BIOMEDICAL INFORMATICS AND MEDICAL EDUCATION

Automating the Creation of the Patient's Digital Twin through Generative AI

Shawn Murphy MD, Ph.D.

UW Medicine Chief Research Information Officer

Director of the UW Institute for Medical Data Science

Professor of Biomedical Informatics and Medical Education and Professor of Neurology

The Learning Healthcare System can be enabled by the Digital Twin

The Digital Twin in healthcare has its roots originating in mechanical engineering.

The creation of the Digital Twins allows a focus on a calculated in-silico representation of an individual who exists in the real world.

The Digital Twin results from high-fidelity computation based upon data from interactions with the patient such as clinical visits and laboratory testing.

Calculating a Digital Twin is crucial for Closed-loop systems in the Learning Healthcare System

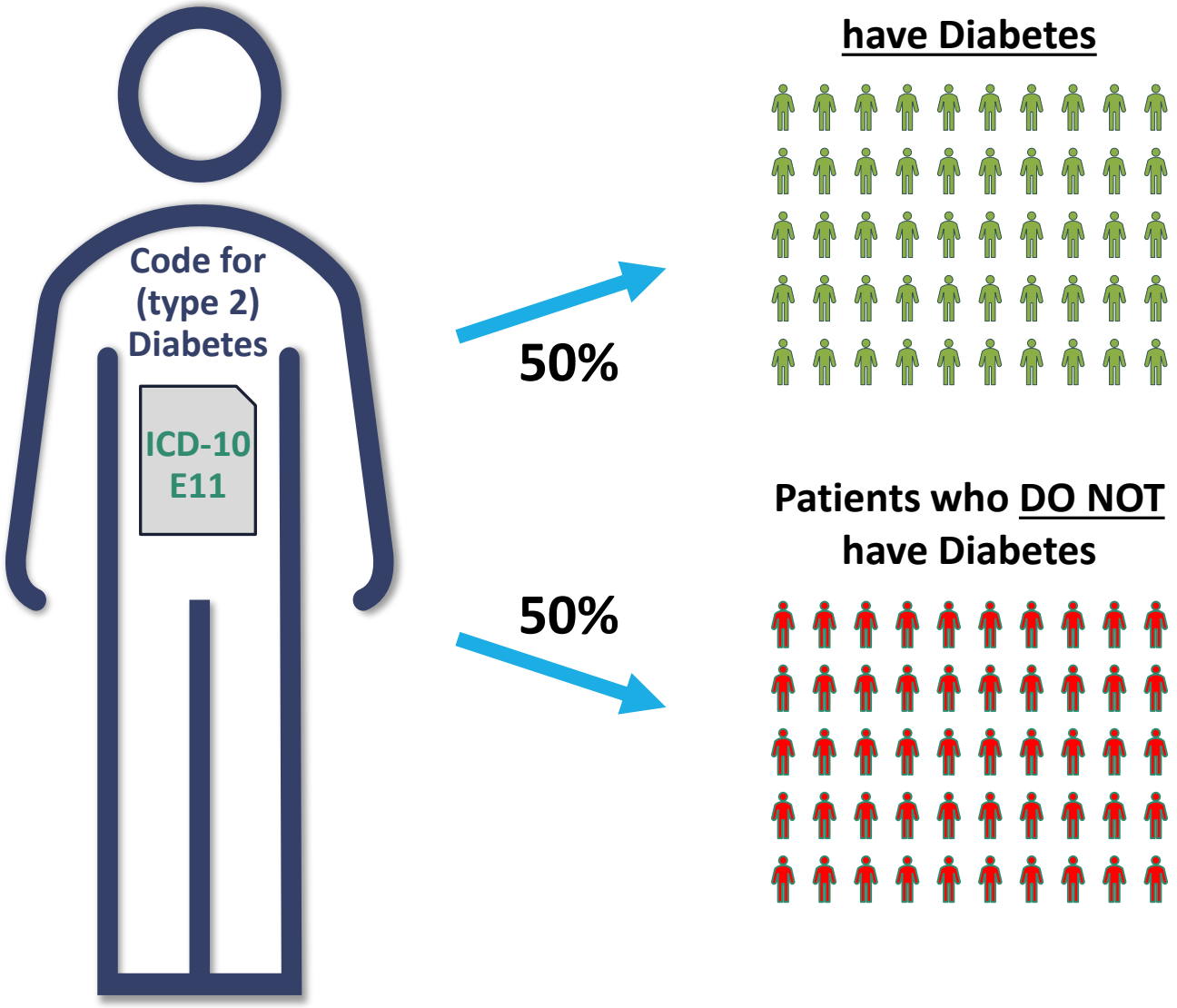
EHR Data Quality Problem

Just because patients have a diagnosis code in the EHR, it doesn't mean they have the medical condition.

Clinicians typically enter ICD diagnosis codes as "rule-out" codes for certain medical conditions like diabetes mellitus. They are often necessary to order a test to determine if a patient has Diabetes.

This results in "low precision" of ICD EHR diagnosis codes. Electronic queries will poorly estimate the number of patients with the disease.

For example, the number of patients who have type 2 diabetes mellitus can be overestimated by 50% or more.



Wei W-Q, et al. J Am Med Inform Assoc 2016;23:e20–e27.

Recall and Precision

Recall and Precision measure how well the patient's condition was captured by the algorithm.

Precision measures "of all items labeled by the algorithm as positive, how many were actually positive?"

Recall measures "of all the actual positive cases, how many did the algorithm correctly identify?

There are **tradeoffs** between recall and precision. An algorithm with better recall produces more cases, but can produce more false positives. A more precise algorithm results in fewer but 'better' cases but can miss other true positives.

ICD used for public health in USA

1. Many ICD codes are sensitive but not very precise (do not have high positive predictive values)
2. Higher precision of ICD codes increases with the number of ICD codes, but:
3. Patients with many visits/notes have an increased chance of incorrect/incidental ICD codes
4. Patients with little data are hard to classify
5. Natural Language Processing increases precision in some disease phenotypes – but not all

The Digital Twin is also important for population studies

Codes assigned for billing and other forms of accounting can be highly inaccurate because they follow the nuances of the business of medicine and are not intended for population science.

The Digital Twins get calculated in a step prior to population analysis because calculations are often nonlinear and need to be validated with high resolution data such as the medical chart before they are used.

Digital Twins include a full set of computational phenotypes for disease covariates and comparators that can be applied to the individuals in population science.

Klann, Jeffrey & Henderson, Darren & Morris, Michele & Estiri, Hossein & Weber, Griffin & Visweswaran, Shyam & Murphy, Shawn. (2023). A broadly applicable approach to enrich electronic-health-record cohorts by identifying patients with complete data: a multisite evaluation. Journal of the American Medical Informatics Association : JAMIA. 30. 10.1093/jamia/ocad166.

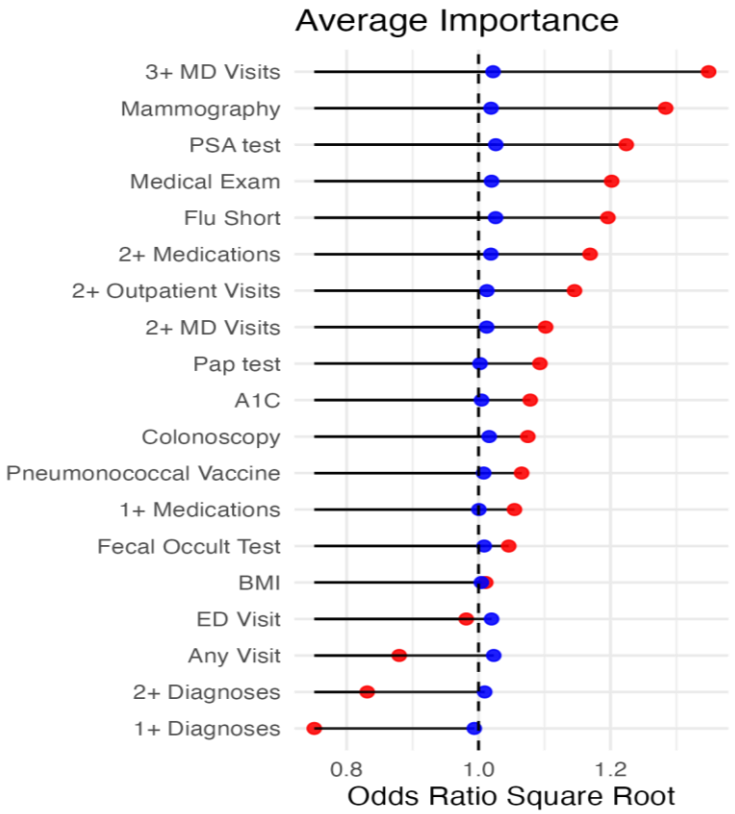
To produce a Digital Twin:

1) Assure data completeness is sufficient

Creation of a Digital Twin is done through calculation of a "loyalty cohort" which is a set of patients where most of their care is received in the hospital system that is used for calculation of the twin

Uses parameters such as:

- **# of routine care visits**
- **# of visits with the same doctor**
- **having a medical exam**
- **# or medication prescriptions**
- **having a colonoscopy**
- **getting a blood test for diabetes (HbA1c)**
- **etc., etc., etc.**



Contribution of each feature, on average across the three sites. Blue is the original regression equation; red is after retraining.

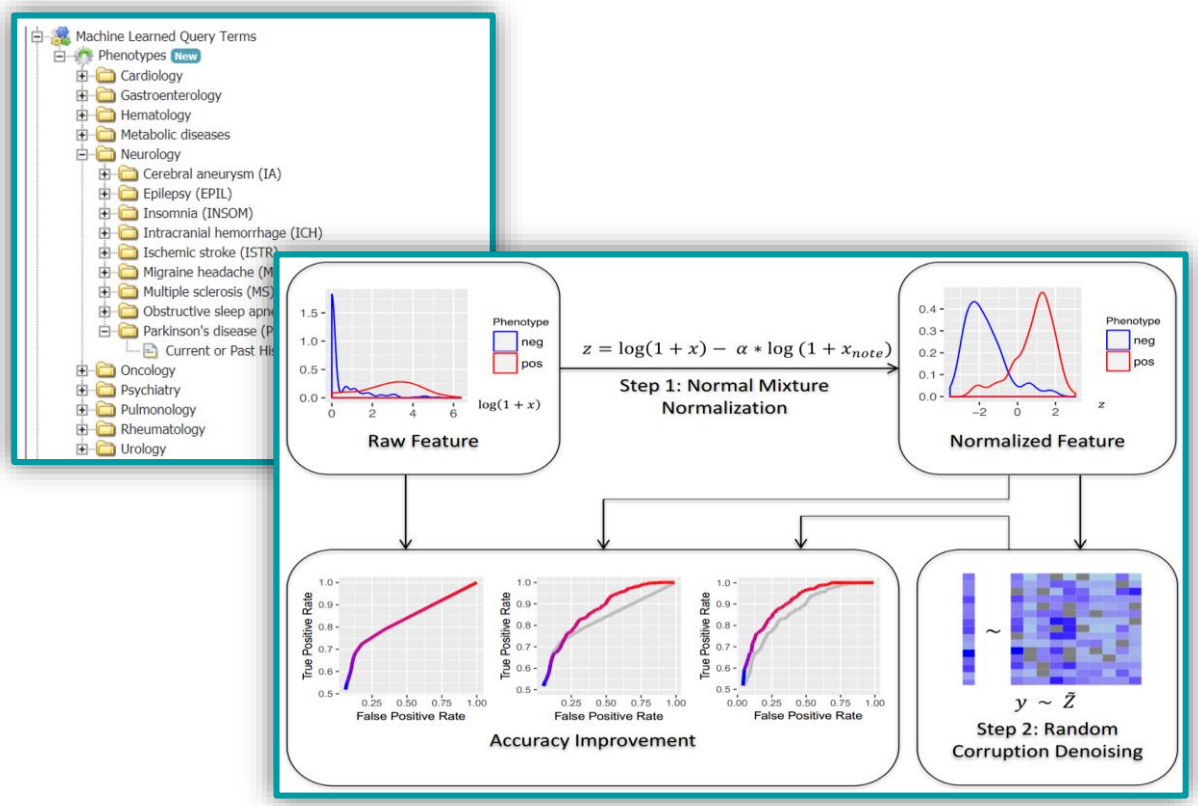
To produce a Digital Twin

2) Calculate a large array of phenotypes per individual

Use computation to create large sets of phenotypes for every individual within the loyalty cohort

Use supervised algorithms to extract conditions in the patient that are unique to a study and have Gold Standards to rely upon

The use of semi-supervised algorithms allow adjustments to be made for each hospital environment automatically and reduce the need to a-priori require many chart reviews as required by the use of supervised algorithms



J Am Med Inform Assoc. 2018 Jan 1;25(1):54-60. doi: 10.1093/jamia/ocx111.

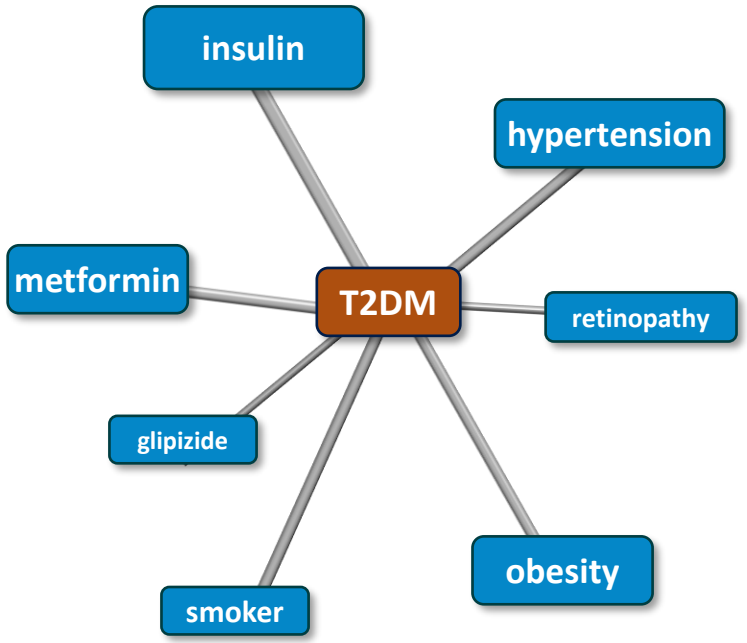
Enabling phenotypic big data with PheNorm.

Yu S[1,2], Ma Y[3], Gronsbell J[4], Cai T[5], Ananthakrishnan AN[6], Gainer VS[7], Churchill SE[8], Szolovits P[9], Murphy SN[7,10], Kohane IS[8], Liao KP[11], Cai T[4].

Keser and Komap computed phenotype pipeline

AI/ML methods to calculate probabilistic models for 1000+ medical conditions

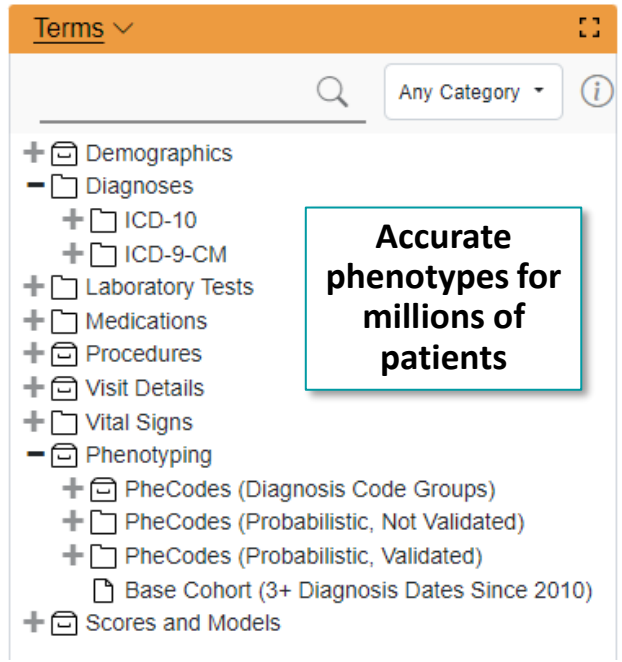
Knowledge Graph Derived from EHR



AI/ML Generated Probabilistic Models

Feature	Weight
Primary features	
Type 2 diabetes	0.5804
Utilization ICD Dates	-0.3662
Additional top features	
metformin	0.0846
Hypertension	0.0615
Insulin pump user	0.0409
Diabetic retinopathy	0.0403
Smoker	-0.0228
BMI 30 plus	-0.0255
Age 65 plus	-0.0366
Female	-0.0613

Computed Medical Conditions in i2b2



Hong C, et al. NPJ Digit Med. 2021 Oct 27;4(1):151

Liao KP, et al. JAMIA. 2019 Nov 1;26(11):1255-1262

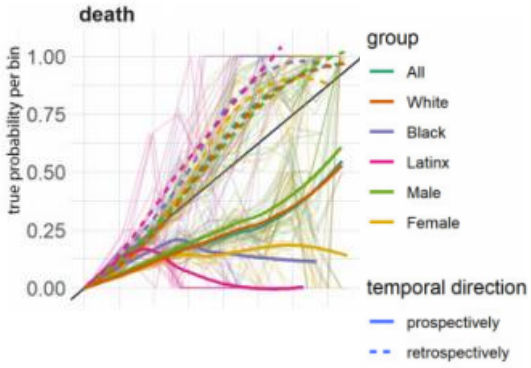
To produce a Digital Twin

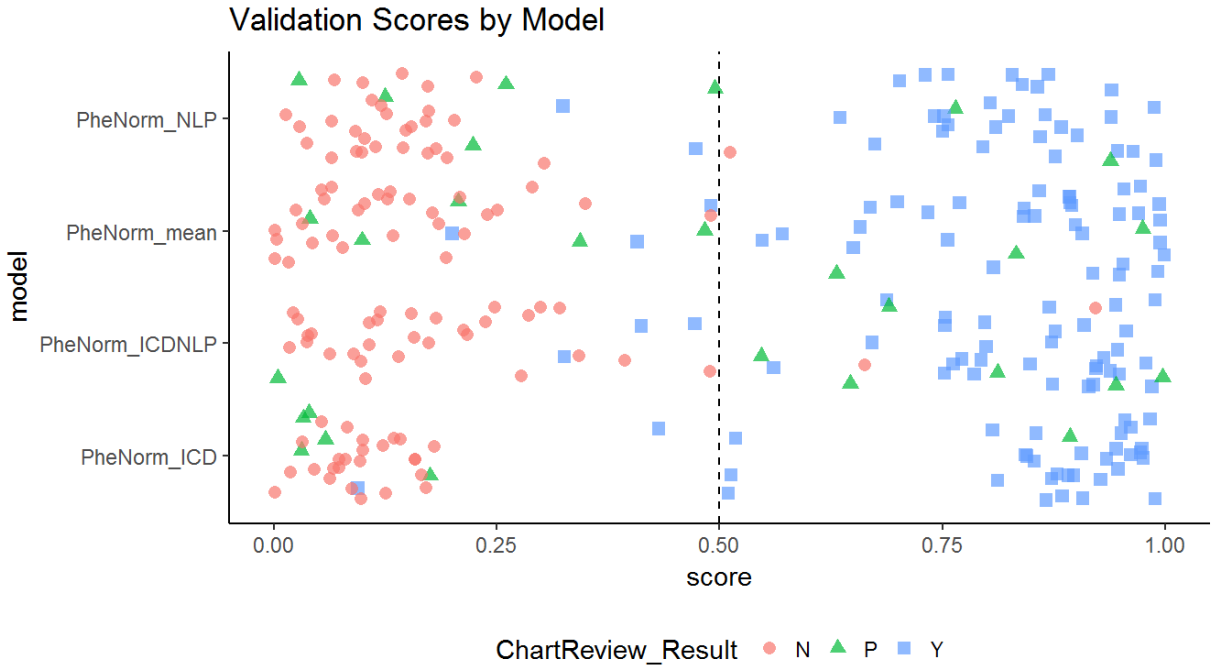
3) Check for accuracy on an ongoing basis

Chart reviews and direct patient confirmation is performed in a subset of individuals to ensure that the process of creating the Digital Twins is working properly

A validation strategy is constantly in motion with the Digital Twin, assuring that there is as little "drift" in accuracy as possible

Bias is measured through subgroup analysis



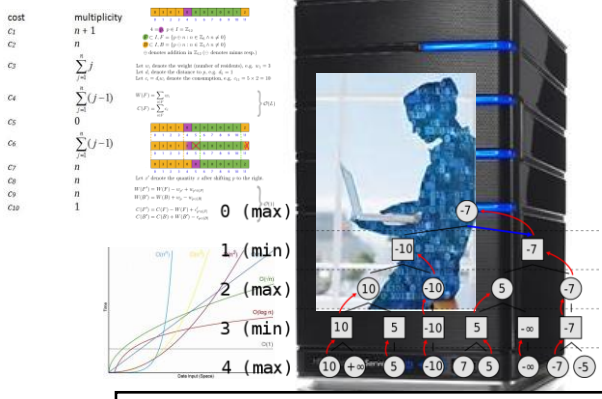


Estiri H, Strasser ZH, Rashidian S, Klann JG, Wagholikar KB, McCoy TH, Murphy SN. An objective framework for evaluating unrecognized bias in medical AI models predicting COVID-19 outcomes. J Am Med Inform Assoc. 2022 Jul 12;29(8):1334-1341. doi: 10.1093/jamia/ocac070. PMID: 35511151; PMCID: PMC9277645.

Enclave enables Digital Twin to empower Studies in Learning Healthcare System and Protect the Patient Data

Hospital Digital and IoT systems continuously output Patient Data

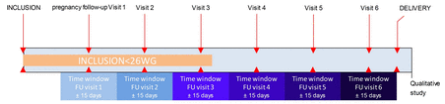
```
1. for ( i = 1; i ≤ a1.length; i++ )
   1.1. j = i
   1.2. while ( a1[i] != a2[j] )
   1.2.1.  if ( j ≥ a1.length)
   1.2.1.1.  return false
   1.2.2.  j++
   1.3. tmp = a2[j]
   1.4. a2[j] = a2[i]
   1.5. a2[i] = tmp
2. return true
```



Data Enclave

PATIENT STUDIES

Digital Twin of patient enables assessment of patient with high-resolution data from hospital electronic health record and instrumentation



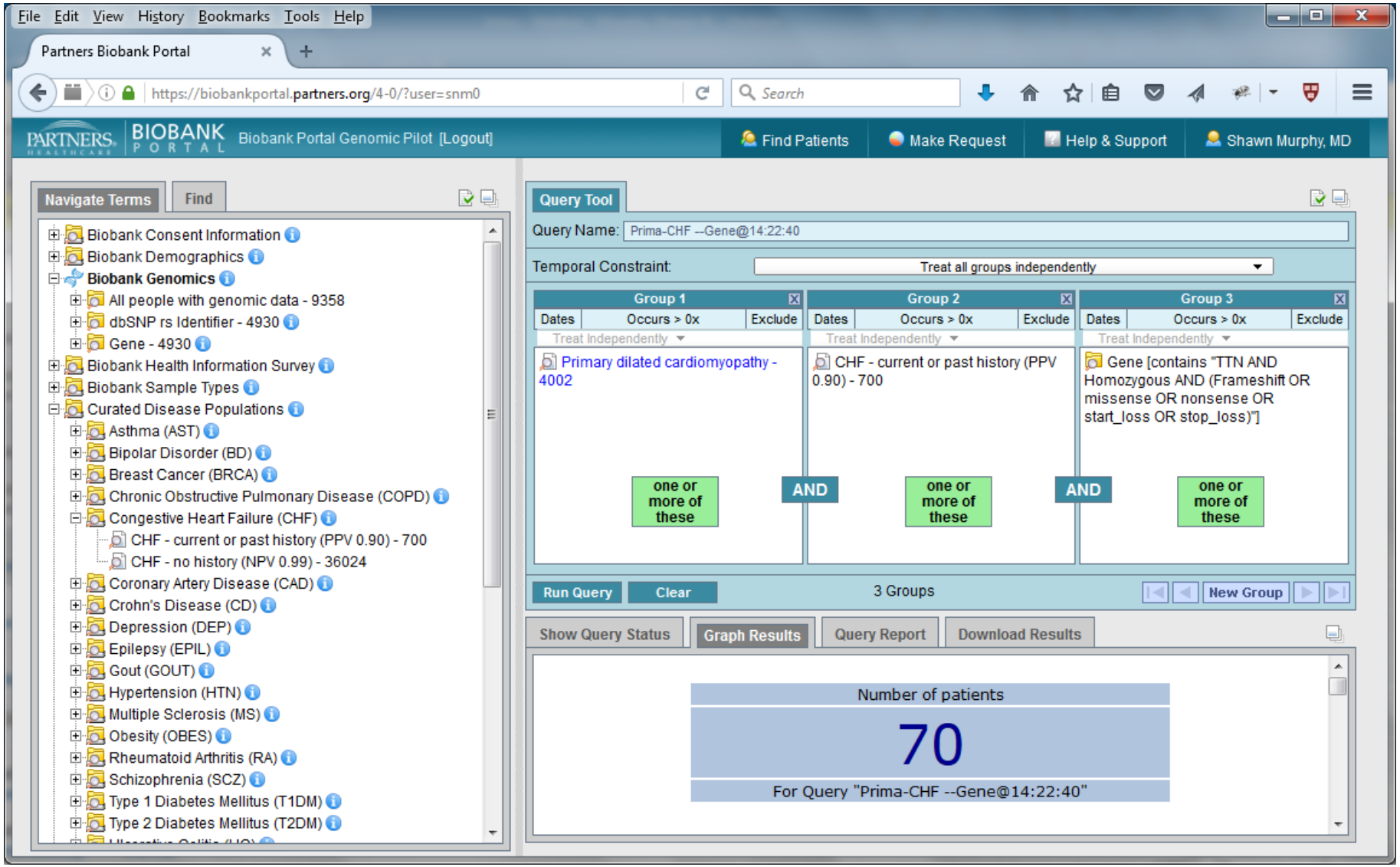
System drives Population Studies and Clinical Trial feasibility and enrollment Leading to Better Learning Healthcare System

FUTURE - Infusion of assessment of patient into clinical care

Machine Learned Phenotypes

- Abdominal hernia
- Acute bronchitis and bronchiolitis
- Acute pancreatitis
- Alcoholism
- Alzheimer's disease
- Aortic aneurysm
- Aplastic anemia
- Atrial fibrillation
- Atrioventricular block
- Autism spectrum disorders
- Basal cell carcinoma
- Bipolar Disease
- Bladder cancer
- Brain cancer
- Breast cancer
- Cerebral aneurysm
- Cholelithiasis
- Chronic pancreatitis
- Chronic sinusitis
- Coronary atherosclerosis
- Crohn's disease
- Deep vein thrombosis
- Depression
- Diverticulosis and diverticulitis
- Eating disorder
- Epilepsy
- Gastroesophageal reflux disease
- Gout
- Heart valve disorders
- Hyperlipidemia
- Hyperparathyroidism
- Hypertension
- Hypothyroidism
- Insomnia
- Intracranial hemorrhage
- Ischemic stroke
- Leukemia
- Lung cancer
- Melanoma
- Migraine headache
- Multiple sclerosis
- Myocardial infarction
- Neutropenia
- Non-Hodgkin lymphoma
- Obesity
- Obsessive compulsive disorder
- Obstructive sleep apnea
- Ovarian cancer
- Pancreatic cancer
- Parkinson's disease
- Peripheral vascular disease
- Pneumonia
- Polycystic ovaries
- Prostate cancer
- Pulmonary heart disease
- Renal cancer
- Renal failure
- Schizophrenia
- Substance addiction
- Suicidal ideation
- Suicide attempt or self-inflicted injury
- Thyroid cancer
- Tobacco use disorder
- Type 1 diabetes
- Type 2 diabetes
- Ulcerative colitis
- Urinary calculus
- Uterine cancer

Enables pervasive use of Digital Twins for Research Studies



Digital Twin helps Find “Normal” Brain MRI’s

Providing Clinically Relevant Brain MRI ADC Maps of Normal Children Ages 0-6

Medical Records Query using i2b2 finds Normal Children with Clinical Brain MRI Images.
NIBIB funded mi2b2 Workbench allows special audited access for investigator to Hospital Clinical Imaging Studies of Patients specified by Institutional Review Board (IRB) approval.
https://www.i2b2.org
http://mi2b2help.partners.org

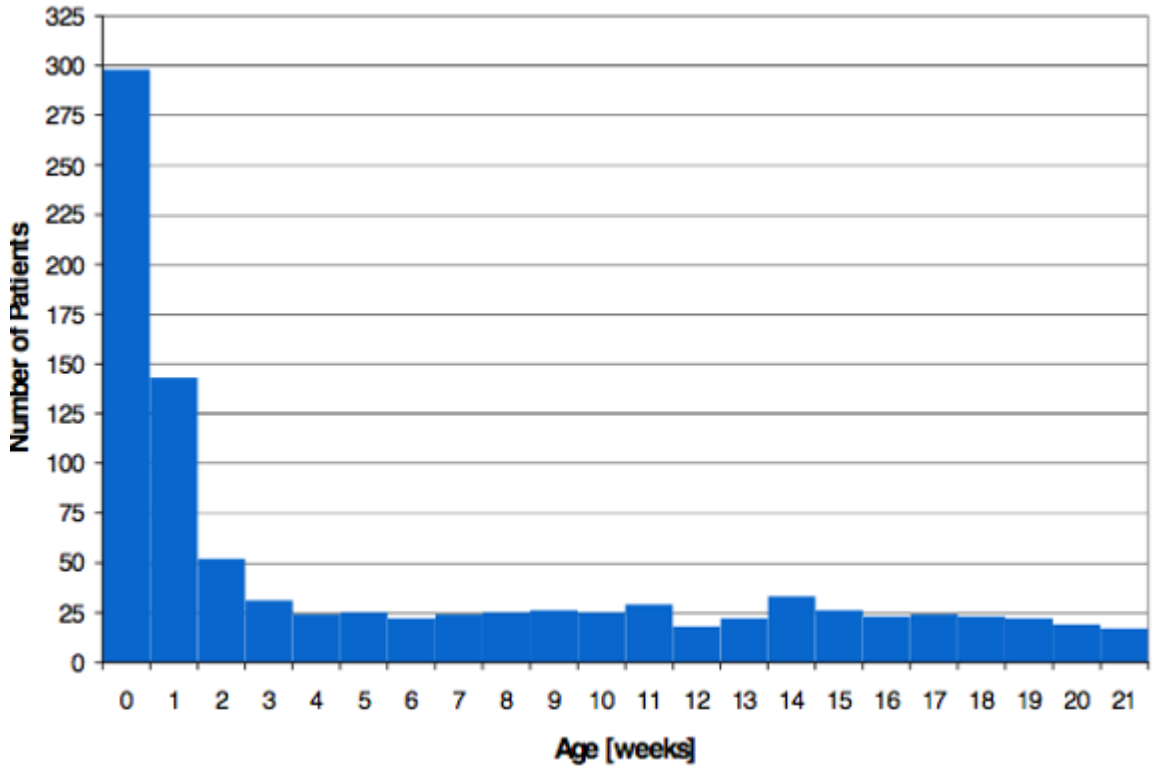
Process to find appropriate clinical imaging studies

N=~100,000
- Brain MRI in MGH

N = 2,871
- Scanned 2006-2013 with ADC maps in Siemens 3T scanner
- 0-6 years old at the time of scan
- Radiological reports suggest no abnormality

N = 1,648
- ADC maps found that are not corrupted

N = 705
- ADC maps and clinical cases re-examined & confirmed to be normal by a neuro-radiologist (Dr. Grant) and a neonatologist (Dr. Bates)

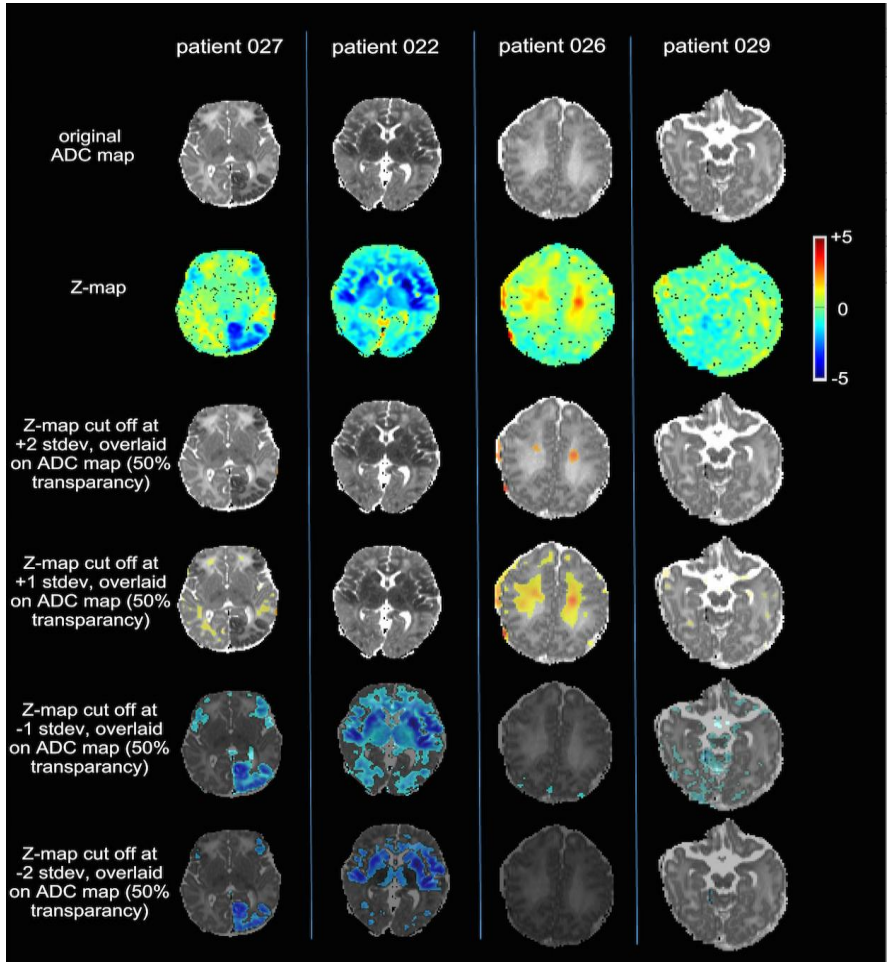


Ou Y, Zöllei L, Retzepi K, Castro V, Bates SV, Pieper S, Andriole KP, Murphy SN, Gollub RL, Grant PE. Using clinically acquired MRI to construct age-specific ADC atlases: Quantifying spatiotemporal ADC changes from birth to 6-year old. Hum Brain Mapp. 2017 Jun;38(6):3052-3068

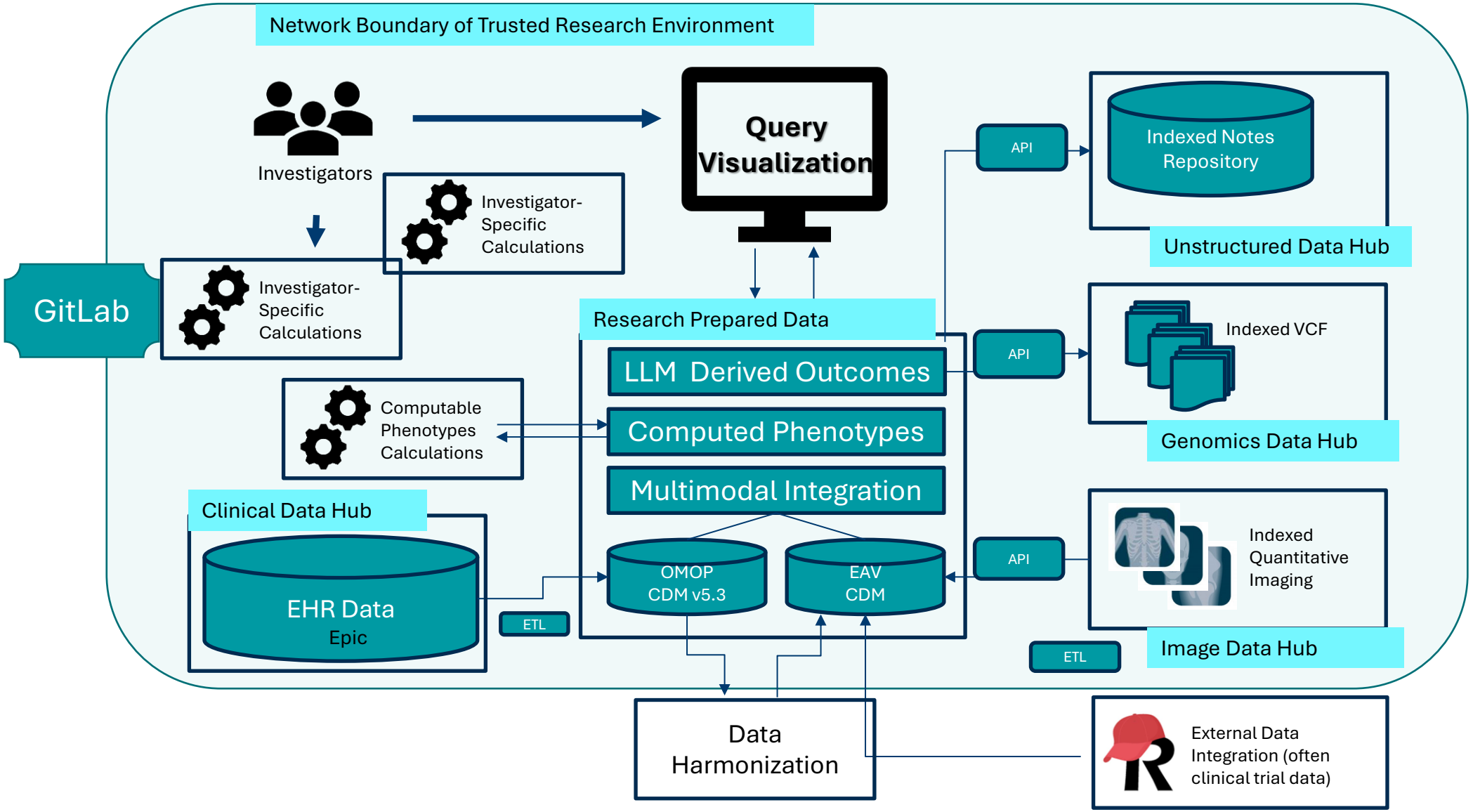
Atlases provide a visual guide for Radiology Decision Support, such as determining Perinatal Hypoxic Ischemic Encephalopathy

ADC map from 4 infants: Each statistically compared to age matched atlas yields visual guide to pathology

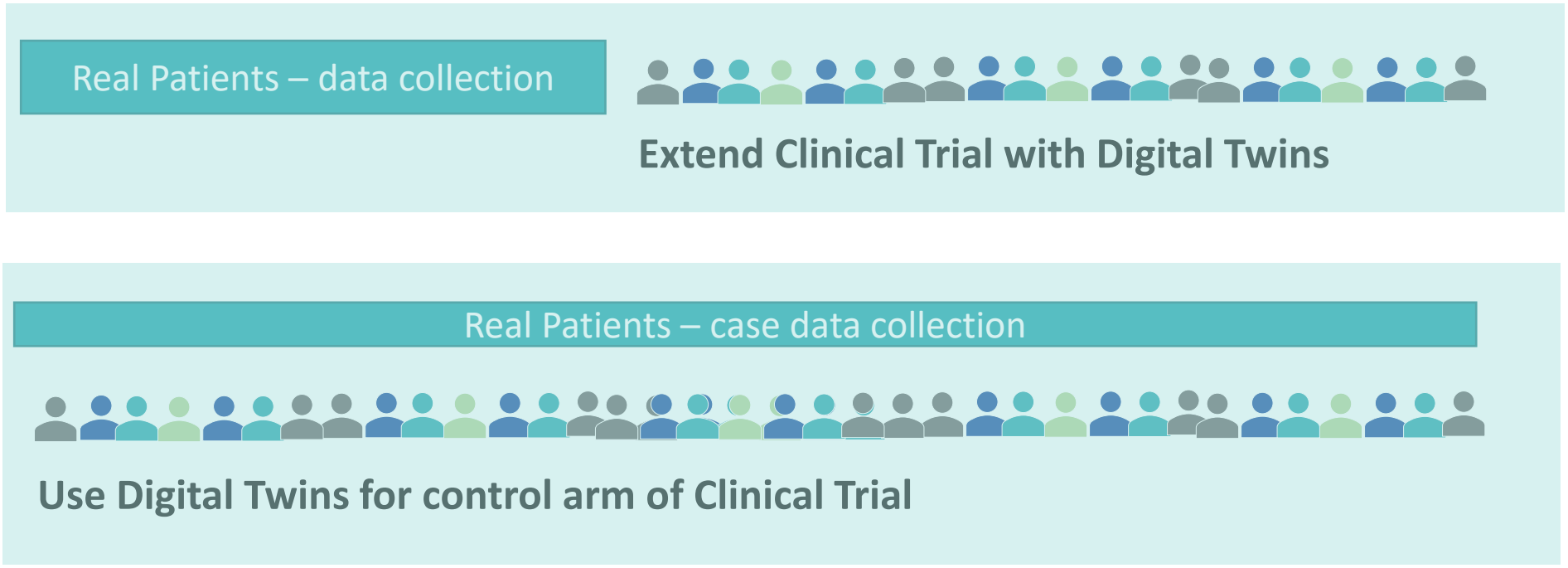
Quantitative analysis tools + large data sets = Great insights for practicing doctors



Sandbox for hosting Digital Twin



Digital Twins allow us to engage with Clinical Trials in new ways:



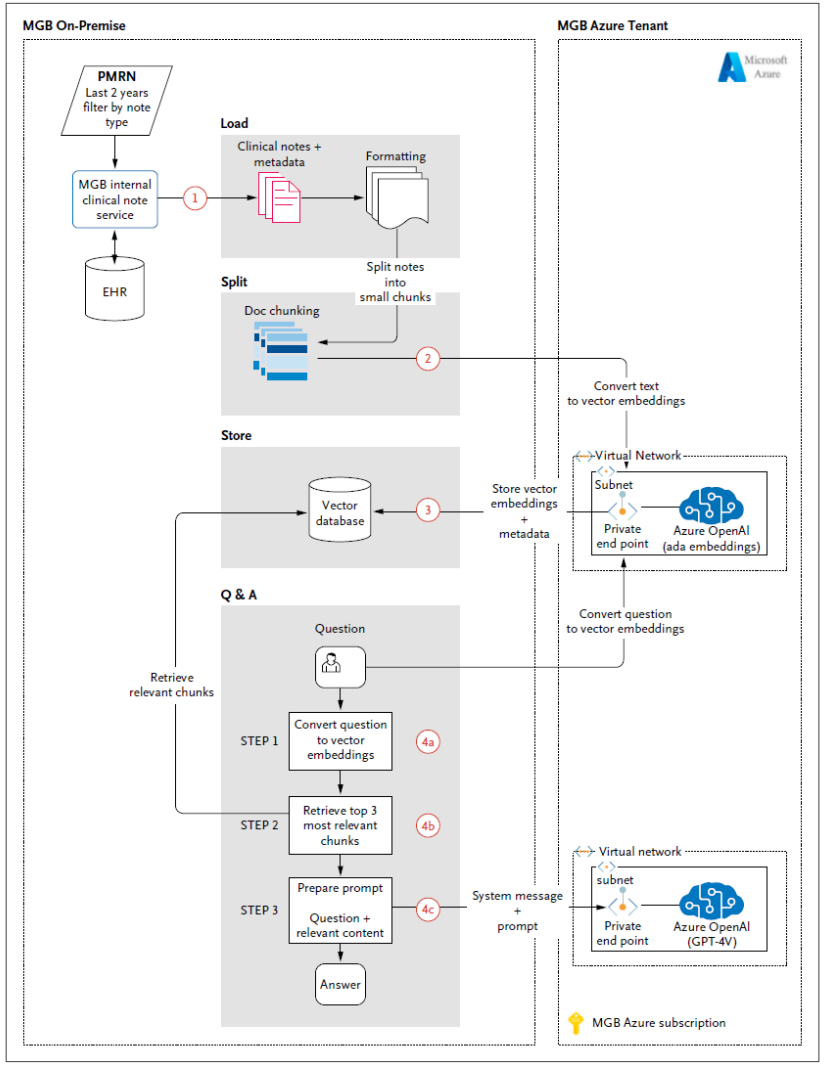
Optimized approach to finding the right patients

Why can't we just use LLMs?

- Cost of computation on millions of patients
- Slow compared to computed phenotypes
- Complex eligibility criteria combine many phenotypes and require multiple processing steps
- Uncertain performance on low prevalence diseases
- General LLM issues: hallucinations, values

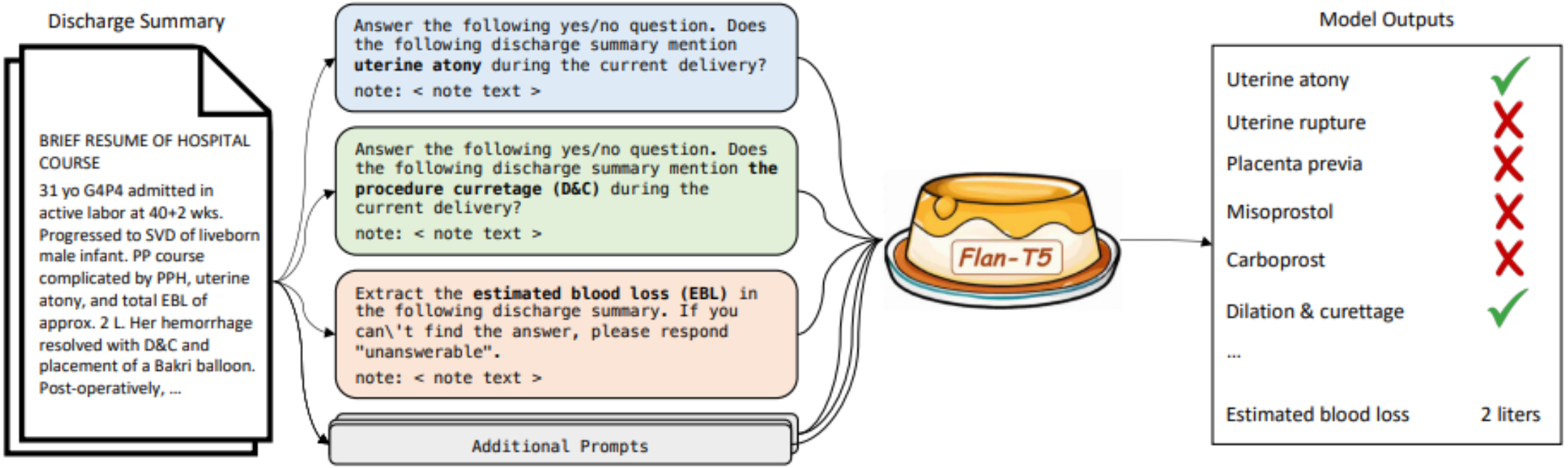
Combining computational phenotypes and LLMs:

- Use low-cost efficient computational phenotypes to quickly narrow population to an enriched cohort
- Refine the cohort using LLMs
- Use LLMs to improve the computed phenotypes



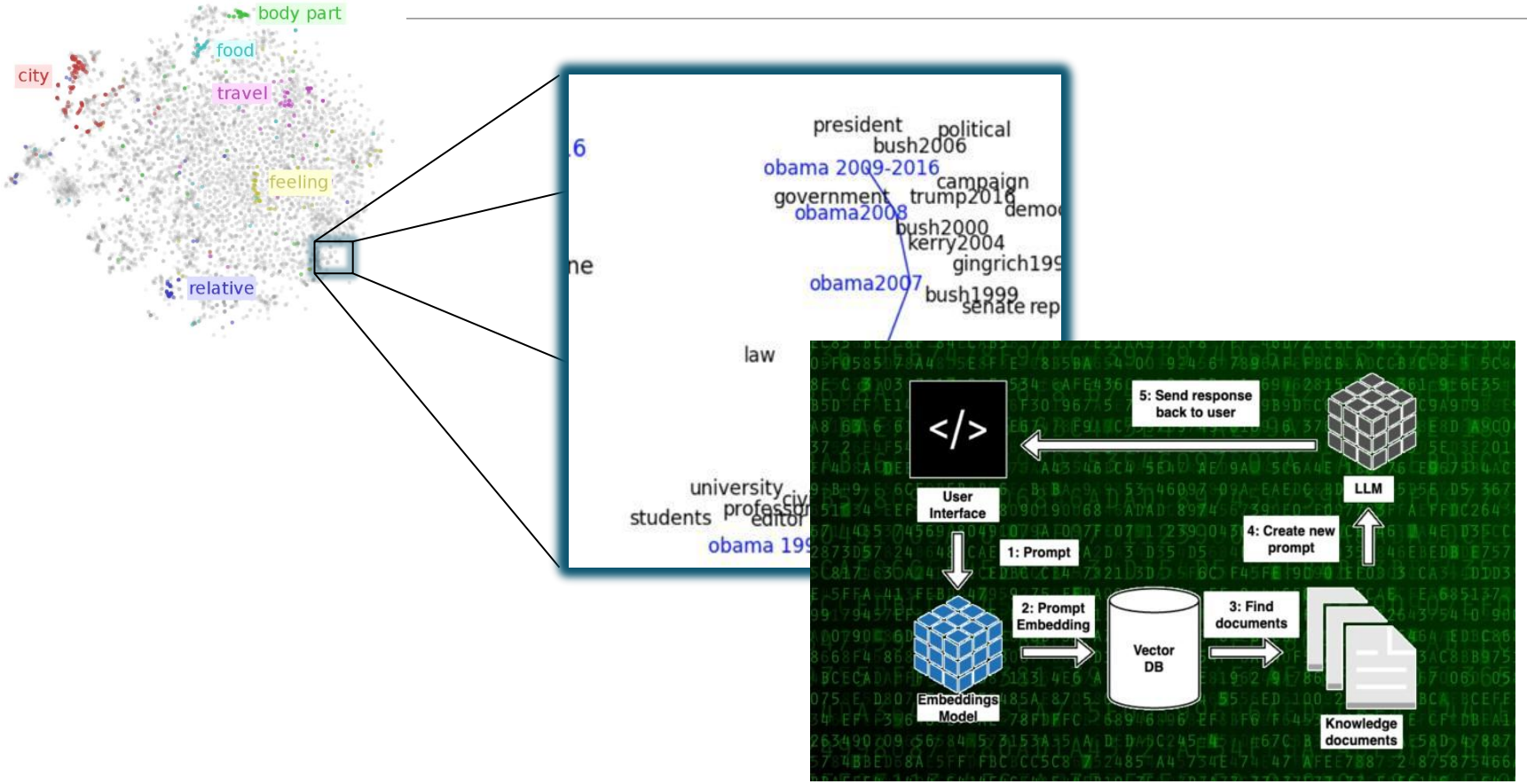
Unlu, et al. NEJM AI. 2024;1(7)

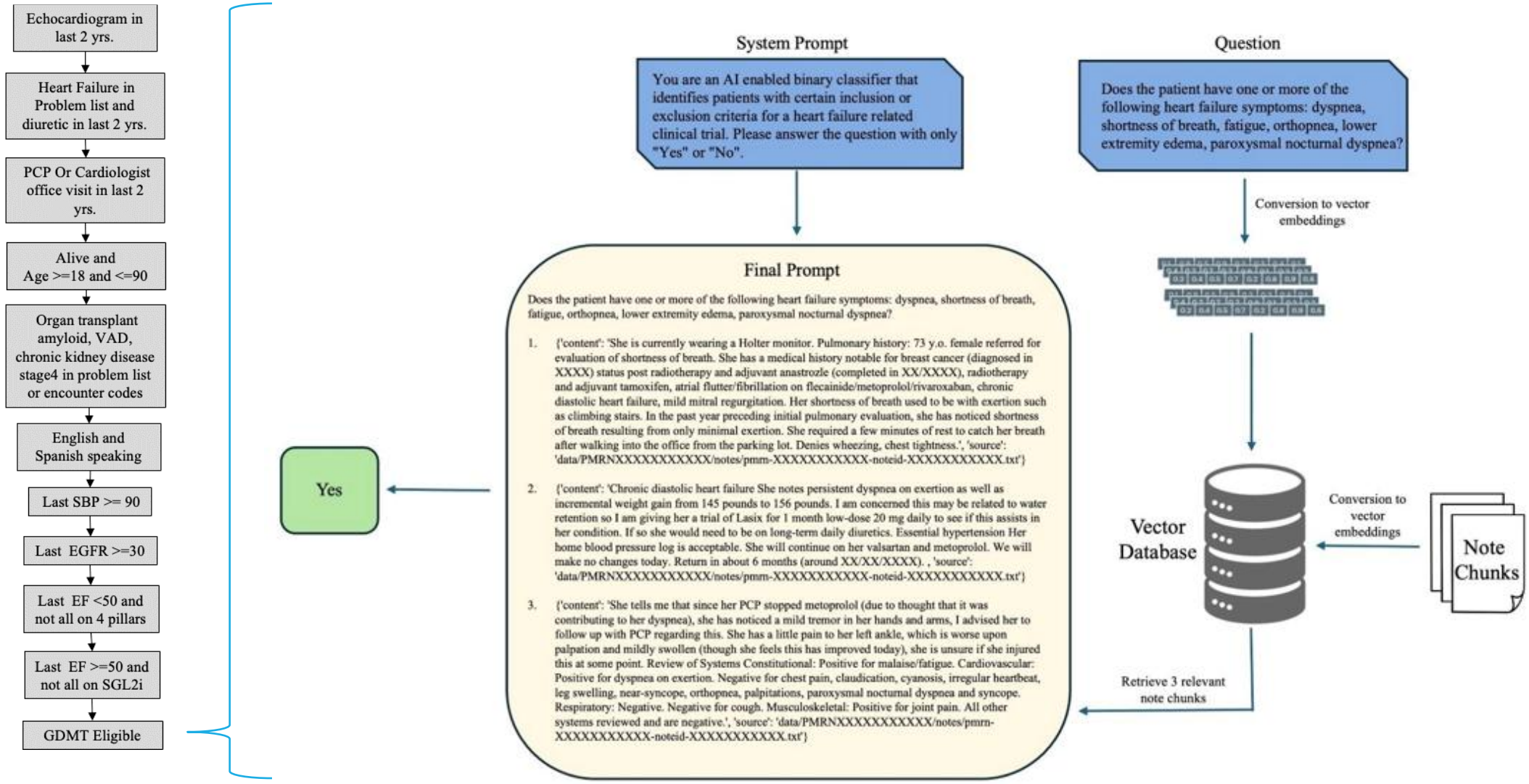
Chart Review supported by Large Language Models



Emily Alsentzer et al
Zero-shot Interpretable Phenotyping of Postpartum Hemorrhage Using Large Language Models
Nature PJ Digit. Med. **6**, 212 (2023).
https://doi.org/10.1038/s41746-023-00957-x

LLM Enhanced interaction with “RAG” Memory





NEJM AI

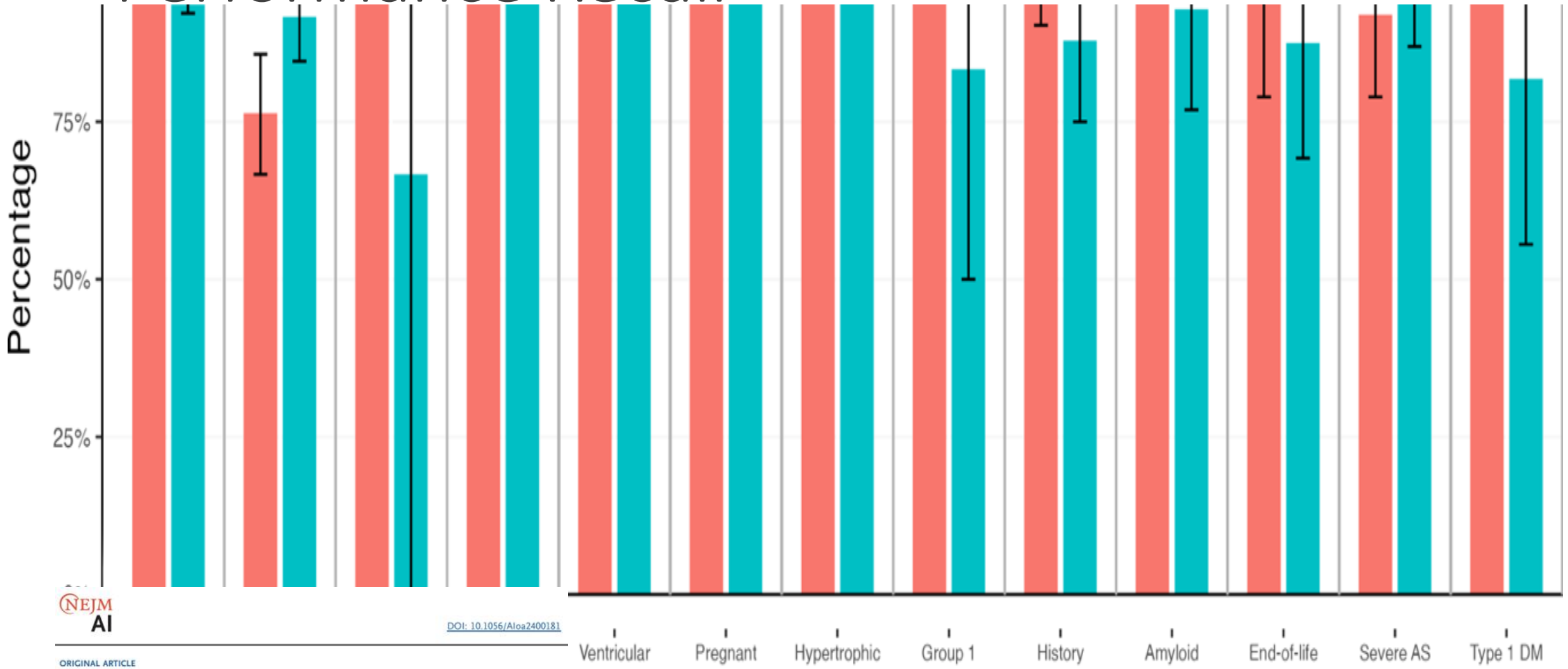
DOI: 10.1056/AIoa2400181

ORIGINAL ARTICLE

Retrieval-Augmented Generation–Enabled GPT-4 for Clinical Trial Screening

Ozan Unlu, M.D.,[1,2,3,4] Jiyeon Shin, A.L.M.,[1,5] Charlotte J. Mailly, B.A.,[1,5] Michael F. Oates,[1,5] Michela R. Tucci, M.P.H.,[1] Matthew Varugheese, M.S.,[1] Kavishwar Wagholikar, M.B.B.S., Ph.D.,[1,4] Fei Wang, M.Sc.,[1,5] Benjamin M. Scirica, M.D., M.P.H.,[1,2,4] Alexander J. Blood, M.D., M.Sc.,[1,2,4] and Samuel J. Aronson, A.L.M., M.A.[1,5]

Performance Recall



NEJM AI

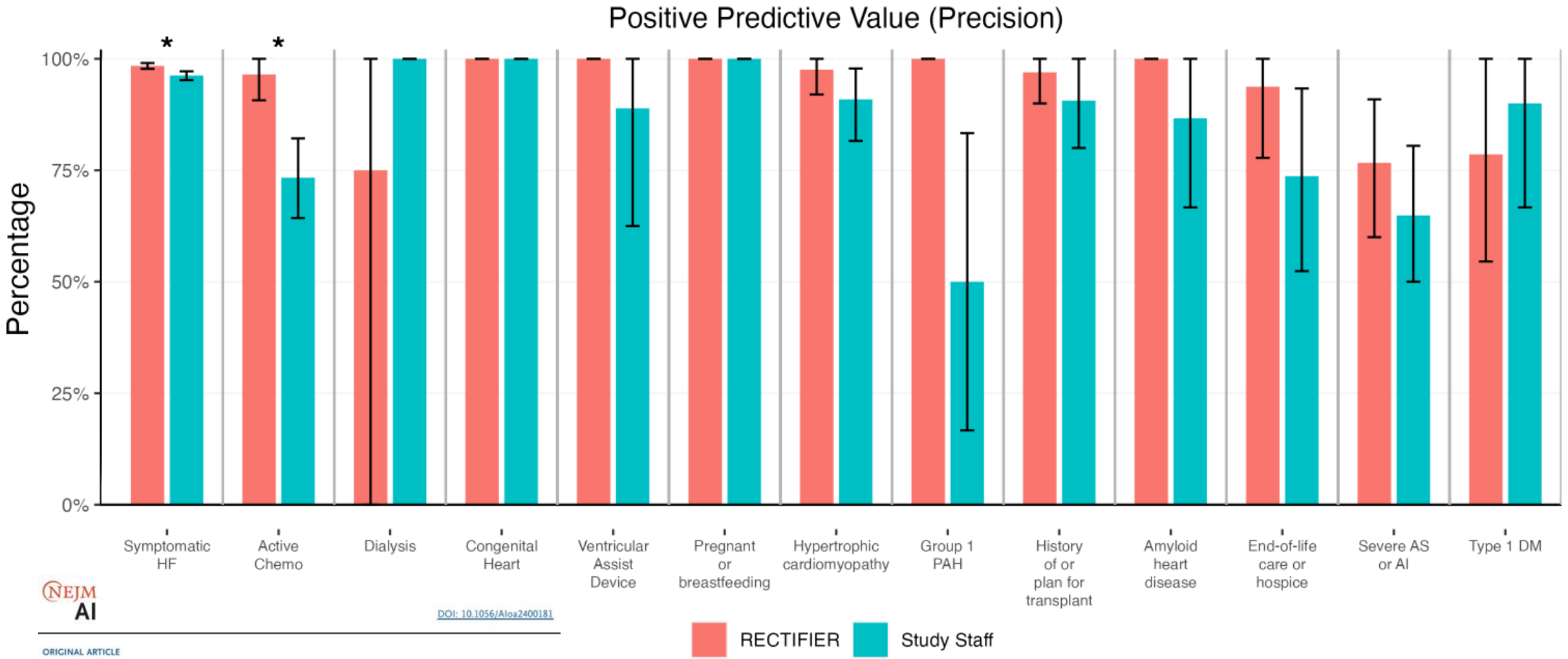
DOI: 10.1056/AIoa2400181

ORIGINAL ARTICLE

Retrieval-Augmented Generation–Enabled GPT-4 for Clinical Trial Screening

Ozan Unlu, M.D.,[1,2,3,4] Jiyeon Shin, A.L.M.,[1,5] Charlotte J. Mailly, B.A.,[1,5] Michael F. Oates,[1,5] Michela R. Tucci, M.P.H.,[1] Matthew Varugheese, M.S.,[1] Kavishwar Wagholikar, M.B.B.S., Ph.D.,[1,4] Fei Wang, M.Sc.,[1,5] Benjamin M. Scirica, M.D., M.P.H.,[1,2,4] Alexander J. Blood, M.D., M.Sc.,[1,2,4] and Samuel J. Aronson, A.L.M., M.A.[1,5]

Performance PPV



NEJM AI

DOI: 10.1056/AIoa2400181

ORIGINAL ARTICLE

Retrieval-Augmented Generation–Enabled GPT-4 for Clinical Trial Screening

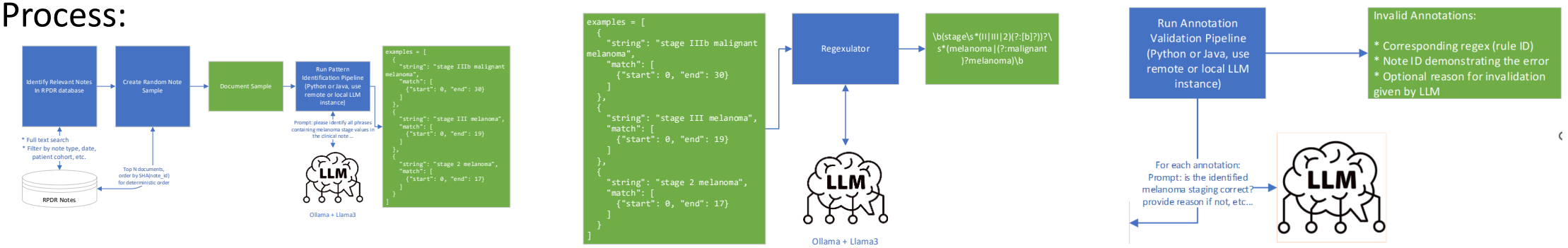
Ozan Unlu, M.D.,[1,2,3,4] Jiyeon Shin, A.L.M.,[1,5] Charlotte J. Mailly, B.A.,[1,5] Michael F. Oates,[1,5] Michela R. Tucci, M.P.H.,[1] Matthew Varugheese, M.S.,[1] Kavishwar Wagholikar, M.B.B.S., Ph.D.,[1,4] Fei Wang, M.Sc.,[1,5] Benjamin M. Scirica, M.D., M.P.H.,[1,2,4] Alexander J. Blood, M.D., M.Sc.,[1,2,4] and Samuel J. Aronson, A.L.M., M.A.[1,5]

Registry Creation with Agentic AI Processing (for truly Big Data)

Example: Extract cancer staging information in a timeline for each cancer patient

Method: Automate Process for using Regular Expressions to extract Staging information from Clinical/Pathology Notes

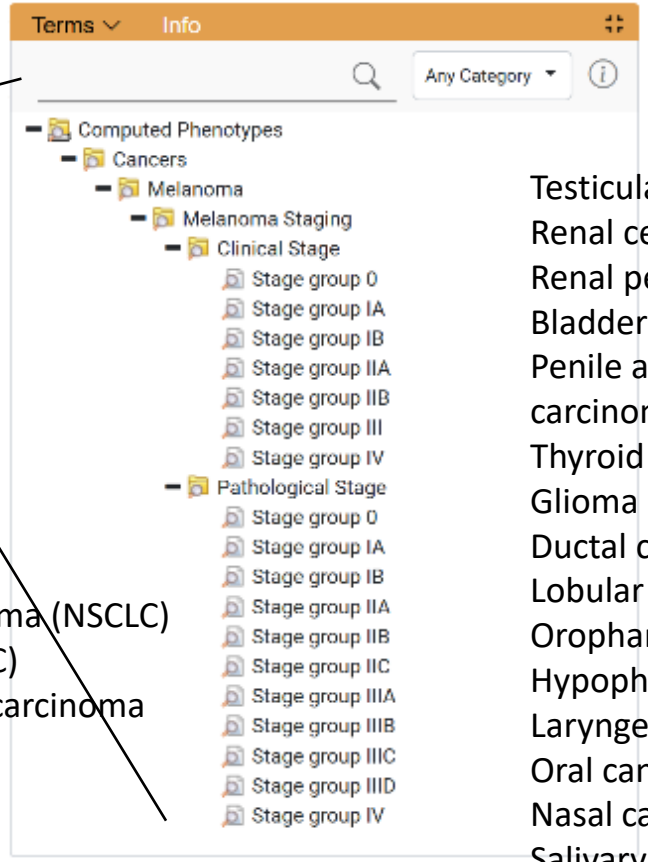
Process:



Outcome: Query-ready staging facts for Multiple Cancer Types

Cancer staging for multiple cancer types

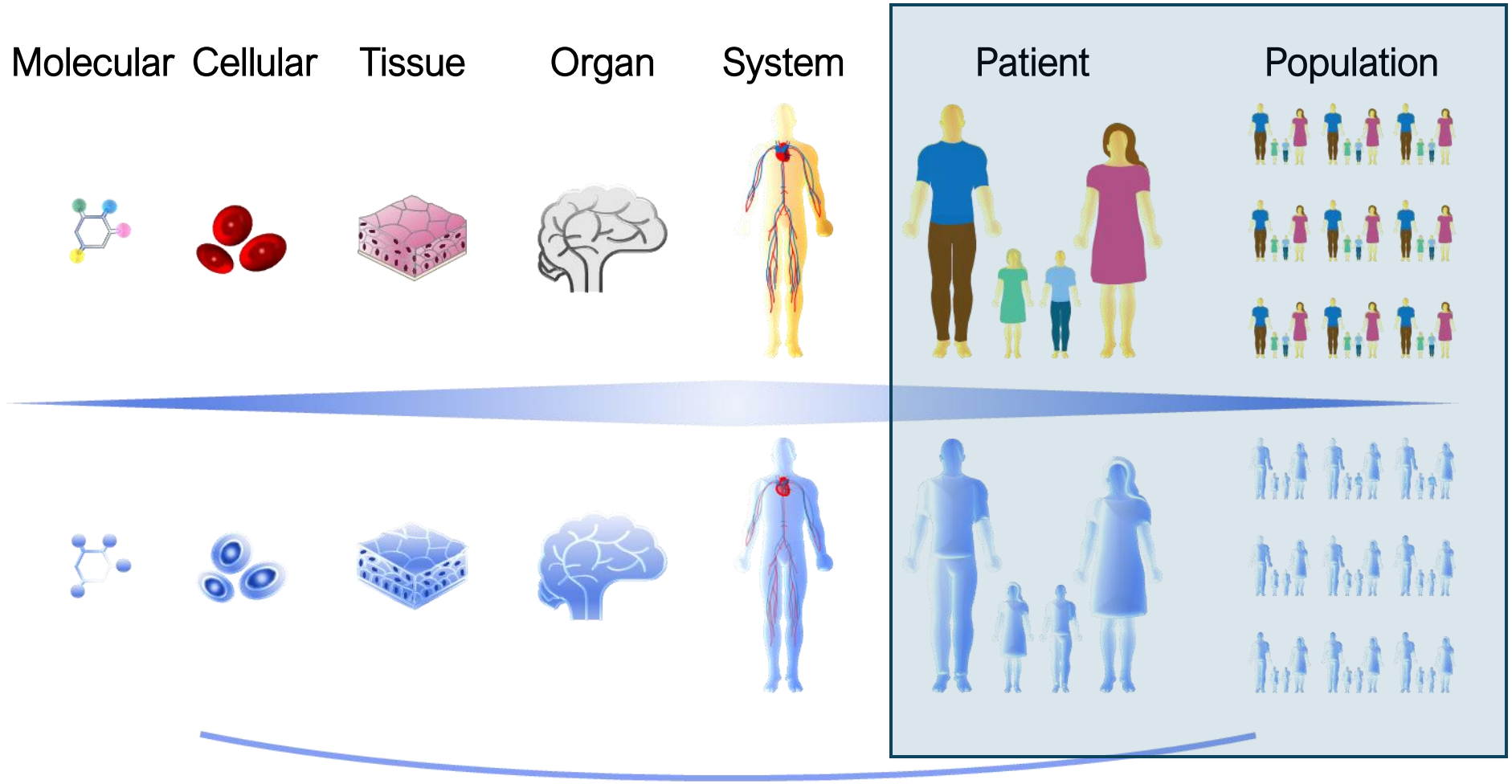
Melanoma
Breast cancer
Prostate cancer
Gastroesophageal cancers
Colorectal cancer
Anal cancer
Hepatocellular carcinoma
Pancreatic cancer
Small intestine cancer
Lung cancer (any histology)
Lung non-small cell carcinoma (NSCLC)
Lung small cell cancer (SCLC)
Large cell neuroendocrine carcinoma
Mesothelioma
Ovarian cancer
Endometrial cancer
Cervical cancer
Vulvovaginal cancer

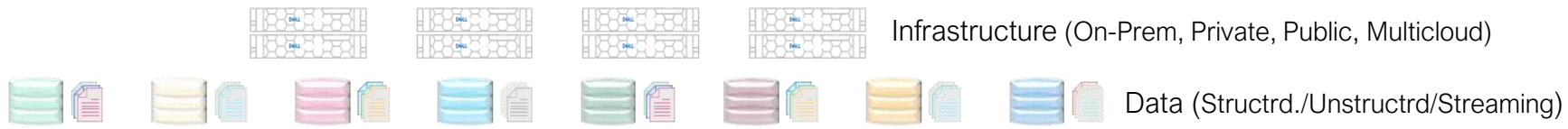


Testicular cancer
Renal cell carcinoma
Renal pelvic and ureter urothelial carcinoma
Bladder urothelial carcinoma
Penile and urethral squamous cell carcinoma
Thyroid cancer
Glioma
Ductal carcinoma in situ
Lobular carcinoma in situ
Oropharynx cancer
Hypopharynx cancer
Laryngeal cancer
Oral cancer
Nasal cavity and paranasal sinuses
Salivary gland cancers
Nasopharynx cancer
Ocular melanoma
Basal cell carcinoma
Squamous cell carcinoma
Merkel cell
Meningiomas

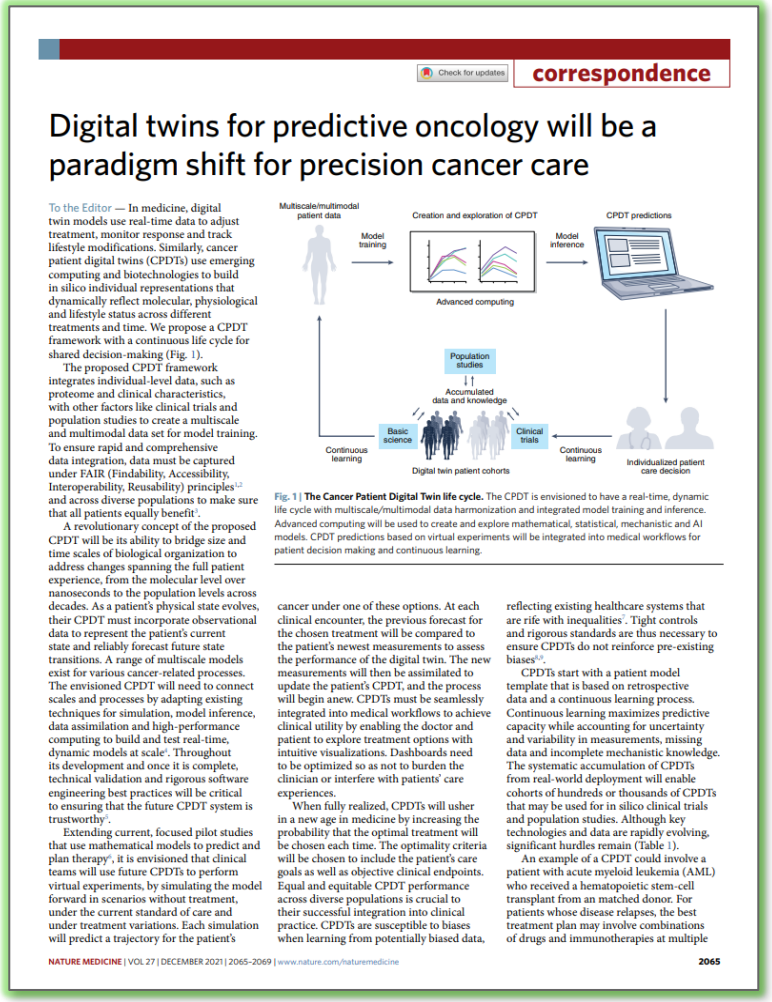
Gastrointestinal stromal tumor (GIST)
CNS lymphoma
Thymic cancer
Acute leukemia (AML/ALL)
Myelodysplastic syndrome
Chronic myeloid leukemia
Myeloproliferative neoplasms (MPN)
Chronic myelomonocytic leukemia (CMML)
Multiple myeloma
Waldenstrom’s macroglobulinemia
Hodgkin’s lymphoma
Biliary tract cancers
GI Neuroendocrine cancer
Non-Hodgkin's lymphoma
Soft tissue sarcoma
Bone cancer
Brain cancer (any histology)
Lymphoma (any histologic type)

Orchestrated Development of Digital Twin





Opportunities for Condition Specific Data Diversity



correspondence

Digital twins for predictive oncology will be a paradigm shift for precision cancer care

To the Editor — In medicine, digital twin models use real-time data to adjust treatment, monitor response and track lifestyle modifications. Similarly, cancer patient digital twins (CPDTs) use emerging computing and biotechnologies to build in silico individual representations that dynamically reflect molecular, physiological and lifestyle status across different treatments and time. We propose a CPDT framework with a continuous life cycle for shared decision-making (Fig. 1).

The proposed CPDT framework integrates individual-level data, such as proteome and clinical characteristics, with other factors like clinical trials and population studies to create a multiscale and multimodal data set for model training. To ensure rapid and comprehensive data integration, data must be captured under FAIR (Findability, Accessibility, Interoperability, Reusability) principles[1,2] and across diverse populations to make sure that all patients equally benefit[3].

A revolutionary concept of the proposed CPDT will be its ability to bridge size and time scales of biological organization to address changes spanning the full patient experience, from the molecular level over nanoseconds to the population levels across decades. As a patient's physical state evolves, their CPDT must incorporate observational data to represent the patient's current state and reliably forecast future state transitions. A range of multiscale models exist for various cancer-related processes. The envisioned CPDT will need to connect scales and processes by adapting existing techniques for simulation, model inference, data assimilation and high-performance computing to build and test real-time, dynamic models at scale[4]. Throughout its development and once it is complete, technical validation and rigorous software engineering best practices will be critical to ensuring that the future CPDT system is trustworthy[5].

Extending current, focused pilot studies that use mathematical models to predict and plan therapy[6], it is envisioned that clinical teams will use future CPDTs to perform virtual experiments, by simulating the model forward in scenarios without treatment, under the current standard of care and under treatment variations. Each simulation will predict a trajectory for the patient's cancer under one of these options. At each clinical encounter, the previous forecast for the chosen treatment will be compared to the patient's newest measurements to assess the performance of the digital twin. The new measurements will then be assimilated to update the patient's CPDT, and the process will begin anew. CPDTs must be seamlessly integrated into medical workflows to achieve clinical utility by enabling the doctor and patient to explore treatment options with intuitive visualizations. Dashboards need to be optimized so as not to burden the clinician or interfere with patients' care experiences.

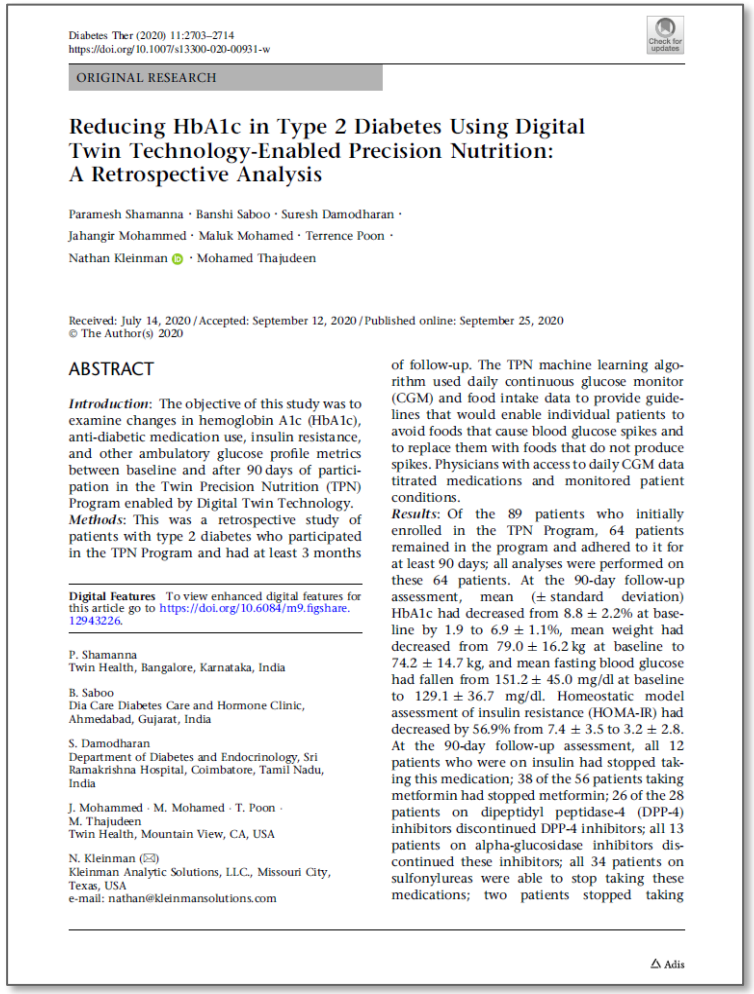
When fully realized, CPDTs will usher in a new age in medicine by increasing the probability that the optimal treatment will be chosen each time. The optimality criteria will be chosen to include the patient's care goals as well as objective clinical endpoints. Equal and equitable CPDT performance across diverse populations is crucial to their successful integration into clinical practice. CPDTs are susceptible to biases when learning from potentially biased data, reflecting existing healthcare systems that are rife with inequalities[7]. Tight controls and rigorous standards are thus necessary to ensure CPDTs do not reinforce pre-existing biases[8,9].

CPDTs start with a patient model template that is based on retrospective data and a continuous learning process. Continuous learning maximizes predictive capacity while accounting for uncertainty and variability in measurements, missing data and incomplete mechanistic knowledge. The systematic accumulation of CPDTs from real-world deployment will enable cohorts of hundreds or thousands of CPDTs that may be used for in silico clinical trials and population studies. Although key technologies and data are rapidly evolving, significant hurdles remain (Table 1).

An example of a CPDT could involve a patient with acute myeloid leukemia (AML) who received a hematopoietic stem-cell transplant from an matched donor. For patients whose disease relapses, the best treatment plan may involve combinations of drugs and immunotherapies at multiple

Fig. 1 | **The Cancer Patient Digital Twin life cycle.** The CPDT is envisioned to have a real-time, dynamic life cycle with multiscale/multimodal data harmonization and integrated model training and inference. Advanced computing will be used to create and explore mathematical, statistical, mechanistic and AI models. CPDT predictions based on virtual experiments will be integrated into medical workflows for patient decision making and continuous learning.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9097784/



Diabetes Ther (2020) 11:2703–2714
https://doi.org/10.1007/s13300-020-00931-w

ORIGINAL RESEARCH

Reducing HbA1c in Type 2 Diabetes Using Digital Twin Technology-Enabled Precision Nutrition: A Retrospective Analysis

Paramesh Shamanna · Banshi Saboo · Suresh Damodharan · Jahangir Mohammed · Maluk Mohamed · Terrence Poon · Nathan Kleinman · Mohamed Thajudeen

Received: July 14, 2020 / Accepted: September 12, 2020 / Published online: September 25, 2020
© The Author(s) 2020

ABSTRACT

Introduction: The objective of this study was to examine changes in hemoglobin A1c (HbA1c), anti-diabetic medication use, insulin resistance, and other ambulatory glucose profile metrics between baseline and after 90 days of participation in the Twin Precision Nutrition (TPN) Program enabled by Digital Twin Technology.

Methods: This was a retrospective study of patients with type 2 diabetes who participated in the TPN Program and had at least 3 months of follow-up. The TPN machine learning algorithm used daily continuous glucose monitor (CGM) and food intake data to provide guidelines that would enable individual patients to avoid foods that cause blood glucose spikes and to replace them with foods that do not produce spikes. Physicians with access to daily CGM data titrated medications and monitored patient conditions.

Results: Of the 89 patients who initially enrolled in the TPN Program, 64 patients remained in the program and adhered to it for at least 90 days; all analyses were performed on these 64 patients. At the 90-day follow-up assessment, mean (± standard deviation) HbA1c had decreased from 8.8 ± 2.2% at baseline by 1.9 to 6.9 ± 1.1%, mean weight had decreased from 79.0 ± 16.2 kg at baseline to 74.2 ± 14.7 kg, and mean fasting blood glucose had fallen from 151.2 ± 45.0 mg/dl at baseline to 129.1 ± 36.7 mg/dl. Homeostatic model assessment of insulin resistance (HOMA-IR) had decreased by 56.9% from 7.4 ± 3.5 to 3.2 ± 2.8. At the 90-day follow-up assessment, all 12 patients who were on insulin had stopped taking this medication; 38 of the 56 patients taking metformin had stopped metformin; 26 of the 28 patients on dipeptidyl peptidase-4 (DPP-4) inhibitors discontinued DPP-4 inhibitors; all 13 patients on alpha-glucosidase inhibitors discontinued these inhibitors; all 34 patients on sulfonylureas were able to stop taking these medications; two patients stopped taking

Digital Features To view enhanced digital features for this article go to https://doi.org/10.6084/m9.figshare.12943226.

P. Shamanna
Twin Health, Bangalore, Karnataka, India

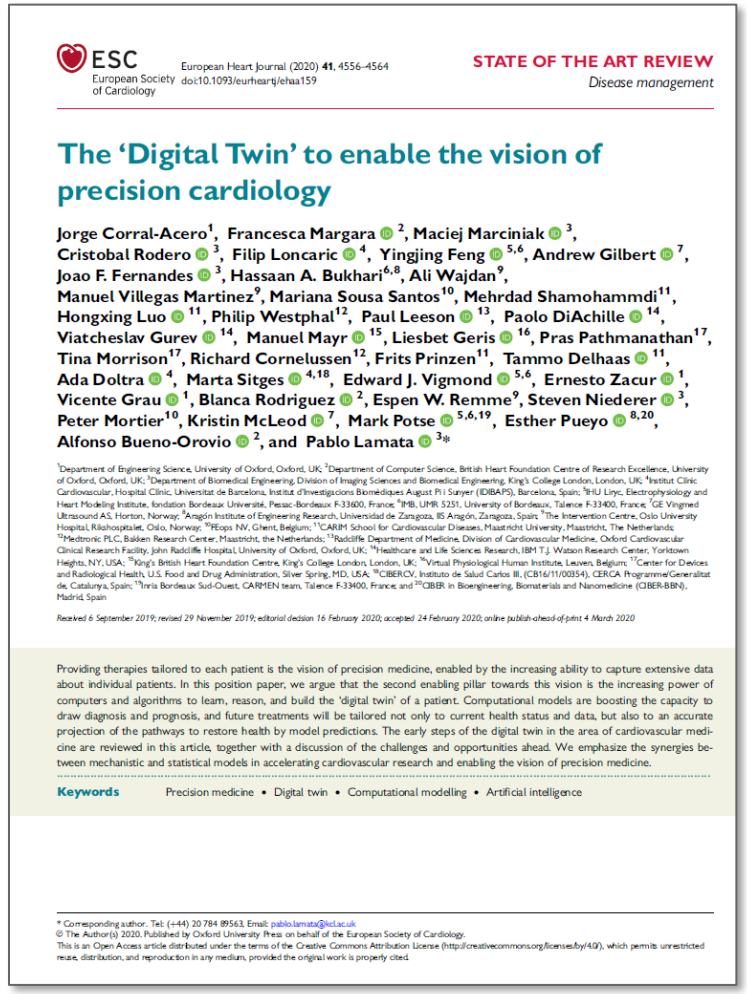
B. Saboo
Dia Care Diabetes Care and Hormone Clinic, Ahmedabad, Gujarat, India

S. Damodharan
Department of Diabetes and Endocrinology, Sri Ramakrishna Hospital, Coimbatore, Tamil Nadu, India

J. Mohammed · M. Mohamed · T. Poon · M. Thajudeen
Twin Health, Mountain View, CA, USA

N. Kleinman (✉)
Kleinman Analytic Solutions, LLC., Missouri City, Texas, USA
e-mail: nathan@kleinmansolutions.com

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7547935/pdf/13300_2020_Article_931.pdf



ESC European Society of Cardiology
European Heart Journal (2020) 41, 4556–4564
doi:10.1093/eurheartj/ehaa159

STATE OF THE ART REVIEW
Disease management

The 'Digital Twin' to enable the vision of precision cardiology

Jorge Corral-Acero[1], Francesca Margara[2], Maciej Marciniak[3], Cristobal Rodero[3], Filip Loncaric[4], Yingjing Feng[5,6], Andrew Gilbert[7], Joao F. Fernandes[3], Hassaan A. Bukhari[6,8], Ali Wajdan[9], Manuel Villegas Martinez[9], Mariana Sousa Santos[10], Mehrdad Shamohammdi[11], Hongxing Luo[11], Philip Westphal[12], Paul Leeson[13], Paolo DiAchille[14], Viatcheslav Gurev[14], Manuel Mayr[15], Liesbet Geris[16], Pras Pathmanathan[17], Tina Morrison[17], Richard Cornelussen[12], Frits Prinzen[11], Tammo Delhaas[11], Ada Doltra[4], Marta Sitges[4,18], Edward J. Vigmond[5,6], Ernesto Zacur[1], Vicente Grau[1], Blanca Rodriguez[2], Espen W. Remme[9], Steven Niederer[3], Peter Mortier[10], Kristin McLeod[7], Mark Potse[5,6,19], Esther Pueyo[8,20], Alfonso Bueno-Orovio[2], and Pablo Lamata[3]*

[1]Department of Engineering Science, University of Oxford, Oxford, UK; [2]Department of Computer Science, British Heart Foundation Centre of Research Excellence, University of Oxford, Oxford, UK; [3]Department of Biomedical Engineering, Division of Imaging Sciences and Biomedical Engineering, King's College London, London, UK; [4]Institut Clinic Cardiovascular, Hospital Clinic, Universitat de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain; [5]IHU Liryc, Electrophysiology and Heart Modeling Institute, fondation Bordeaux Université, Pessac-Bordeaux F-33600, France; [6]IMB, UMR 5251, University of Bordeaux, Talence F-33400, France; [7]GE Vingmed Ultrasound AS, Horton, Norway; [8]Aragón Institute of Engineering Research, Universidad de Zaragoza, IIS Aragón, Zaragoza, Spain; [9]The Intervention Centre, Oslo University Hospital, Rikshospitalet, Oslo, Norway; [10]Ecops NV, Ghent, Belgium; [11]CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht, The Netherlands; [12]Medtronic PLC, Bakken Research Center, Maastricht, the Netherlands; [13]Radcliffe Department of Medicine, Division of Cardiovascular Medicine, Oxford Cardiovascular Clinical Research Facility, John Radcliffe Hospital, University of Oxford, Oxford, UK; [14]Healthcare and Life Sciences Research, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA; [15]King's British Heart Foundation Centre, King's College London, London, UK; [16]Virtual Physiological Human Institute, Leuven, Belgium; [17]Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, USA; [18]CIBERCV, Instituto de Salud Carlos III, (CB16/11/00354), CERCA Programme/Generalitat de, Catalunya, Spain; [19]Inria Bordeaux Sud-Ouest, CARMEN team, Talence F-33400, France; and [20]CIBER in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain

Received 6 September 2019; revised 29 November 2019; editorial decision 16 February 2020; accepted 24 February 2020; online publish-ahead-of-print 4 March 2020

Providing therapies tailored to each patient is the vision of precision medicine, enabled by the increasing ability to capture extensive data about individual patients. In this position paper, we argue that the second enabling pillar towards this vision is the increasing power of computers and algorithms to learn, reason, and build the 'digital twin' of a patient. Computational models are boosting the capacity to draw diagnosis and prognosis, and future treatments will be tailored not only to current health status and data, but also to an accurate projection of the pathways to restore health by model predictions. The early steps of the digital twin in the area of cardiovascular medicine are reviewed in this article, together with a discussion of the challenges and opportunities ahead. We emphasize the synergies between mechanistic and statistical models in accelerating cardiovascular research and enabling the vision of precision medicine.

Keywords Precision medicine • Digital twin • Computational modelling • Artificial intelligence

* Corresponding author. Tel: (+44) 20 784 89563, Email: pablo.lamata@kcl.ac.uk
© The Author(s) 2020. Published by Oxford University Press on behalf of the European Society of Cardiology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7774470/pdf/ehaa159.pdf

Opportunities for Multiplanar Data Depth

npj | digital medicine

www.nature.com/npjdigitalmed

PERSPECTIVE OPEN

Check for updates

Building digital twins of the human immune system: toward a roadmap

R. Laubenbacher[1✉], A. Niarakis[2,3], T. Helikar[4], G. An[5], B. Shapiro[1], R. S. Malik-Sheriff[6], T. J. Sego[7], A. Knapp[1], P. Macklin[7] and J. A. Glazier[7]

Digital twins, customized simulation models pioneered in industry, are beginning to be deployed in medicine and healthcare, with some major successes, for instance in cardiovascular diagnostics and in insulin pump control. Personalized computational models are also assisting in applications ranging from drug development to treatment optimization. More advanced medical digital twins will be essential to making precision medicine a reality. Because the immune system plays an important role in such a wide range of diseases and health conditions, from fighting pathogens to autoimmune disorders, digital twins of the immune system will have an especially high impact. However, their development presents major challenges, stemming from the inherent complexity of the immune system and the difficulty of measuring many aspects of a patient's immune state in vivo. This perspective outlines a roadmap for meeting these challenges and building a prototype of an immune digital twin. It is structured as a four-stage process that proceeds from a specification of a concrete use case to model constructions, personalization, and continued improvement.

npj Digital Medicine (2022)5:64 ; https://doi.org/10.1038/s41746-022-00610-z

INTRODUCTION

Today, nearly every industry that deals with complex technologies use sophisticated "digital twin" computer simulations to forecast how individual pieces of equipment should perform under ever-changing real-world conditions. For instance, a digital twin of a jet engine receives continuously updated operational data streamed over the "internet of things" while the engine is in flight. Any deviation between the digital twin's prediction and the actual engine's state can provide an early warning of a potential problem, which can be resolved before it becomes critical/catastrophic.

The medical analog is only in its infancy. Medical digital twins could find many uses and take many forms. They could predict disease trajectories in individual patients, allowing diagnosis before the onset of serious symptoms. They could be used to optimize the timing of suggested medical care and to investigate the effects of potential treatments in a patient-tailored manner. They could help identify biomarkers or elucidate drug mechanisms of action. They could be data-driven or based on mechanistic computational models of biological function, or a combination of both. Currently, digital twins of the human heart improve diagnosis, prognosis, and therapies[1,2]. Automated workflows for generating cardiac digital twins could serve as a blueprint for the generation of other types of medical digital twins. Another example of an operational medical digital twin is the artificial pancreas that aids Type I diabetic patients in insulin management[3]. Of particular importance are digital twins that capture key features of the immune system, with its ubiquitous influence on human disease. Being able to predict an individual's immune response to infection or injury could be lifesaving in many ways, for instance in designing optimal individualized treatments[4].

This perspective describes a high-level outline of a roadmap for building an immune digital twin (IDT), depicted in Fig. 1. It advocates a partnership between the international not-for-profit scientific community, the commercial sector, and government funding agencies. Such an approach requires a collaborative infrastructure and broad agreement on the nature of the final product. For instance, the recently formed European-led DigiTwin consortium, which includes academic, clinical, and industrial partners from 32 countries, aims to create digital twins for every European citizen for a range of conditions[5], and could become a model for some aspects of a global consortium for immune digital twins.

DEVELOPMENT STAGES

The construction of an immune digital twin will broadly follow the four-stage paradigm for industrial digital twin development (see Fig. 2).

Stage 1: Define a specific application and construct an appropriate generic template model.

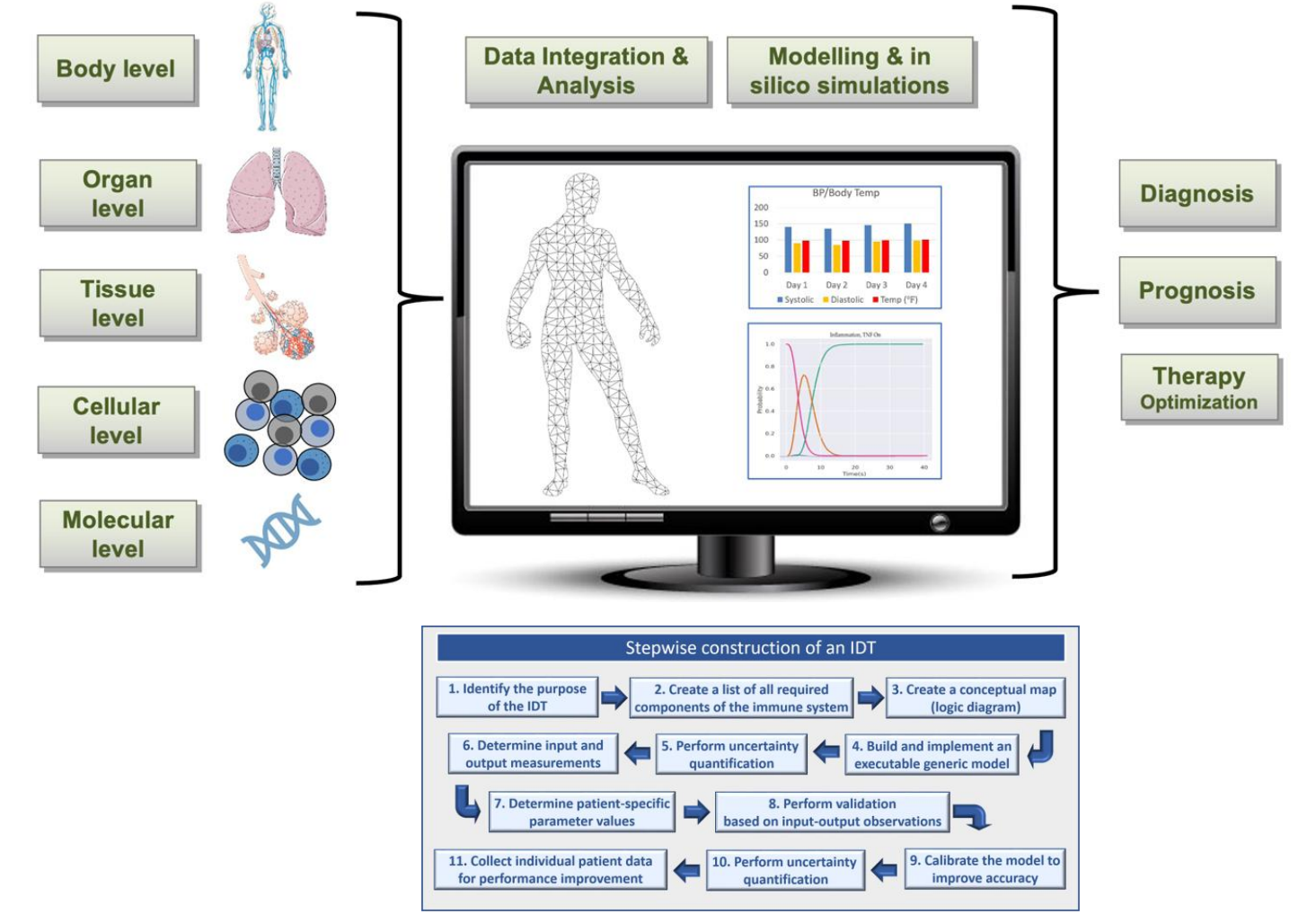
The crucial first step is to determine a specific application of a personalized immune simulation model. Applications can be "online," e.g., to predict the efficacy of a particular treatment based on frequent measurements of a patient's condition, or "offline," e.g., to develop novel drugs using simulated patient populations. Defining the application helps to narrow down the components of the immune system to include in the model, and to define the data available for model construction and operation, resulting in a generic "specification" of the digital twin design.

As there are many patient-specific immune system measurements that are currently infeasible, the choice of application is crucial for success.

[1]Department of Medicine, University of Florida, Gainesville, FL, USA. [2]Université Paris-Saclay, Laboratoire Européen de Recherche pour la Polyarthrite rhumatoïde - Genhotel, Univ Evry, Evry, France. [3]Lifeware Group, Inria, Saclay-Ile de France, 91120 Palaiseau, France. [4]Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, USA. [5]Department of Surgery, Larner College of Medicine, University of Vermont, Burlington, VT, USA. [6]European Bioinformatics Institute, European Molecular Biology Laboratory (EMBL-EBI), Hinxton, Cambridge, UK. [7]Biocomplexity Institute and Department of Intelligent Systems Engineering, Indiana University, Bloomington, IN, USA.
✉email: reinhard.laubenbacher@medicine.ufl.edu

Published in partnership with Seoul National University Bundang Hospital

npj

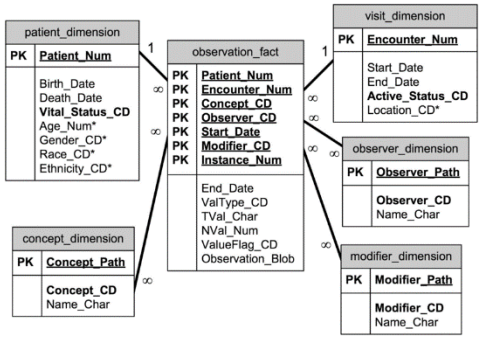


https://www.nature.com/articles/s41746-022-00610-z

Laubenbacher, R., Niarakis, A., Helikar, T. *et al.* Building digital twins of the human immune system: toward a roadmap. *npj Digit. Med.* **5**, 64 (2022). https://doi.org/10.1038/s41746-022-00610-z

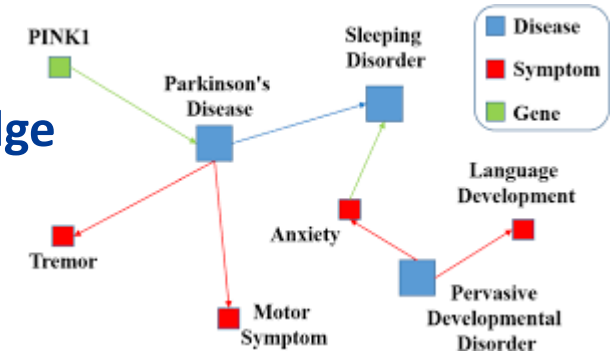
Data Architectures for Integration across Datasets

Star Schema for FACTS



- Ensure all data is in same time dimension
- Diverse concepts representable
- Progressively higher accuracy overlays in same structure
- Very fast lookup with indexes
- Memory for prior extraction by LLM
- Efficient use of context window

Graph DB for Knowledge



- Explicit Relationships
- Context Preservation
- Hallucination Reduction
- Multi-Hop Reasoning
- Deterministic Retrieval
- Dynamic Schema
- Explainable AI

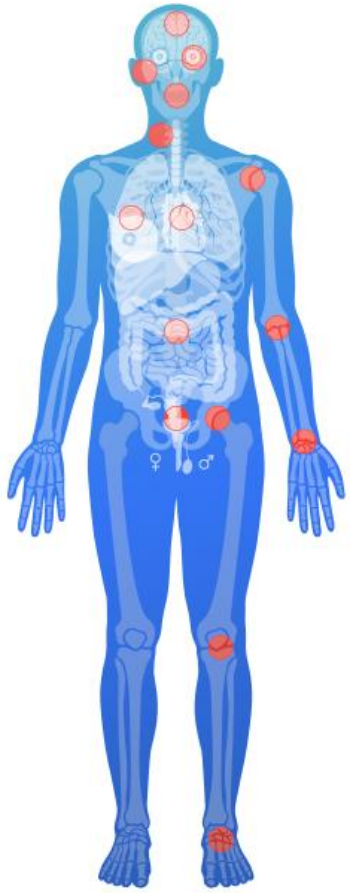
Trust in multi-modal Digital Twin unlocks discovery

Computed Phenotypes

Statistically validated definitions accurately identify medical conditions

Clinical Context

Understanding metrics and measures taken in clinical workflows



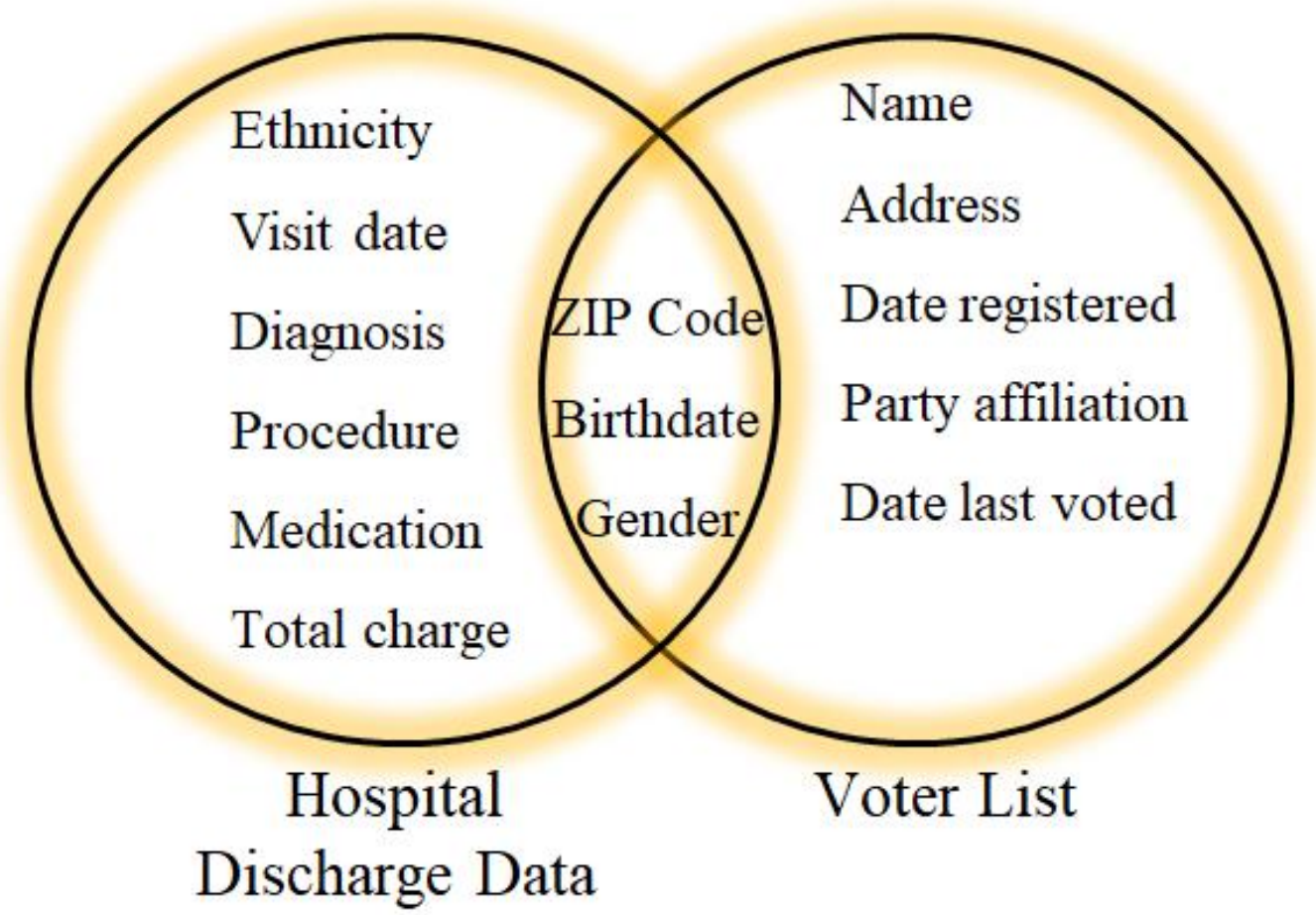
Multi-modal Data

Merge structured EHR data with unstructured notes, imaging, genomics, etc.

Gold Standards

Creation and validation of gold standards led by on staff clinicians and enabled by LLMs

What Are We Afraid Of – Has AI introduced Greater Risk ?



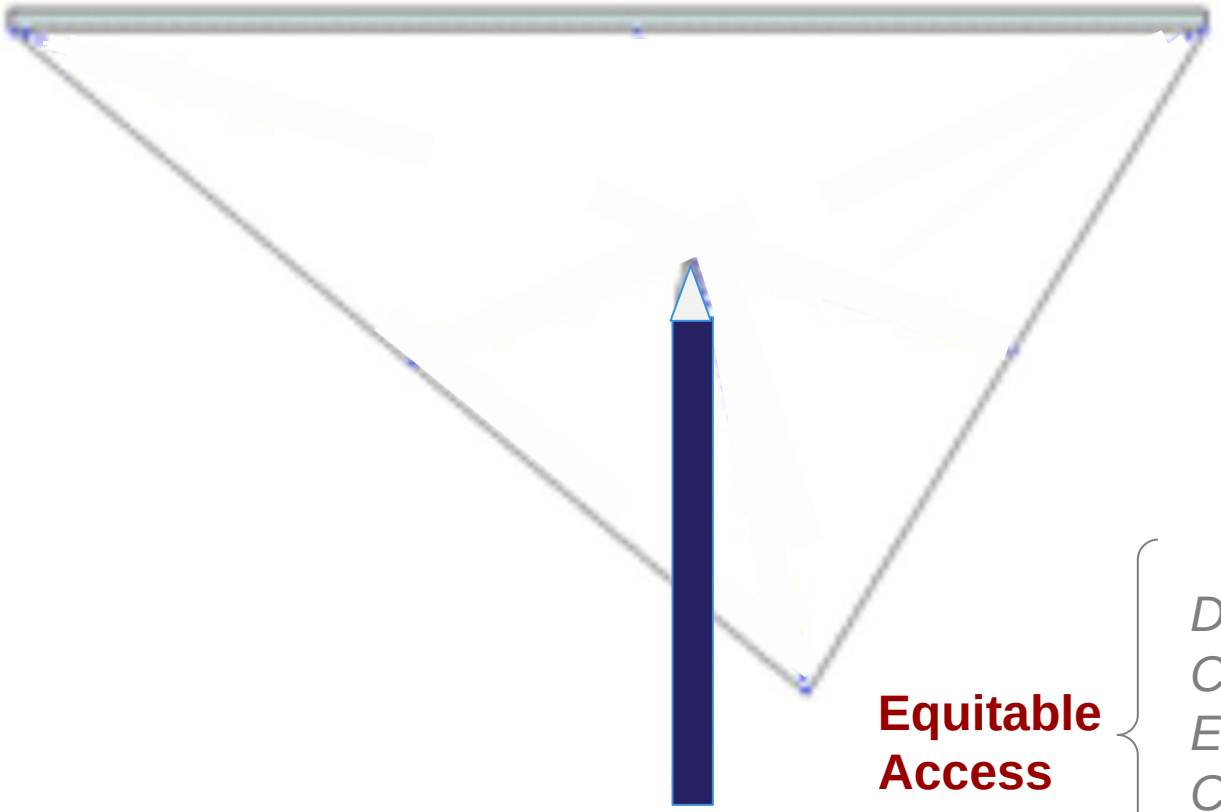
Sweeney JLME 1997

MA Governor William Weld

Balancing Competing Priorities

Scientific Utility

Security & Privacy



Equitable Access

Diverse Research Institutions
Citizen Scientists
Educators
Commercial Organizations
International Scientists

Mitigating Bias in Digital Twins

Data quality improvement: Cleaning and standardizing EHR data to address missing values and coding inconsistencies

Diverse data collection: Ensuring that the data used to train algorithms is representative of the wider patient population.

Sensitivity analysis: Evaluating the robustness of the algorithm by testing different feature sets and parameter configurations to assure convergence

Fairness metrics: Incorporating metrics to assess the fairness of the algorithm across different demographic groups.

Regulation and Data Governance: Ensuring that those who curate and control data sets act in the wider public interest.

Summary

Machine learning algorithms can be effectively and efficiently applied to a large populations to accurately phenotype patients to produce Digital Twins

Validation is critical to ensure the accuracy and precision of the unsupervised algorithms.

You need to think through all the steps and acknowledge the limits of your algorithms, especially when matching performance to workload

References

- Multi-modal Data Hosting

  Castro VM, Gainer V, Wattanasin N, Benoit B, Cagan A, Ghosh B, Goryachev S, Metta R, Park H, Wang D, Mendis M, Rees M, Herrick C, Murphy SN. The Mass General Brigham Biobank Portal: an i2b2-based data repository linking disparate and high-dimensional patient data to support multimodal analytics. J Am Med Inform Assoc. 2022 Mar 15;29(4):643-651. doi: 10.1093/jamia/ocab264. PMID: 34849976; PMCID: PMC8922162.

- Computational Phenotypes

  Yu S, Ma Y, Gronsbell J, Cai T, Ananthakrishnan AN, Gainer VS, Churchill SE, Szolovits P, Murphy SN, Kohane IS, Liao KP, Cai T. Enabling phenotypic big data with PheNorm. J Am Med Inform Assoc. 2018 Jan 1;25(1):54-60. doi: 10.1093/jamia/ocx111. PMID: 29126253; PMCID: PMC6251688.

- LLMs

  Alsentzer E, Rasmussen MJ, Fontoura R, Cull AL, Beaulieu-Jones B, Gray KJ, Bates DW, Kovacheva VP. Zero-shot interpretable phenotyping of postpartum hemorrhage using large language models. NPJ Digit Med. 2023 Nov 30;6(1):212. doi: 10.1038/s41746-023-00957-x. PMID: 38036723; PMCID: PMC10689487..