exposed	Compliers	6,133
Exposed	Noncompliers	3,722

Thiamine

Unexposed	Compliers	38,810
Unexposed	Noncompliers	14,128
Exposed	Compliers	1,604
Exposed	Noncompliers	3,766

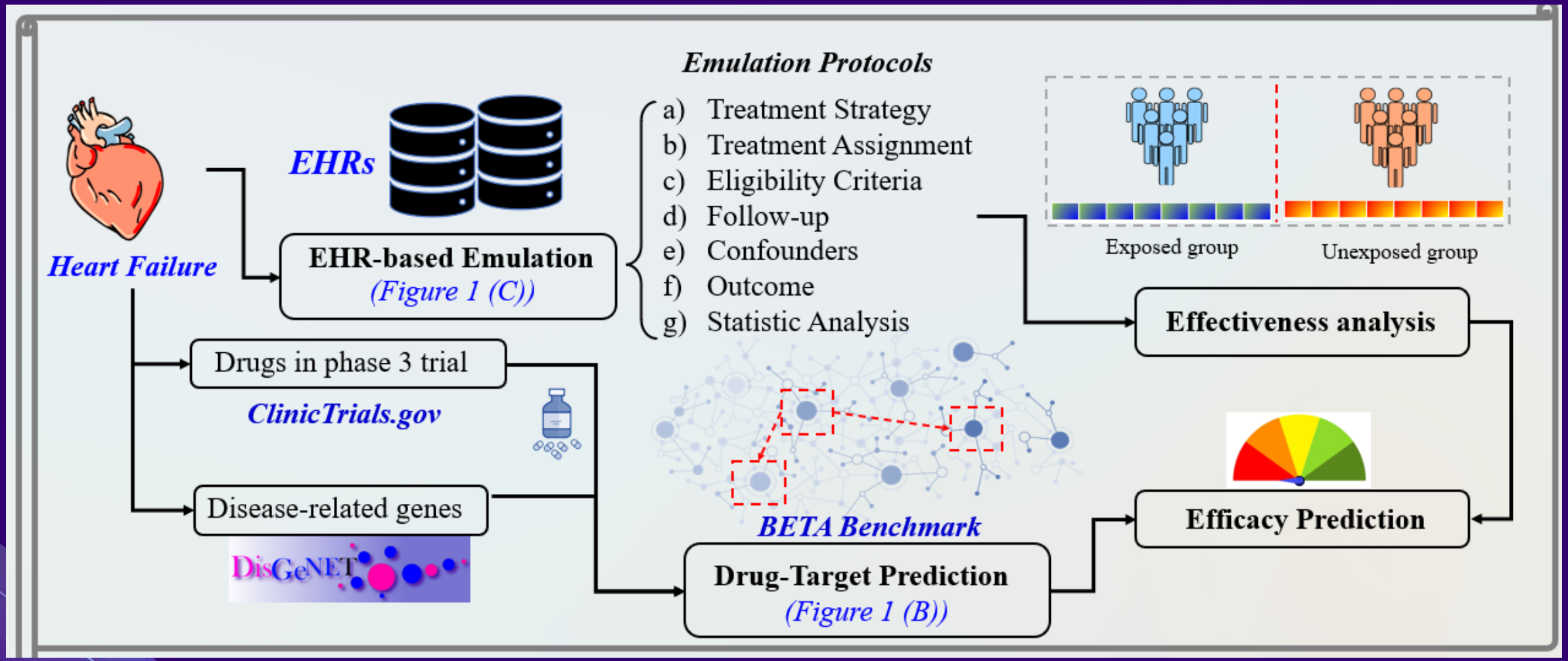
Warfarin

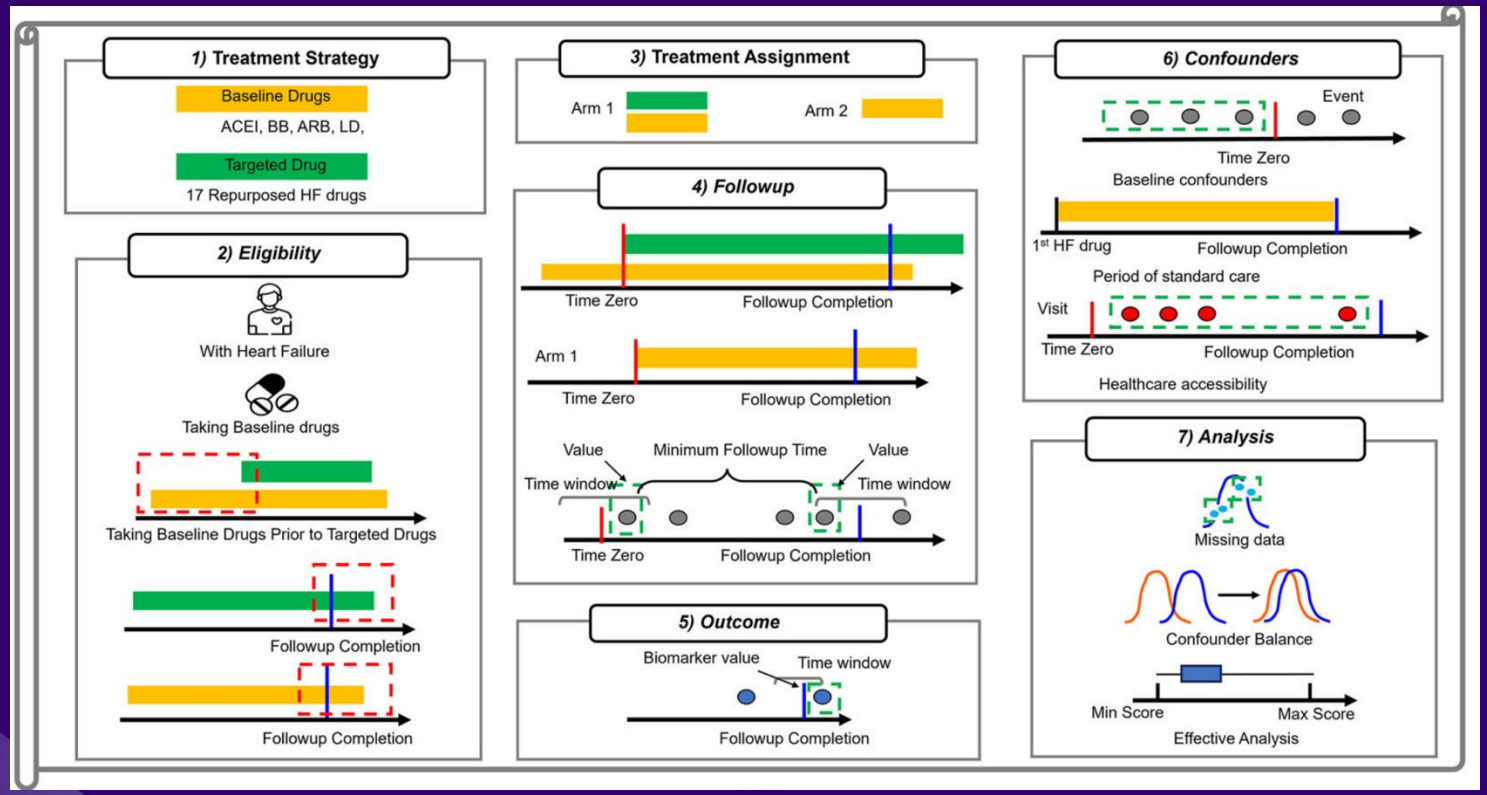
Unexposed	Compliers	20,772
Unexposed	Noncompliers	16,704
Exposed	Compliers	8,846
Exposed	Noncompliers	11,655

Table 1 | Statistics of the HF patients from the Mayo Clinic EHR

Variables	# patient	Prognostic Markers	# patient
Age ± SD	54.0 ± 31.5	NT-proBNP	5845
Ethnicity		Creatinine	58,350
Non-Hispanic White	52,329	BUN	58,315
Others	6674	Hemoglobin	58,160
Baseline HF drugs		Aldosterone	2941
ACEI	26,719	BNP	36,733
BB	46,081	CRP	35,360
ARB	18,276	eGFR	53,784
LD	53,565	Renin	2674
HF subtypes		TnI	5989
HFpEF	24,427	TnT	37,603
HFrEF	21,134	**Repurposed drugs**	
Confounders		Albuterol	28,652
Arrhythmias (including baseline variations, Atrial fibrillation, Flutter)	11,566	Amiodarone	15,805
Cardiac Arrest (including history of)	726	Aspirin	41,504
Myocardial infarction (past, prior)	4492	Cholecalciferol	12,502
Hypertension (including long-standing)	19,955	Dexamethasone	26,010
Pulmonary hypertension	1680	Macitentan	200
Ischemic heart disease	10,298	Metformin	5142
Hypertensive heart and Renal disease	1597	Methotrexate	642
Medications affecting heart health	456	Mirabegron	400
Congestive heart failure, unspecified	12,074	Prednisone	15,272
Renal Failure (including end-stage, acute, and chronic)	7776	Ranolazine	834
Diabetes Mellitus (type 1 and type 2)	9083	Rosuvastatin	7611
		Sertraline	6060
		Sildenafil	771
		Simvastatin	9855
		Thiamine	5370
# HF in total	59,102	Warfarin	20,501

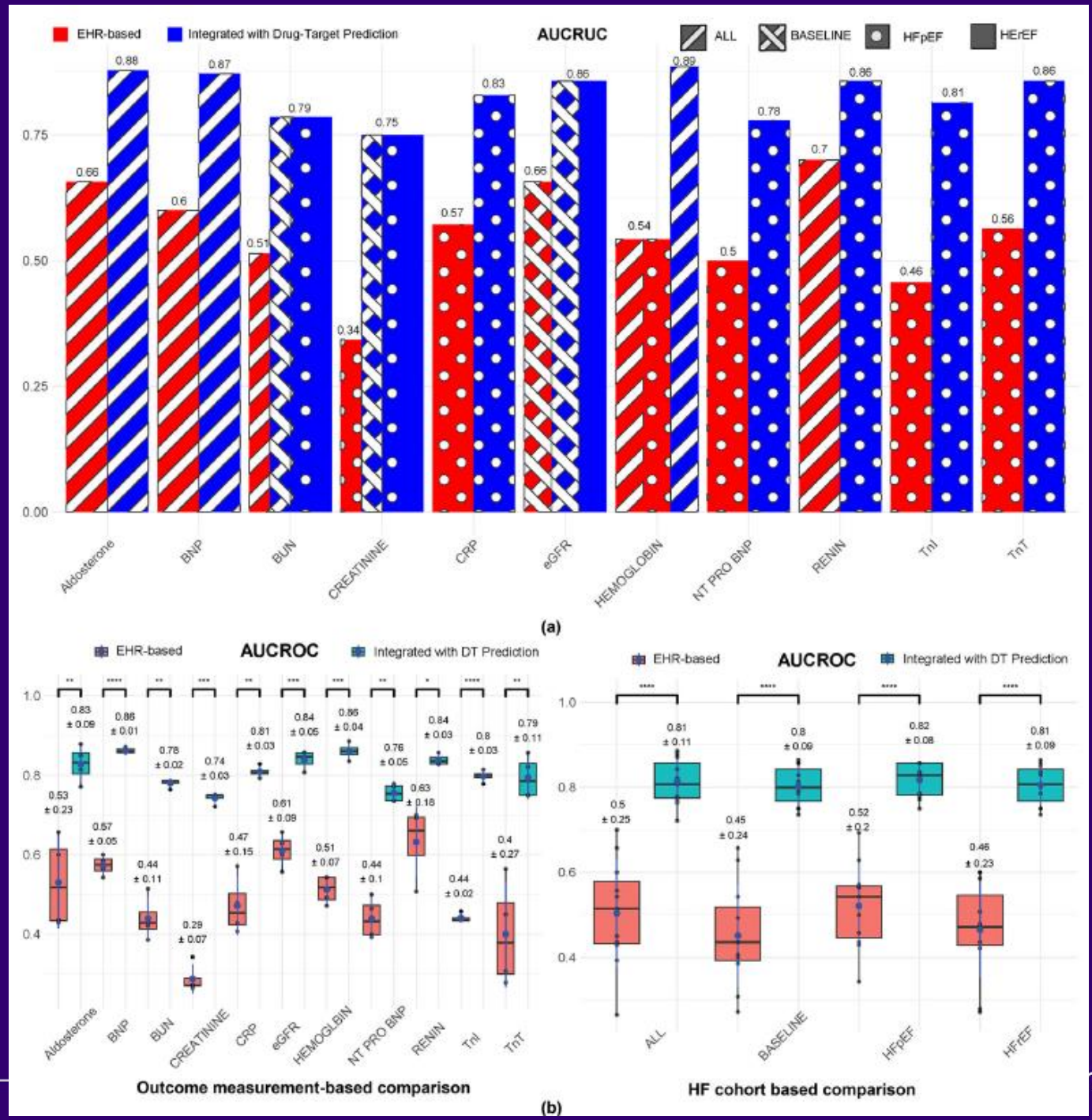
Overview

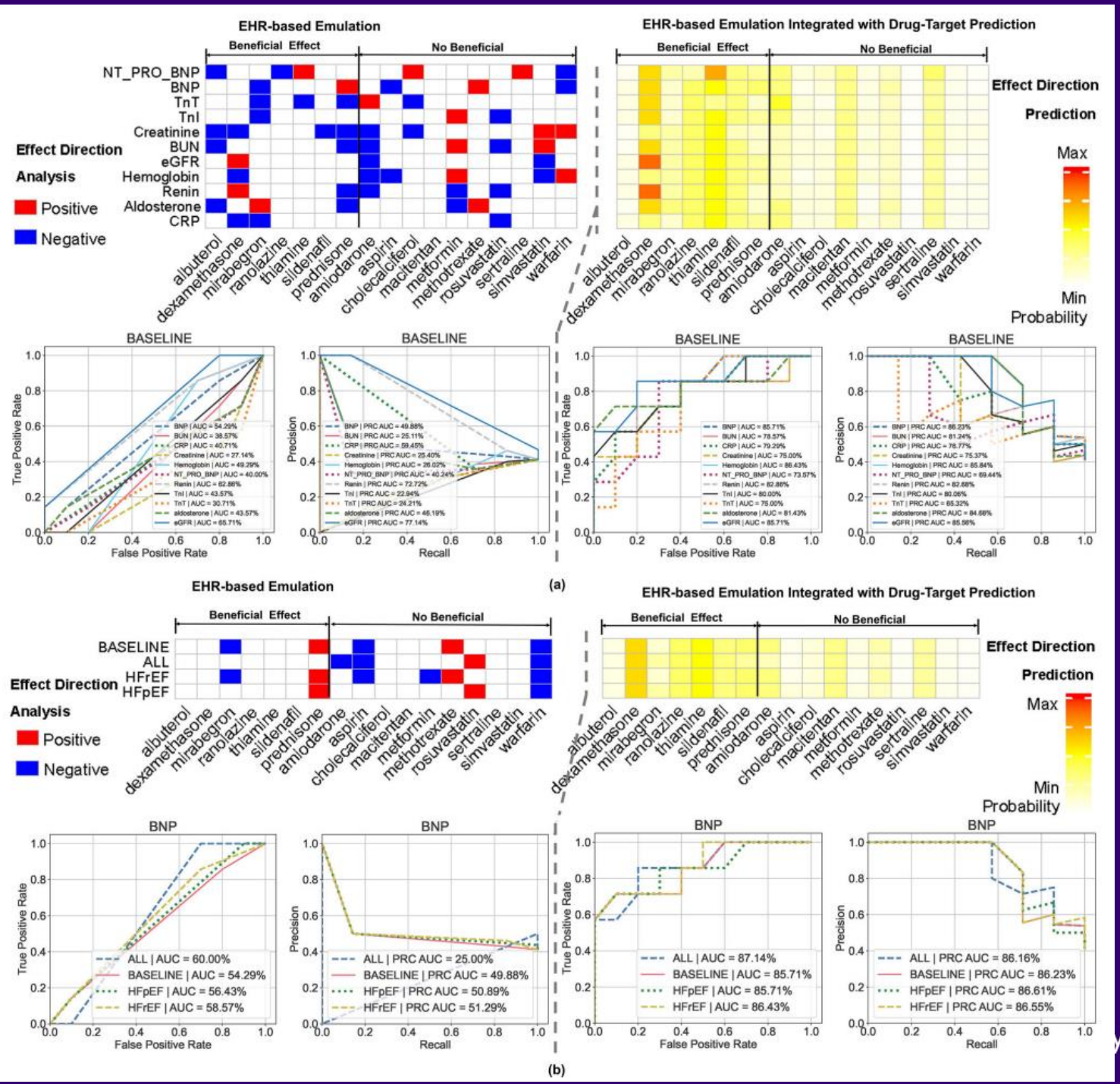

Emulation Protocols
a) Treatment Strategy
b) Treatment Assignment
c) Eligibility Criteria
d) Follow-up
e) Confounders
f) Outcome
g) Statistic Analysis
EHRs
Heart Failure
EHR-based Emulation
(Figure 1 (C))
Exposed group
Unexposed group
Effectiveness analysis
Drugs in phase 3 trial
ClinicTrials.gov
Disease-related genes
DisGeNET
BETA Benchmark
Drug-Target Prediction
(Figure 1 (B))
Efficacy Prediction



Supplementary Table 1. Emulation protocol

Protocol component	Description
Aim	To evaluate the improvement in outcomes for HF patients receiving additional targeted repurposed drugs alongside standard heart failure care, compared to those receiving standard care alone, we assessed the causal effect on 11 key prognostic markers by estimating the odds of achieving normal outcomes. Please note the distinction between the purpose of emulation and our study objective. While the odds ratio captures both the magnitude and direction of the association between taking a repurposed drug and the likelihood of achieving a normal prognostic marker, it is later transformed and integrated with drug-target predictions to estimate the probability of the drug demonstrating beneficial efficacy compared to the standard of care (i.e., efficacy direction).
Eligibility criteria	• The HF patient has received standard-of-care treatment with HF medications, including ACE inhibitors (ACEI), beta-blockers (BB), angiotensin receptor blockers (ARB), and loop diuretics (LD) • The HF patient has not received any targeted repurposed drugs before Time Zero. • The HF patient has completed all follow-up visits (refer to Supplementary Table 6 for the detailed follow-up schedule for each drug). • The HF patient is over 18 years old • The HF patient has consented to provide their EHR for research purposes.
Treatment strategies	Arm 1 (Treatment Group): • Continuous therapy with any HF medications (e.g., ACEI, BB, ARB, LD). • Continuous therapy with the targeted repurposed drugs. Arm 2 (Standard-of-Care Group): • Continuous therapy with any HF medications (e.g., ACEI, BB, ARB, LD). • No administration of the targeted repurposed drugs.
Treatment Assignment	The treatment arms are determined on a case-by-case basis in clinical practice based on the specific condition of the patients and the clinical judgment of healthcare providers.
Follow-up	The follow-up begins at time zero, defined as the point when the eligibility criteria are met and treatment is assigned by healthcare providers. For Arm 2, the treatment at time zero specifically refers to any of the HF drugs (e.g., ACEI, BB, ARB, LD), and for Arm 1, the treatment refers to a targeted repurposed drug at time zero. The follow-up ends after a pre-defined period for each targeted repurposed drug (see Supplementary Table 6 for details).
Confounders	Baseline Confounder: Arrhythmias, cardiac arrest, myocardial infarction, hypertension, pulmonary hypertension, ischemic heart disease, hypertensive heart and renal disease, medications affecting heart health, congestive heart failure, renal failure, and diabetes mellitus. Others: • Prognostic markers at time zero and the corresponding elapsed time, defined as the baseline biomarker values measured at time zero and the interval from time zero until the end of the follow-up period. • HF subtypes (Baseline, All, HFpEF, HFrEF), HF medication types (ACEI, BB, ARB, LD), sex, age, and ethnicity. • Length of HF medication consumption, defined as the total duration of HF drug usage through the end of follow-up. • Frequency of medication orders, defined as the count of HF- and repurposed-drug orders from time zero until the end of the follow-up period, based on treatments administered during follow-up visits.
Outcome	11 prognostic outcome measures (NT-proBNP, BNP, TnT, TnI, Creatinine, BUN, eGFR, Hemoglobin, Renin, Aldosterone, CRP)
Causal Contrast	Observational analog of PP and ITT effect.
Statistical Analysis	A logistic regression model was applied after balancing confounders using propensity scores and inverse probability weighting, with further adjustment for confounders included as covariates.

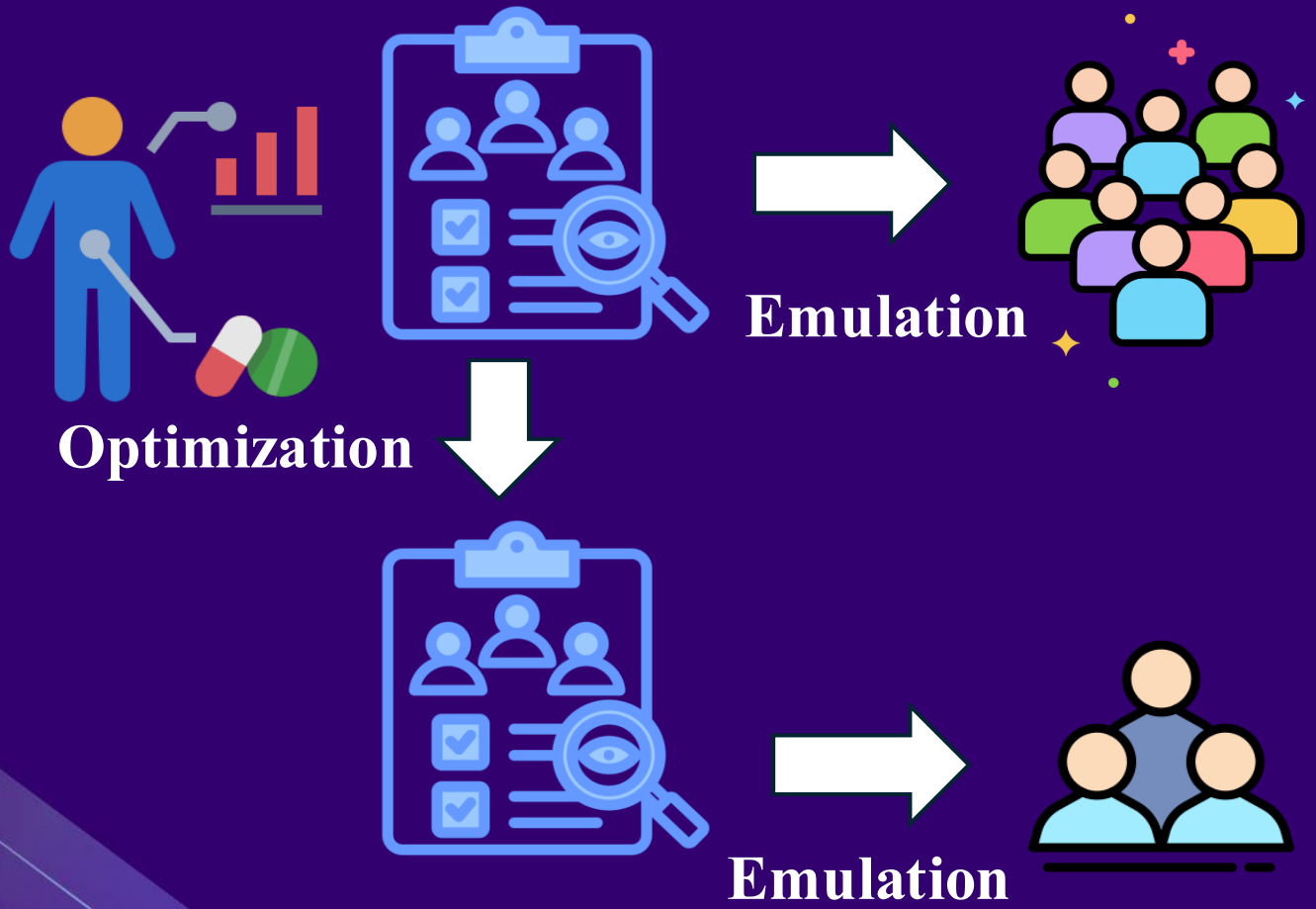

EHR-based
Integrated with Drug-Target Prediction
AUCRUC
ALL
BASELINE
HFpEF
HErEF
0.00
0.25
0.50
0.75
0.66
0.88
0.6
0.87
0.51
0.79
0.34
0.75
0.57
0.83
0.66
0.86
0.54
0.89
0.5
0.78
0.7
0.86
0.46
0.81
0.56
0.86
Aldosterone
BNP
BUN
CREATININE
CRP
eGFR
HEMOGLOBIN
NT PRO BNP
RENIN
TnI
TnT
(a)
EHR-based
AUCROC
Integrated with DT Prediction
1.0
0.8
0.6
0.4
0.53 ± 0.23
0.83 ± 0.09
0.57 ± 0.05
0.86 ± 0.01
0.44 ± 0.11
0.78 ± 0.02
0.29 ± 0.07
0.74 ± 0.03
0.47 ± 0.15
0.81 ± 0.03
0.61 ± 0.09
0.84 ± 0.05
0.51 ± 0.07
0.86 ± 0.04
0.44 ± 0.1
0.76 ± 0.05
0.63 ± 0.18
0.84 ± 0.03
0.44 ± 0.02
0.8 ± 0.03
0.4 ± 0.27
0.79 ± 0.11
Aldosterone
BNP
BUN
CREATININE
CRP
eGFR
HEMOGLOBIN
NT PRO BNP
RENIN
TnI
TnT
Outcome measurement-based comparison
EHR-based
AUCROC
Integrated with DT Prediction
0.5 ± 0.25
0.81 ± 0.11
0.45 ± 0.24
0.8 ± 0.09
0.52 ± 0.2
0.82 ± 0.08
0.46 ± 0.23
0.81 ± 0.09
ALL
BASELINE
HFpEF
HFrEF
HF cohort based comparison
(b)


EHR-based Emulation
Beneficial Effect
No Beneficial
NT_PRO_BNP
BNP
TnT
TnI
Creatinine
BUN
eGFR
Hemoglobin
Renin
Aldosterone
CRP
Effect Direction
Analysis
Positive
Negative
albuterol
dexamethasone
mirabegron
ranolazine
thiamine
sildenafil
prednisone
amiodarone
aspirin
cholecalciferol
macitentan
metformin
methotrexate
rosuvastatin
sertraline
simvastatin
warfarin
EHR-based Emulation Integrated with Drug-Target Prediction
Effect Direction
Prediction
Max
Min
Probability
BASELINE
True Positive Rate
False Positive Rate
Precision
Recall
(a)
BNP
ALL
HFrEF
HFpEF
ALL | AUC = 60.00%
BASELINE | AUC = 54.29%
HFpEF | AUC = 56.43%
HFrEF | AUC = 58.57%
ALL | PRC AUC = 25.00%
BASELINE | PRC AUC = 49.88%
HFpEF | PRC AUC = 50.89%
HFrEF | PRC AUC = 51.29%
ALL | AUC = 87.14%
BASELINE | AUC = 85.71%
HFpEF | AUC = 85.71%
HFrEF | AUC = 86.43%
ALL | PRC AUC = 86.16%
BASELINE | PRC AUC = 86.23%
HFpEF | PRC AUC = 86.61%
HFrEF | PRC AUC = 86.55%
(b)

Optimization of Clinical Trial



Lead by:
Dr. Xiaodi Li



11/11/2026 | *9:45 AM – 11:00 AM* | Cortez C

S111: Open Trails & Open Trials: Research Without Borders (Oral Presentations)

Presentation Type: Oral Presentations

Session Credits: 1.25

Date: Wednesday (11/11)
Time: 9:45 AM to 11:00 AM
Room: Cortez C

AERO: An AI Agent for Adaptive Eligibility Refinement and Optimization of Clinical Trial Criteria in Real-World Trial Emulation

Presentation Type: Paper - Student

Optimization of Clinical Trial

Study to Evaluate the Effect of Dapagliflozin on the Incidence of Worsening Heart Failure or Cardiovascular Death in Patients With Chronic Heart Failure (DAPA-HF)

ClinicalTrials.gov ID NCT03036124

Sponsor AstraZeneca

Information provided by AstraZeneca (Responsible Party)

Last Update Posted 2020-09-01

Study Overview

Brief Summary

The purpose of this study is to evaluate the effect of dapagliflozin on the incidence of worsening heart failure or cardiovascular death in patients with chronic heart failure with reduced ejection fraction

Detailed Description

This is an international, multicentre, parallel group, event-driven, randomized, double-blind, placebo-controlled study in patients with chronic heart failure with reduced ejection fraction (HFrEF), evaluating the effect of dapagliflozin versus placebo, given once daily in addition to background regional standard of care therapy, for the prevention of cardiovascular (CV) death or reduction of heart failure (HF) events.

Official Title

Study to Evaluate the Effect of Dapagliflozin on the Incidence of Worsening Heart Failure or Cardiovascular Death in Patients With Chronic Heart Failure With Reduced Ejection Fraction

Conditions

Chronic Heart Failure With Reduced Ejection Fraction (HFrEF)

Intervention/Treatment

- Drug: Dapagliflozin
- Drug: Placebo

Other Study ID Numbers

- D1699C00001
- 2016-003897-41 (EudraCT Number)
- 2016 (U.S. NIH Grant/Contract)

Study Start (Actual)
2017-02-08

Primary Completion (Actual)
2019-07-17

Study Completion (Actual)
2019-07-17

Enrollment (Actual)
4744

Study Type
Interventional

Phase
Phase 3

Safety of Switching From Donepezil to Rivastigmine Patch in Patients With Probable Alzheimer's Disease

ClinicalTrials.gov ID NCT00428389

Sponsor Novartis

Information provided by Novartis (Responsible Party)

Last Update Posted 2014-06-11

Study Overview

Brief Summary

This study was designed to evaluate the safety and tolerability of switching from donepezil to an initial dose of 5cm^2 rivastigmine patch formulation in patients with probable Alzheimer's Disease (MMSE 10-24). The study included a 5-week, open-label, randomized period followed by a 20-week open-label extension period. Patients were randomized to either an immediate switch from donepezil to rivastigmine patch formulation or to a switch to rivastigmine patch formulation following a 7-day withdrawal period.

Official Title

A Prospective, 5-Week, Open-Label, Randomized, Multi-Center, Parallel-Group Study With a 20-Week, Open-Label Extension Evaluating the Tolerability and Safety of Switching From Donepezil to an Initial Dose of 5 cm2 Rivastigmine Patch Formulation in Patients With Probable Alzheimer's Disease

Conditions

Alzheimer's Disease

Intervention/Treatment

- Drug: Rivastigmine 5 cm^2 transdermal patch
- Drug: Rivastigmine 10 cm^2 transdermal patch

Other Study ID Numbers

- CENA713DUS38
- CENA713DUS38E1 (Other Identifier) (OTHER: Novartis)

Study Start

2007-01

Primary Completion (Actual)

2008-02

Study Completion (Actual)

2008-02

Enrollment (Actual)

262

Study Type

Interventional

Phase

Phase 3

Future Works and Collaboration

MAYO CLINIC | News Network

News Releases | Health Topics | Medical Research | Media Contacts | About

RESEARCH

Researchers Improve AI Models Used in Drug Discovery

Mayo Clinic Staff

June 3, 2022

MAYO CLINIC | News Network

News Releases | Health Topics | Medical Research | Media Contacts | About

RESEARCH

'Virtual clinical trials' may predict success of hea failure drugs

Marla Broadfoot, Ph.D.

September 16, 2025

HEALTHCARE BREW

Stories | Topics | Resources | Events | Subscribe

TECH

How Mayo Clinic researchers are using algorithms to predict drug efficacy

The tool may aid the selection of drug candidates in the not-too-distant future.

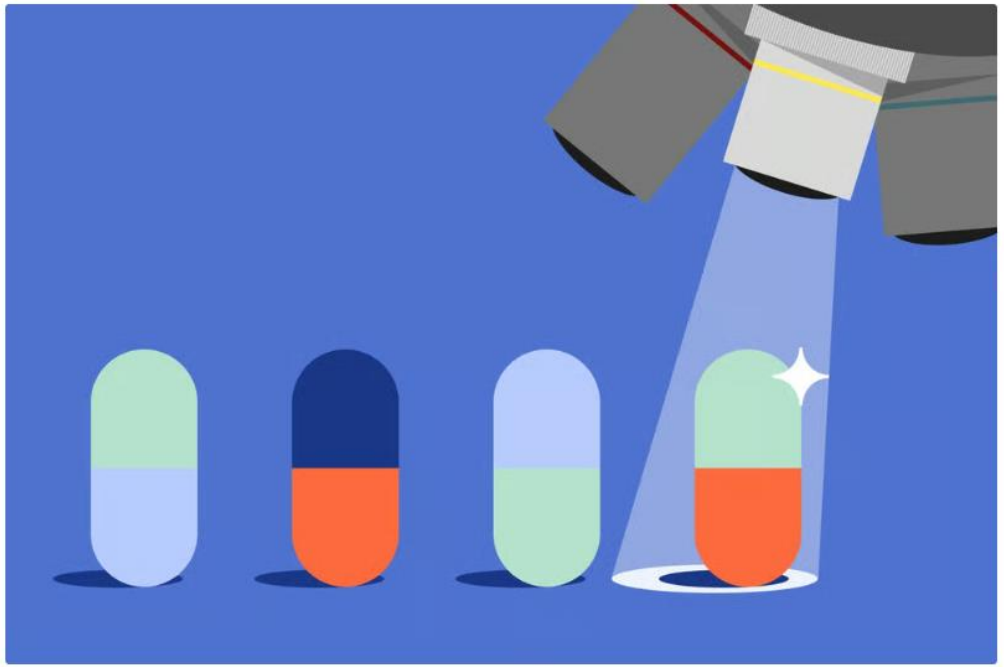


Illustration: Francis Scialabba

Acknowledgment

Cui Tao, Ph.D.
Xiaodi Li, Ph.D.
Shaika Chowdhury, Ph.D.
Shibo Zhou, Ph.D.
Sivaraman Rajaganapathy, Ph.D
Hongfang Liu, Ph.D.
Guoqian Jiang , Ph.D.
Ming Huang, Ph.D.
Wen Andrew, M.S.
Yue Yu, Ph.D.
Richard Weinshilboum, M.D.
Liewei Wang, M.D., Ph.D.
Qiying Dai, MD
Suzette J. Bielinski, Ph.D., M.Ed.
Yongbin Chen, MD, Ph.D
James R. Cerhan, M.D.,Ph.D.

He Zhe, Ph.D.

K99/R00 (BOND); R01

R01 (HOPE)

Thank you