Plenary #1: Federated Learning Meets Negative Control Calibration: Toward Reliable Multi-Site Evidence Generation

Yong Chen (University of Pennsylvania),
Chongliang Luo (Washington University in Saint Louis),
Fleur Vereijken (Erasmus MC),
Fan Bu (University of Michigan),
Yumou Qiu (Peking University),
Bingyu Zhang (University of Pennsylvania)

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Overview



Yong Chen

Modern methods for federated learning



Chongliang Luo

PDA/MOSAiC in action



Fleur Vereijken

Model stability



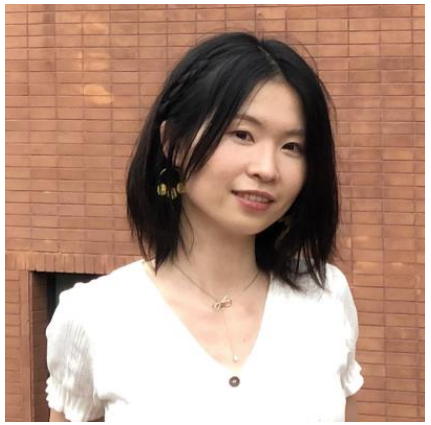
Fan Bu

Joint evidence synthesis and bias calibration



Yumou Qiu

Systematic Error Reference Outcomes



Bingyu Zhang

AI-enabled study implementation





MOSAiC: Multi-site One-Shot Aggregation of Compressed Risk Functions

Yong Chen (University of Pennsylvania)



~500 million patients from 80 countries

Distributed Research Networks (DRN)



Promise and challenges of multi-site studies using RWD



- Covers broader population
  - Results are more generalizable
  - Better statistical power
  - Opportunities of studying rare diseases

- Ecosystem of bringing together expertise from different investigators

- Privacy: sharing patient level data is challenging



(distributed algorithms are needed)

- Efficiency: need iterative coordination/analyses among investigators from different institutes

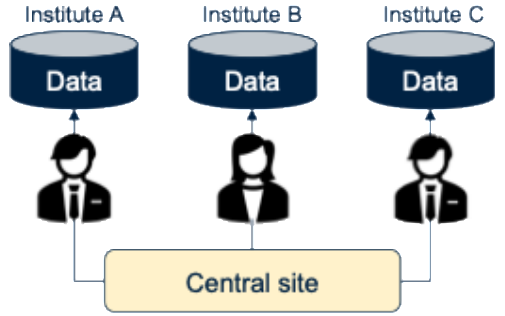
(design communication-efficient protocol is critical)

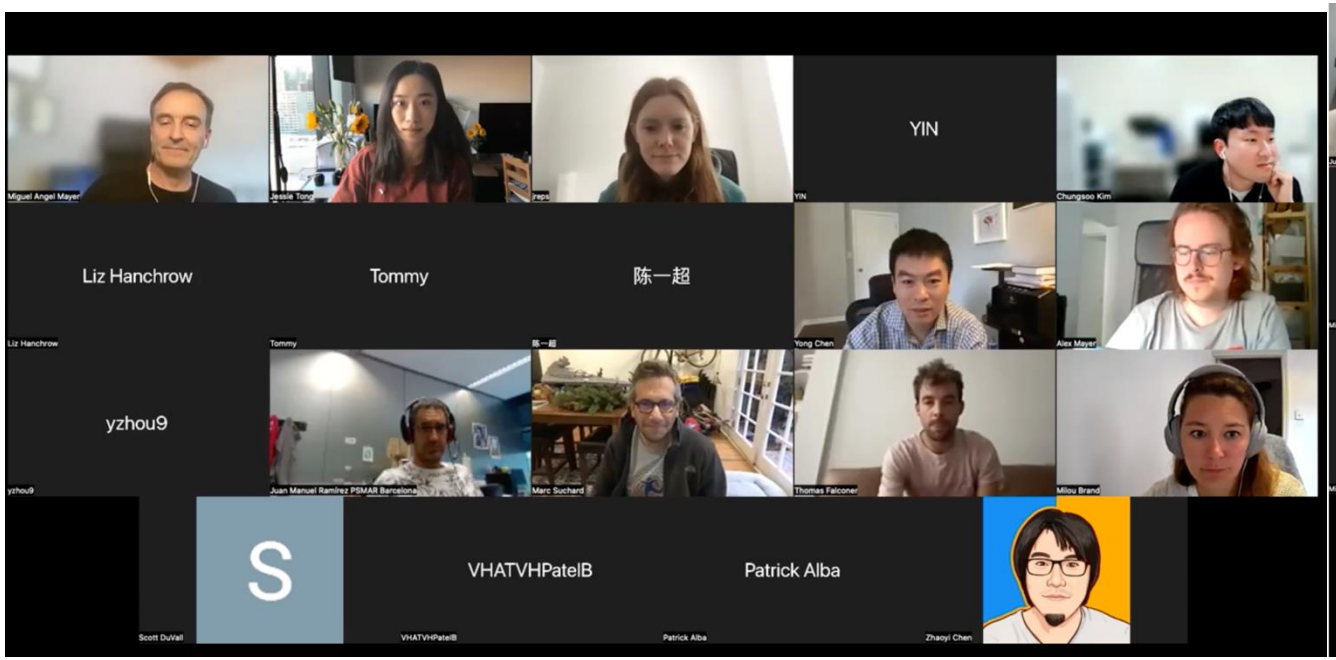
A federated learning example of RWD in action

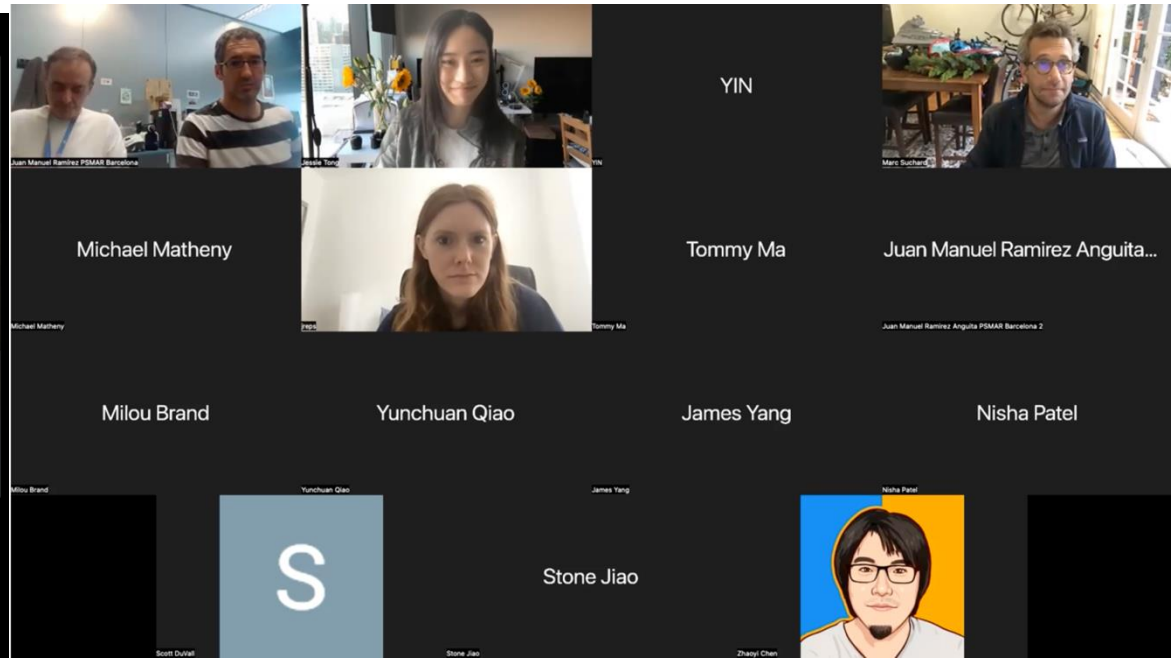


- 12 sites across 3 countries

A federated learning example of RWD in action

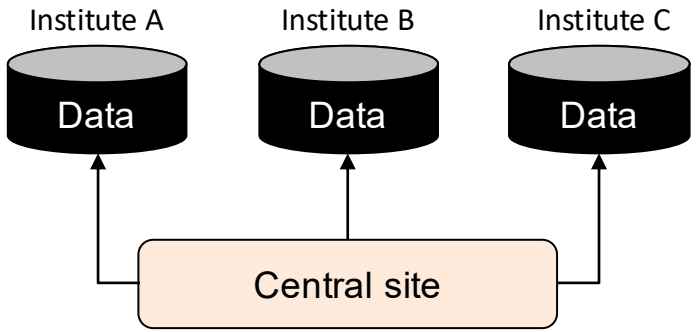






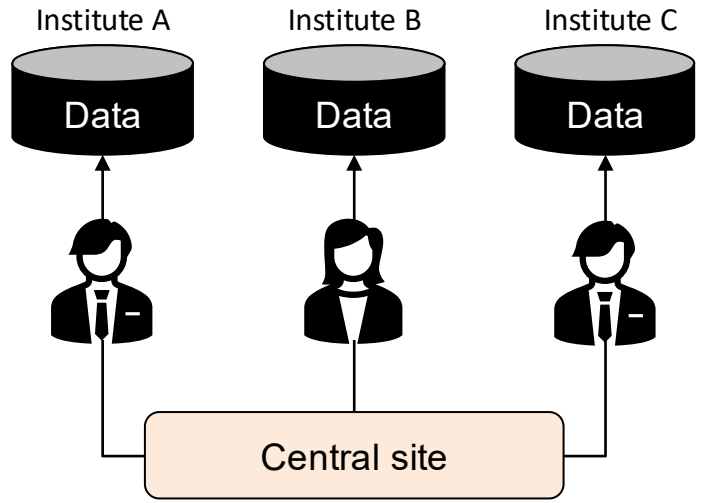
Static system





Hybrid system





Statistics (by Oct 2026)

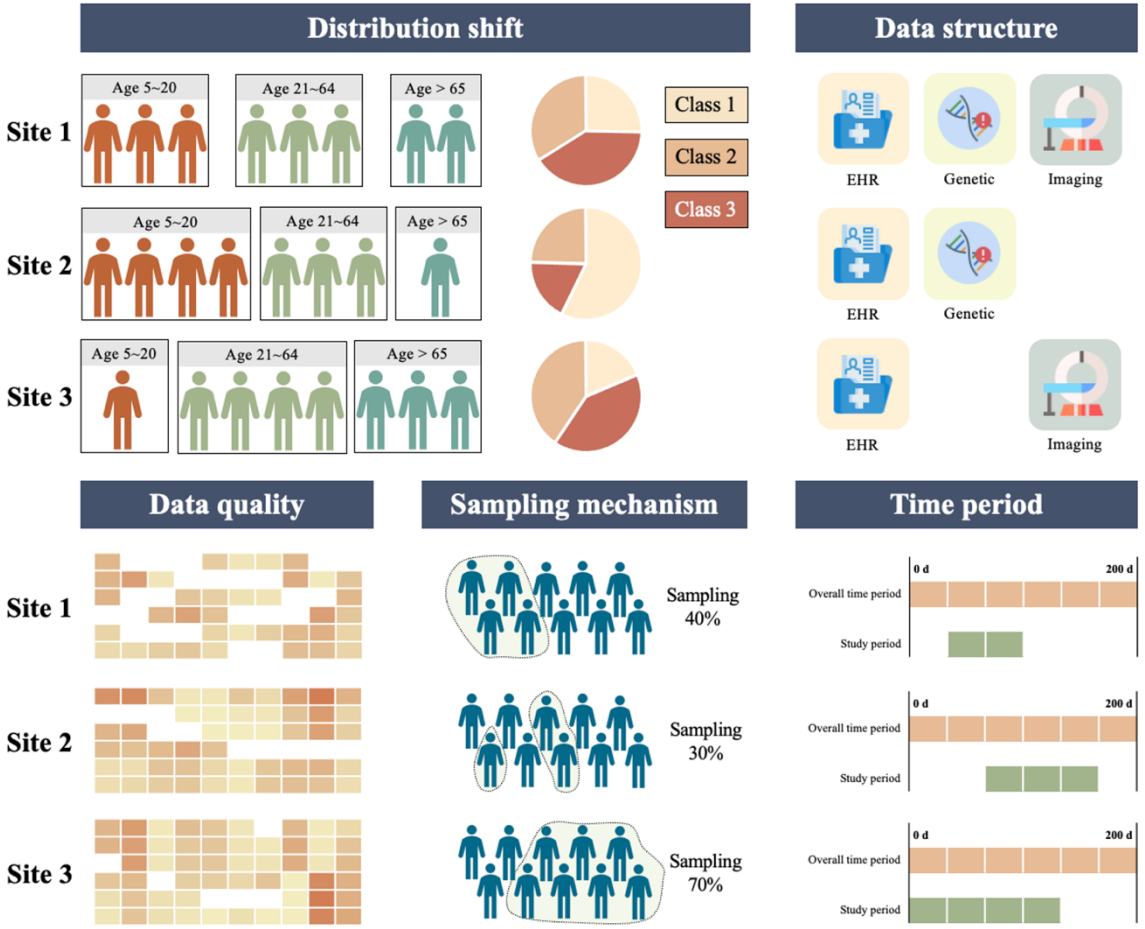
- **30** national and international studies
- **75** method papers and clinical papers
- **15** ongoing studies
- **> 50** data partners

PennCIL
A Computing · Inference · Learning lab at University of Pennsylvania

PDA
Privacy-preserving Distributed Algorithms

Developed by PennCIL lab
From 2015 to present

Heterogeneity? Yes, but what type?



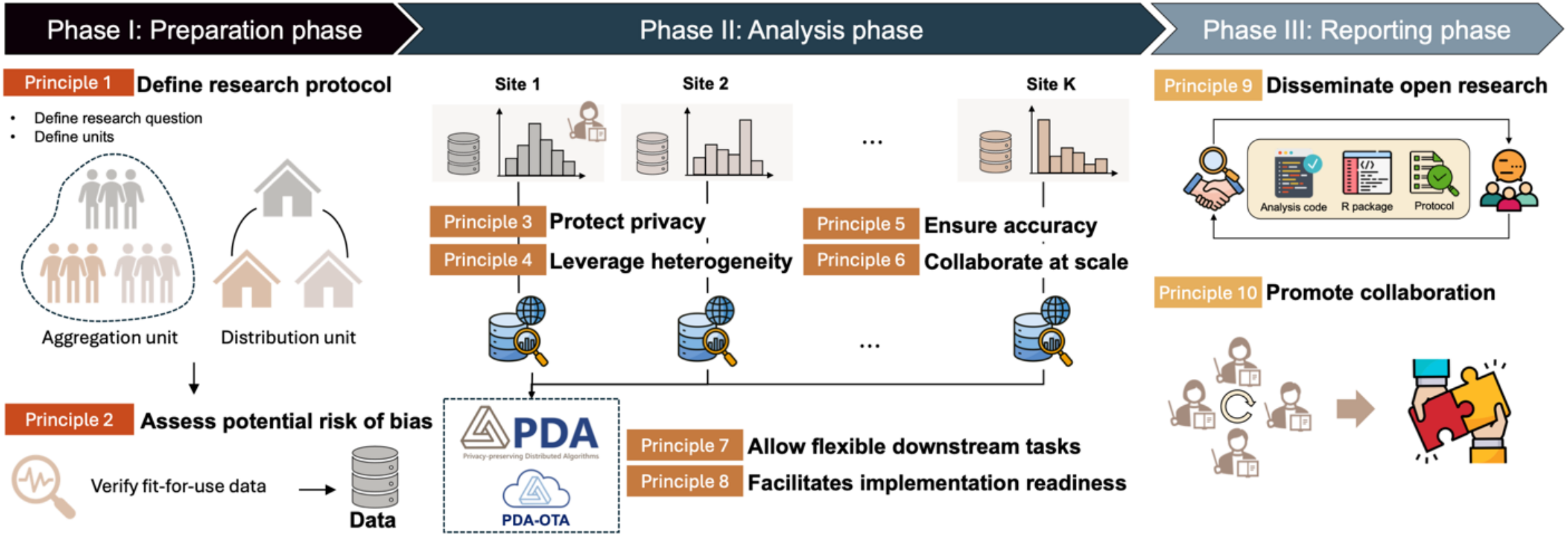
- Heterogeneity in **relationship between outcome and covariates, i.e., *f(y|x)* *"domain shift"***
- Heterogeneity in **distributions in covariates, *f(x)*. *"distribution-shift"***
- Heterogeneity in **data structure across sites. "missing modality"**
- Heterogeneity in **data quality**
- Heterogeneity in sampling mechanism/**design-aware** PDA
- Heterogeneity in **time**
- Other…

Ten principles for PDA

PDA (Privacy-Preserving Distributed Algorithms) in action: ten principles for high-quality multi-site clinical evidence generation

Yong Chen, PhD ✉, Jiayi Tong, PhD ✉, Yiwen Lu, PhD, Rui Duan, PhD, Chongliang Luo, PhD, Marc A Suchard, MD, PhD, Patrick B Ryan, PhD, Andrew E Williams, PhD, John H Holmes, PhD, Jason H Moore, PhD, Hua Xu, PhD, Yun Lu, PhD, Raymond J Carroll, PhD, Scott L Zeger, PhD, George Hripcsak, MD, MS, Martijn J Schuemie, PhD
Author Notes

Journal of the American Medical Informatics Association, Volume 33, Issue 9, September 2026, Pages 1636–1648, https://doi.org/10.1093/jamia/ocag119
Published: 13 July 2026 **Article history** ▾

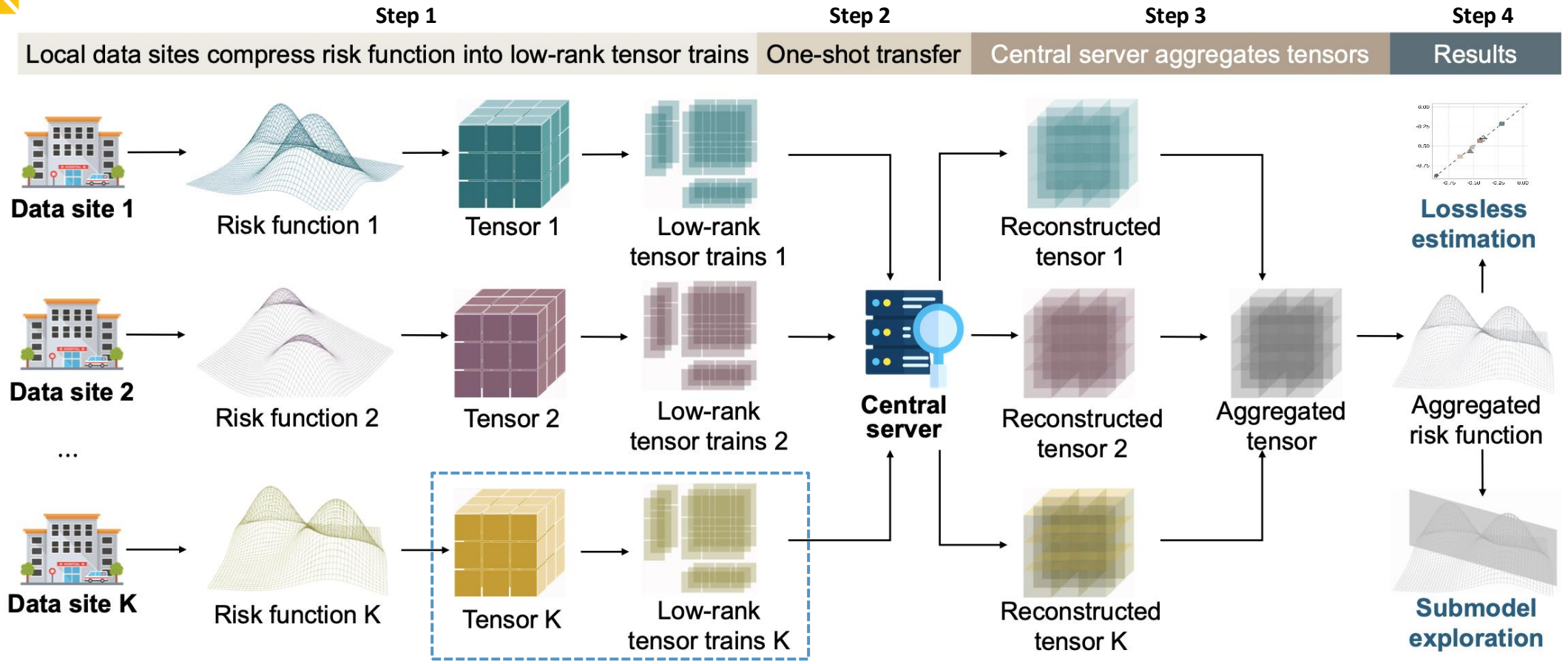


Chen et al. (2026) JAMIA. PDA in action

What would be a dream FL algorithm for OHDSI?

- **One-shot**: It takes a single round of summary-level data communication.
- **Lossless**: The resulting estimator can be controlled to be arbitrarily close to the pooled estimator.
- **Analytic submodel exploitability**: It supports one-shot submodel analysis.
- **Inclusiveness**: It does not require local estimation in any local site, so allows edge cases such as complete separation and zero events.

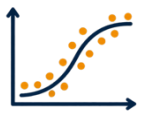
Tensor Train-powered MOSAiC workflow [8]



TT-cross algorithm:
highly efficient, no need to calculate all tensor entries

Our solutions

Models



Generalized linear models

Logistic and other GLMs
Site-specific nuisance parameters
All covariate types



Stratified Cox models

Time-to-event outcomes
Site-specific baseline hazards



Target Trial Emulation

Target trial emulation with
binary and time-to-event outcomes

Applications



Alzheimer's drug repurposing

LATTE and FLAT-WHITE
Federated high-throughput causal machine learning



Genomics

Federated learning using Biobank data
Comprehensive Genomic Analysis Toolbox



Rare disease

Predictive Analytics via Networked Distributed Algorithms for multi-system diseases (PANDA)

Predictive Analytics via Networked Distributed Algorithms for multi-system diseases (PANDA)







- A multi-institutional initiative (Penn + UF) leveraging AI to accelerate diagnosis of rare, multi-system diseases through distributed learning on EHR data.



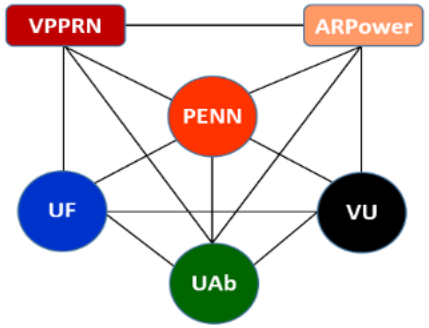


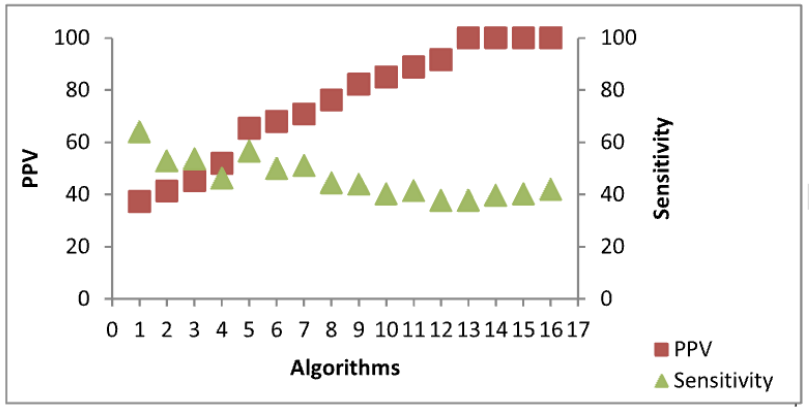
PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2016; **25**: 1368–1374
Published online 1 November 2016 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.4116

ORIGINAL REPORT

Development and validation of case-finding algorithms for the identification of patients with anti-neutrophil cytoplasmic antibody-associated vasculitis in large healthcare administrative databases

Antoine G. Sreih[1]*, Narender Annapureddy[2], Jason Springer[3], George Casey[4], Kevin Byram[2], Andy Cruz[5], Maya Estephan[3], Vince Frangiosa[5], Michael D. George[1], Mei Liu[6], Adam Parker[3], Sapna Sangani[1], Rebecca Sharim[1] and Peter A. Merkel[1,7] for the Vasculitis Patient-Powered Research Network





1. ICD9
2. ICD9+Medications
3. ICD9+Encounter
4. ICD9+Encounter+Medications
5. ICD9+Specialty
6. ICD9+Encounter+Specialty
7. ICD9+Specialty+Medications
8. **ICD9+Encounter+Specialty+Medications**
9. ICD9+pANCA
10. ICD9+Medications+pANCA
11. ICD9+Encounter+pANCA
12. ICD9+Encounter+Medications+pANCA
13. ICD9+Encounter+Specialty+Medications+pANCA
14. ICD9+Encounter+Specialty+pANCA
15. ICD9+Specialty+Medications+pANCA
16. ICD9+Specialty+pANCA

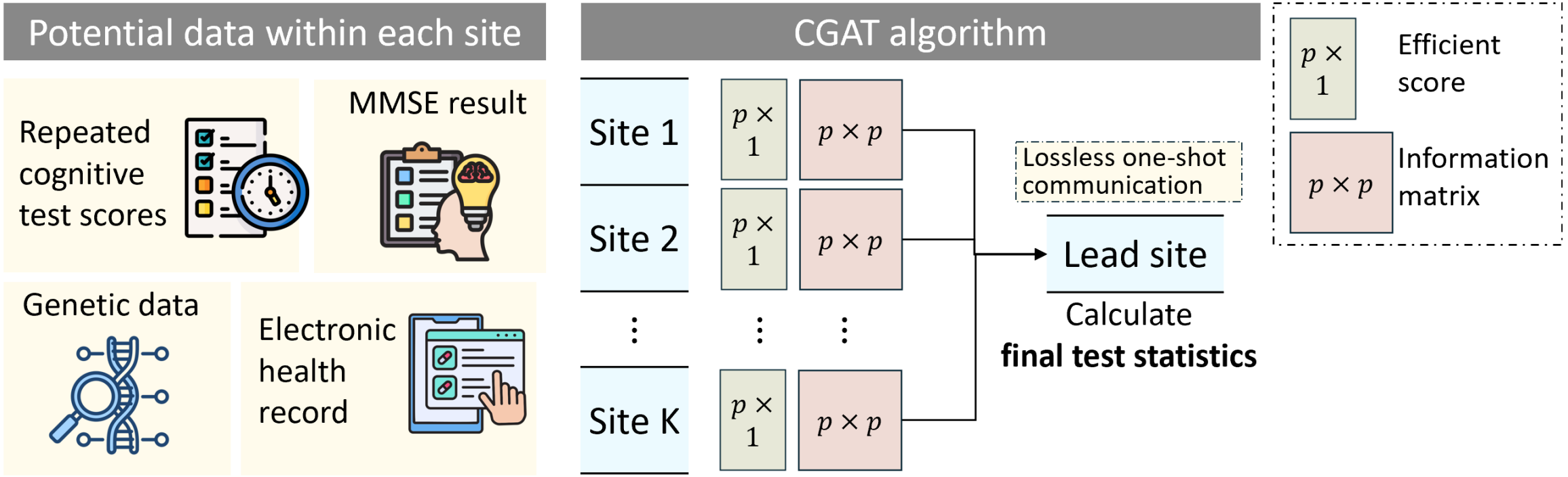
CGAT
(Comprehensive Genomic Analysis Toolbox)


Yong Chen


Adam Naj


Jason Moore

Local clinical and genomic data support joint tests across participating studies







PDA/MOSAiC in Action

Chongliang Luo (Washington University in Saint Louis)



Continuous outcome:
COVID length of stay



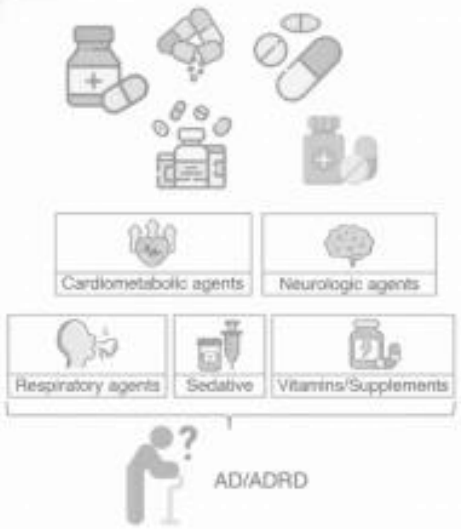
Time-to-event outcome:
cancer dx and proteomics

Binary outcome:
ADRD disease onset
and drug repurposing

1

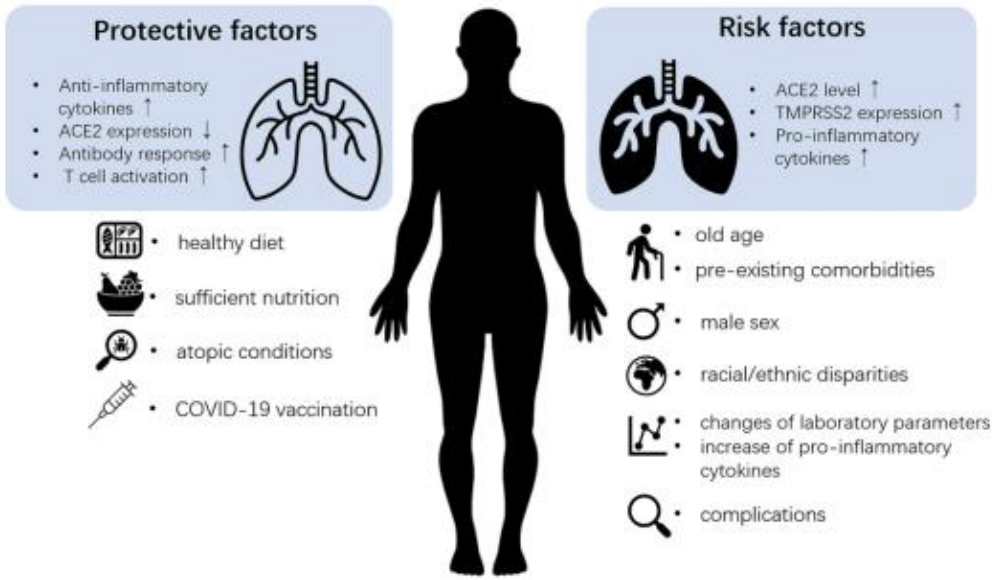
2

3



COVID length of stay

- Length of stay (days): continuous outcome
- Linear regression model: LoS ~ risk factors (age, gender, admission time, comorbidities)

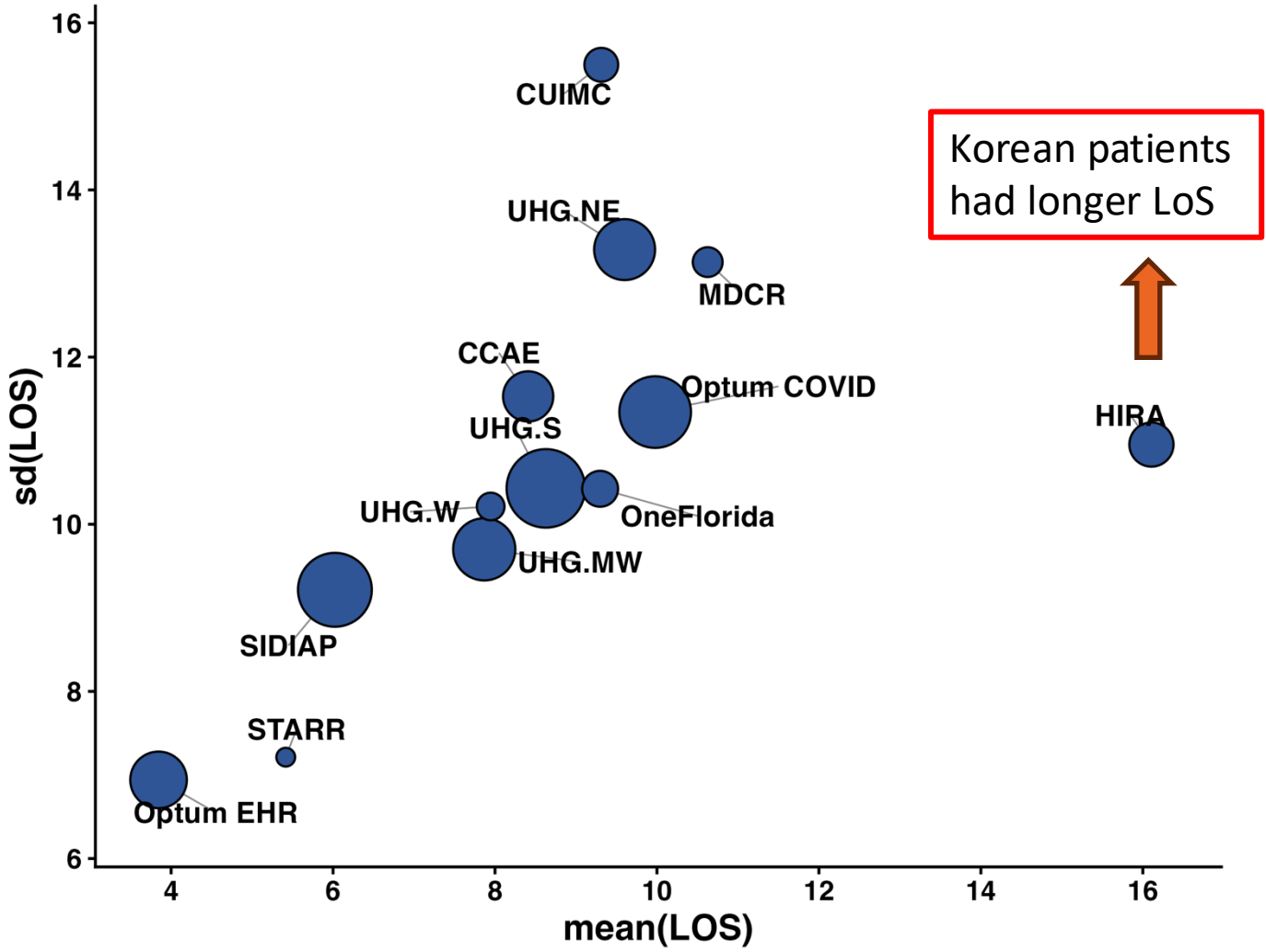




120,609 COVID-19 patients from 11 collaborative data sources worldwide

COVID length of stay (DLMM)

- Heterogeneity across sites.
- Linear mixed-effects model (LMM).
- Allows effects of risk factors to vary across sites.
- **DLMM:** distributed LMM.



COVID length of stay (DLMM)

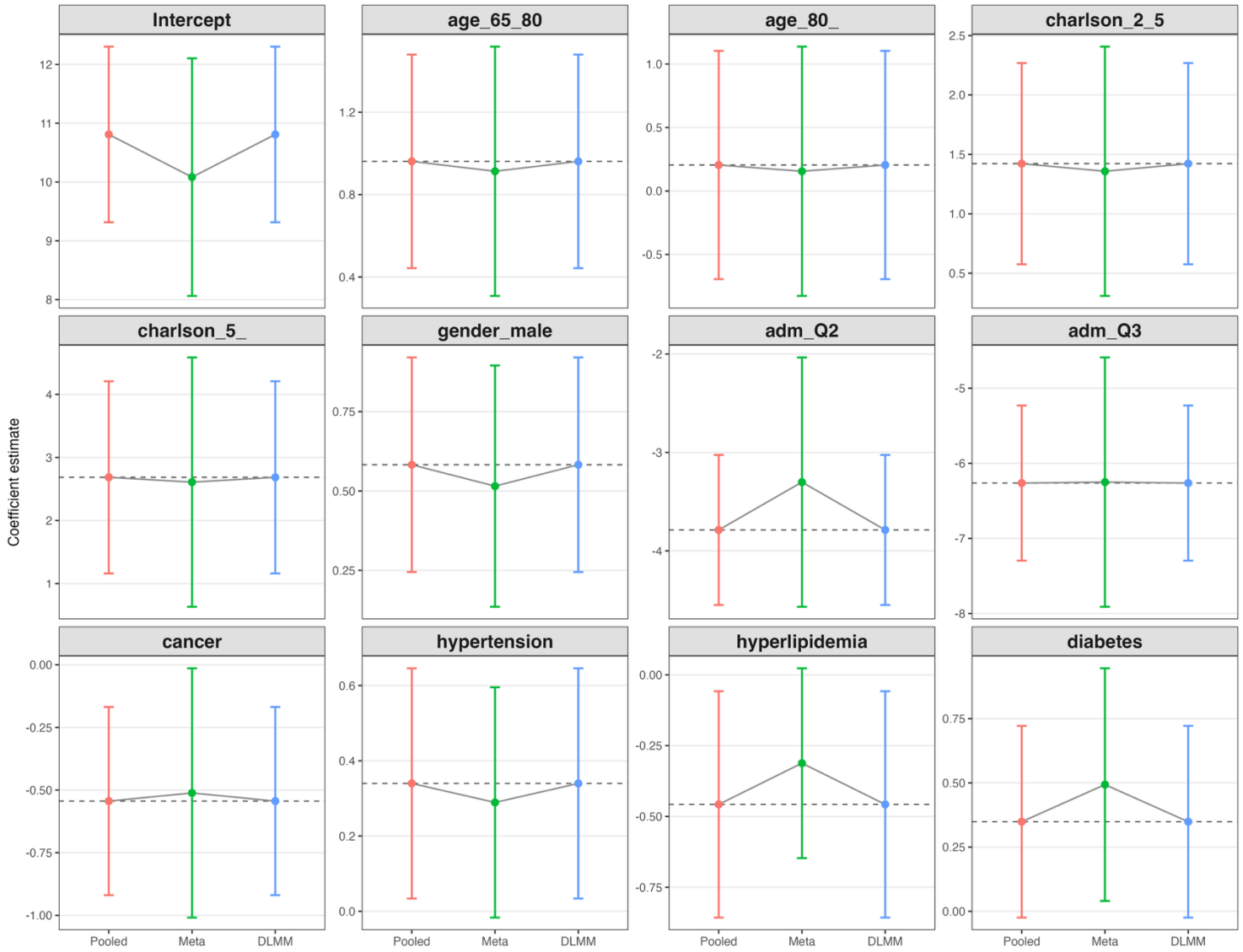
AD = aggregate data

R code:

```
library(pda)
pda(site_id = 'CCAE',
      ipdata = covid_CCAE,
      control = control)
```

AD for a site with 2000 patients
and 16 risk factors and LoS outcome:

	Intercept	age_65_80	age_80_	...	obesity	heart_disease	diabetes	LOS
Intercept	2000	958	531	...	595	634	598	15983
age_65_80	958	958	0	...	278	295	292	7579
age_80_	531	0	531	...	166	176	156	4246
...	...	...	...	...	...	...	...	...
obesity	595	278	166	...	595	195	181	4714
heart_disease	634	295	176	...	195	634	191	5106
diabetes	598	292	156	...	181	191	598	4818
LOS	15983	7579	4246	...	4714	5106	4818	141125



COVID length of stay (DLMM)

nature > nature communications > articles > article

Article | Open access | Published: 30 March 2022

DLMM as a lossless one-shot algorithm for collaborative multi-site distributed linear mixed models

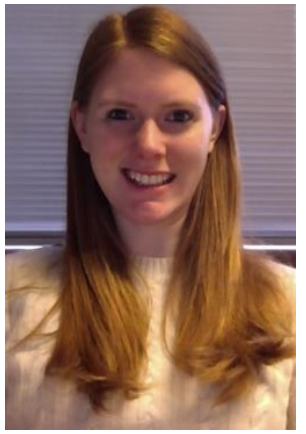
Chongliang Luo, Md. Nazmul Islam, Natalie E. Sheils, John Buresh, Jenna Reps, Martijn J. Schuemie, Patrick B. Ryan, Mackenzie Edmondson, Rui Duan, Jiayi Tong, Arielle Marks-Anglin, Jiang Bian, Zhaoyi Chen, Talita Duarte-Salles, Sergio Fernández-Bertolín, Thomas Falconer, Chungsoo Kim, Rae Woong Park, Stephen R. Pfohl, Nigam H. Shah, Andrew E. Williams, Hua Xu, Yujia Zhou, Ebbing Lautenbach, ... Yong Chen + Show authors

Nature Communications **13**, Article number: 1678 (2022) | Cite this article

https://www.nature.com/articles/s41467-022-29160-4


Chongliang Luo


Nazmul Islam


Jenna Reps


Yong Chen

OHDSI

Penn

COLUMBIA | COLUMBIA UNIVERSITY IRVING MEDICAL CENTER

Tufts | CTSI Tufts Clinical and Translational Science Institute

UnitedHealth Group

Stanford MEDICINE

janssen PHARMACEUTICAL COMPANIES OF Johnson & Johnson

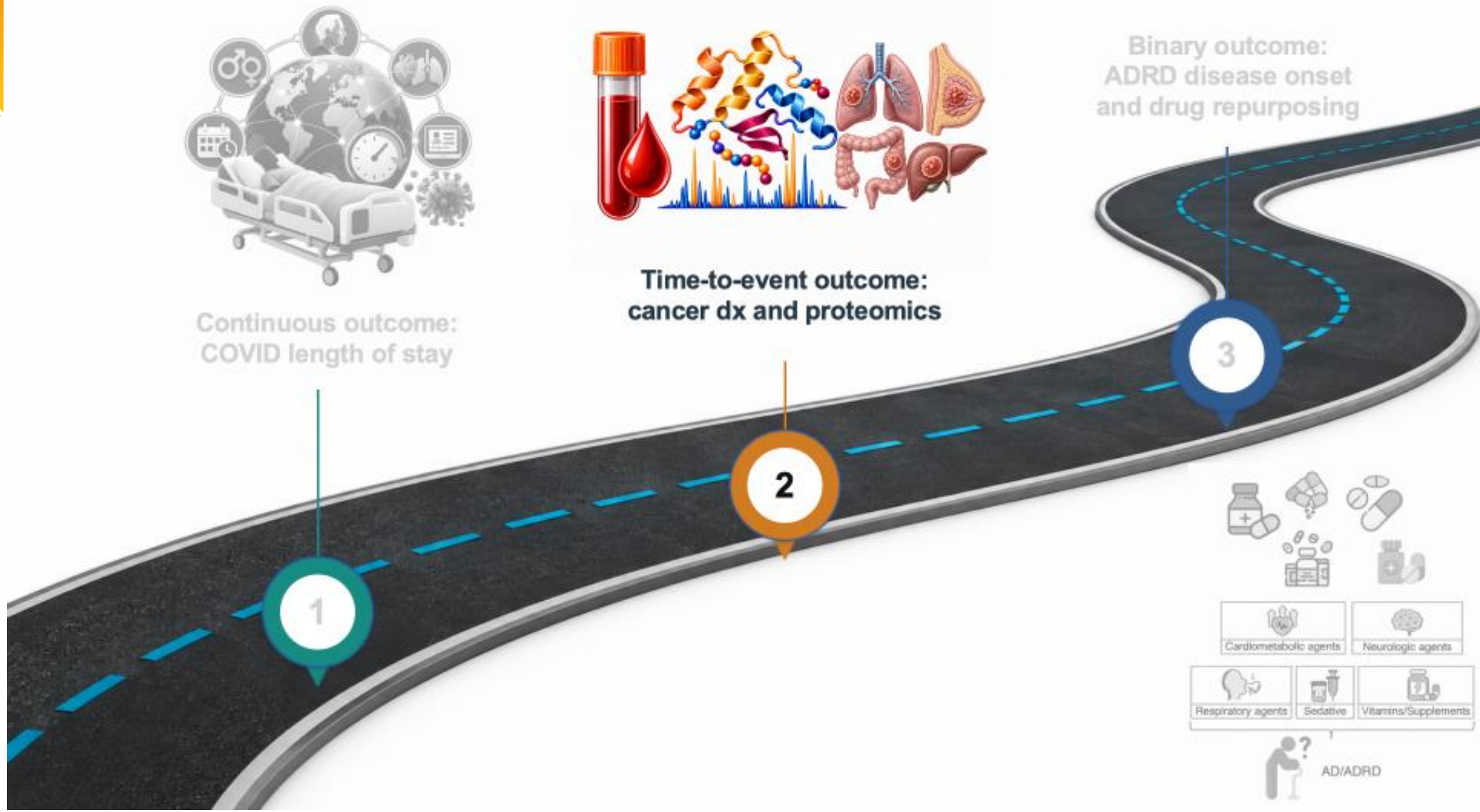
OneFlorida Clinical Research Consortium

IBM Watson Health

UTHealth | The University of Texas Health Science Center at Houston

SIDIAP

HIRA


Continuous outcome:
COVID length of stay
Time-to-event outcome:
cancer dx and proteomics
Binary outcome:
ADRD disease onset
and drug repurposing
1
2
3
Cardiometabolic agents
Neurologic agents
Respiratory agents
Sedative
Vitamins/Supplements
AD/ADRD

Federated learning for Cox regression

ODAC: federated Cox regression

Journal of the American Medical Informatics Association, 0(0), 2020, 1–9
doi: 10.1093/jamia/ocaa044
Research and Applications

Research and Applications

Learning from local to global: An efficient distributed algorithm for modeling time-to-event data

Rui Duan,[1,†] Chongliang Luo,[1,†] Martijn J. Schuemie,[2,†] Jiayi Tong,[1] C. Jason Liang,[3] Howard H. Chang,[4] Mary Regina Boland,[1] Jiang Bian,[5,6] Hua Xu,[7] John H. Holmes,[1] Christopher B. Forrest,[8] Sally C. Morton,[9] Jesse A. Berlin,[10] Jason H. Moore,[1] Kevin B. Mahoney,[11] and Yong Chen[1]

ODACH: federated Cox regression with site-specific baseline hazards

scientific reports

Explore content | About the journal | Publish with us

nature > scientific reports > articles > article

Article | Open access | Published: 22 April 2022

ODACH: a one-shot distributed algorithm for Cox mod with heterogeneous multi-center data

Chongliang Luo, Rui Duan, Adam C. Naj, Henry R. Kranzler, Jiang Bian & Yong Chen

ODACT: federated Cox regression with time-varying effects

Journal of the American Medical Informatics Association, 2025, **32(4)**, 656–664
https://doi.org/10.1093/jamia/ocae313
Advance access publication 26 January 2025
Research and Applications

Research and Applications

Communication-efficient federated learning of temporal effects on opioid use disorder with data from distributed research networks

C. Jason Liang, PhD[1], Chongliang Luo, PhD[2,3], Henry R. Kranzler, MD[4], Jiang Bian, PhD[5], Yong Chen, PhD*[,3,6]



Rui Duan



Chongliang Luo



Martijn Schuemie



C. Jason Liang

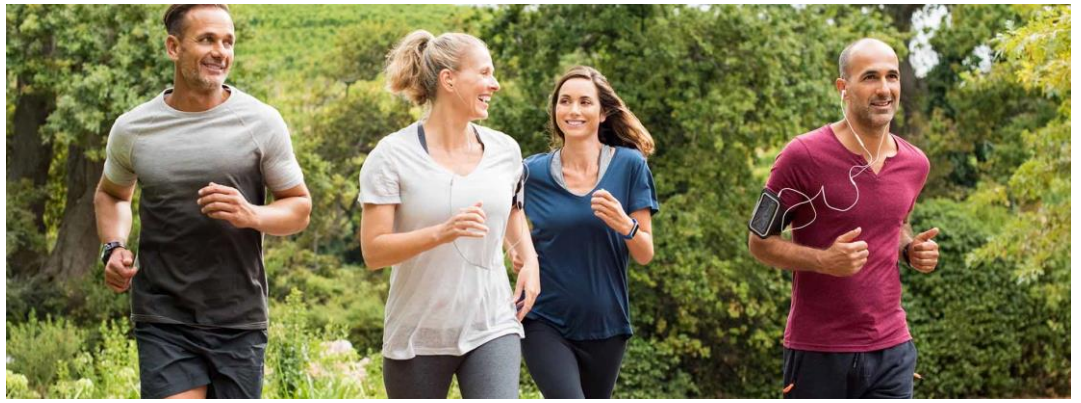


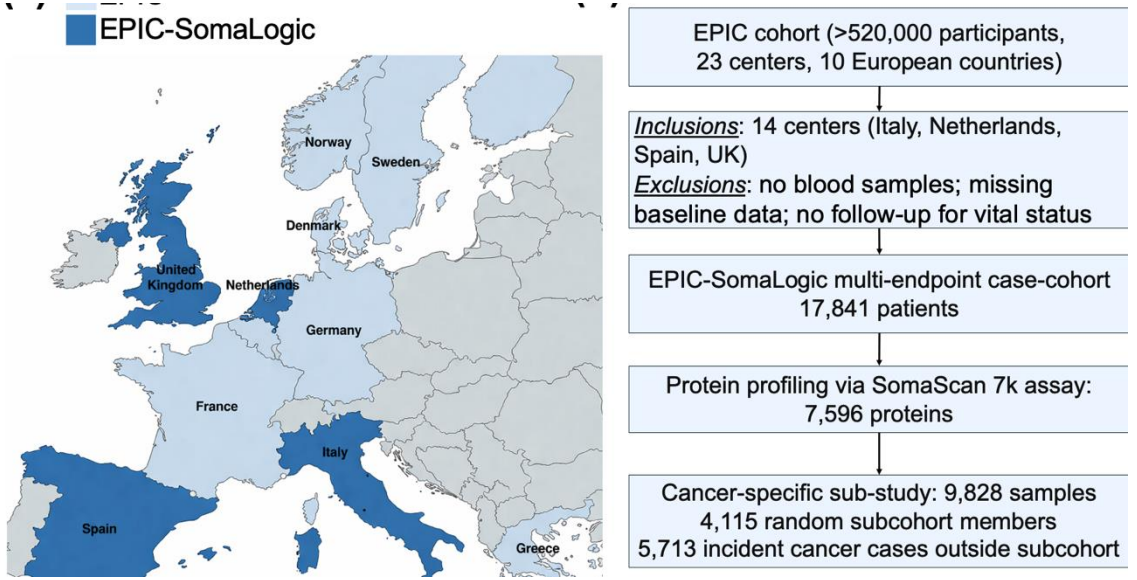
Yong Chen

Pan-cancer proteomics biomarkers

- Time to cancer Dx: time-to-event outcome.

- Cox proportional hazards regression: time to Dx ~ covariates + proteins.

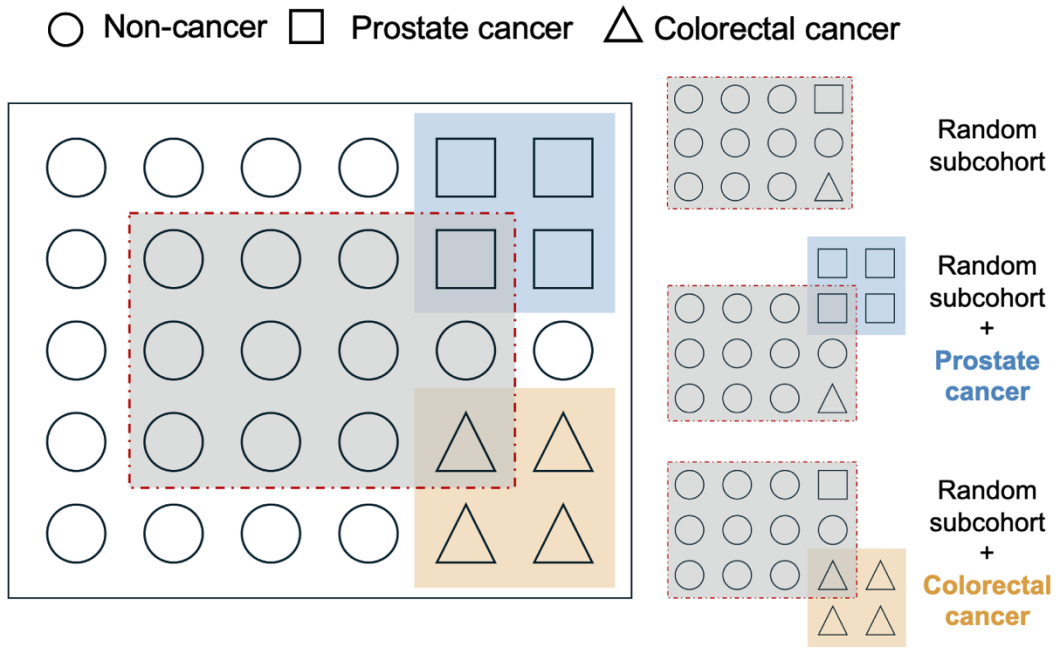




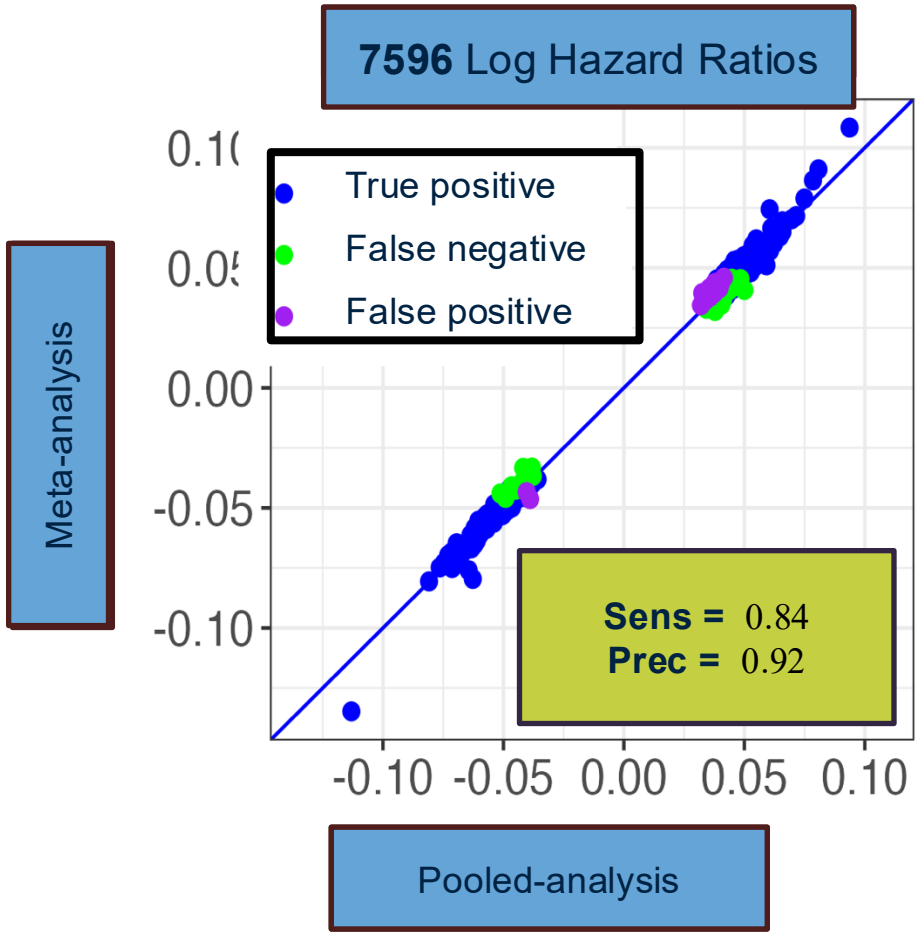
The European Prospective Investigation into Cancer and Nutrition (EPIC-Europe) study

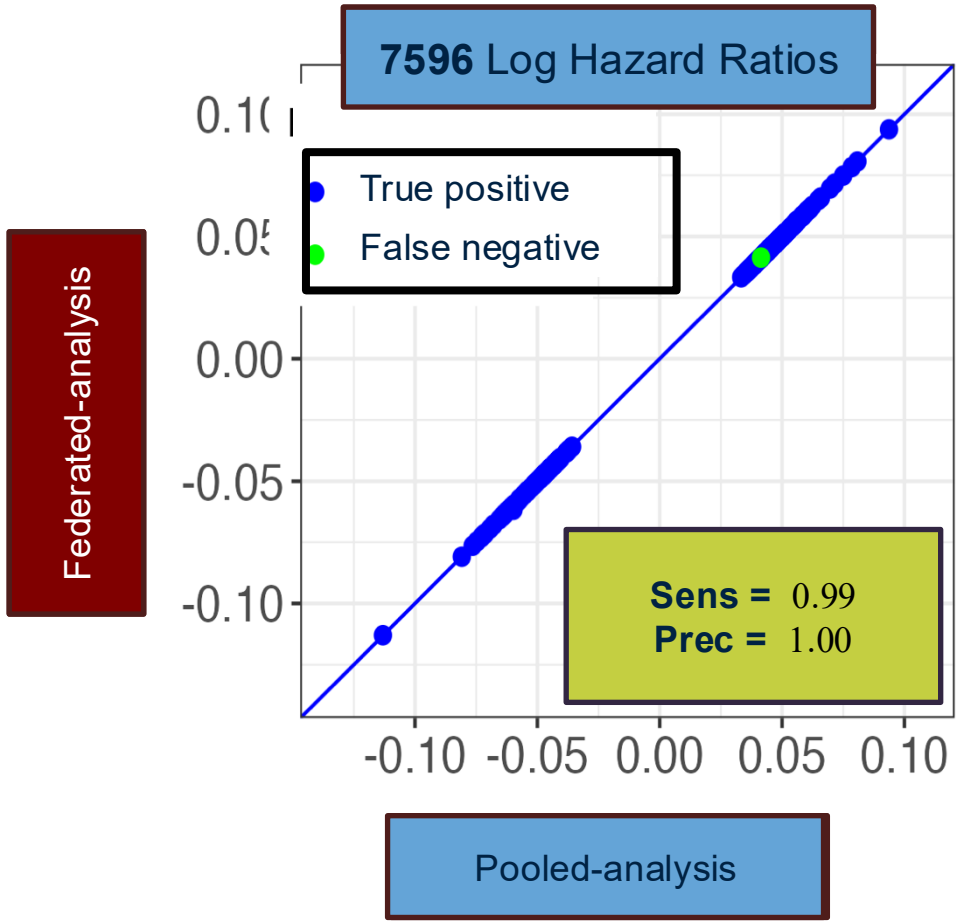
Pan-cancer proteomics biomarkers (DRAFT)

- Case-cohort design for pan-cancer analyses: a fixed subcohort + different cancer cases.
- Extend ODACH for case-cohort design.
- **DRAFT**: Design-informed Regression Algorithms for Federated-learning Toolbox



Pan-cancer proteomics biomarkers (DRAFT)





Meta-analysis: inverse-variance weighted estimates

Pan-cancer proteomics biomarkers (DRAFT)

- Wang, Y., et al. 2026. Design-aware Federated Learning for Large-scale Cost-effective Studies: Advancing Biomarker Discovery for Cancer Prevention. Nature Communications. Under review.

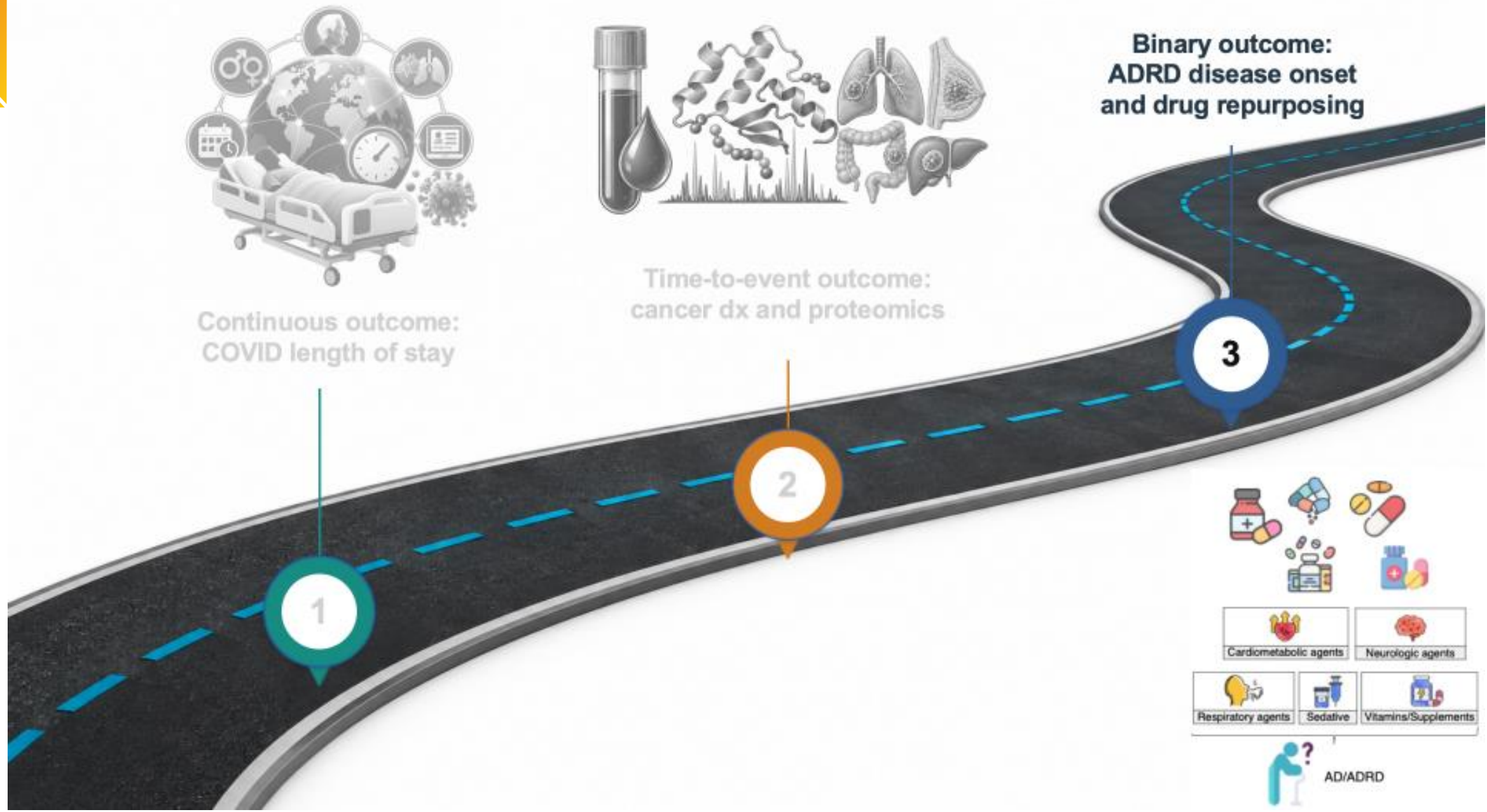

Yudong Wang


Chongliang Luo


Ali Farnudi


Vivian Viallon


Yong Chen


Continuous outcome:
COVID length of stay
Time-to-event outcome:
cancer dx and proteomics
Binary outcome:
ADRD disease onset
and drug repurposing
1
2
3
Cardiometabolic agents
Neurologic agents
Respiratory agents
Sedative
Vitamins/Supplements
AD/ADRD

Target Trial Emulation

- TTE: use observational data (e.g. EHR) to emulate RCT. (LEGEND framework)
- Drug repurposing disease Dx (ADRD): binary outcomes

Clinical questions

- ❑ Treatment effectiveness
- ❑ Medication safety
- ❑ Comparative effectiveness
- ❑ Heterogeneous treatment effect
- ❑ etc.

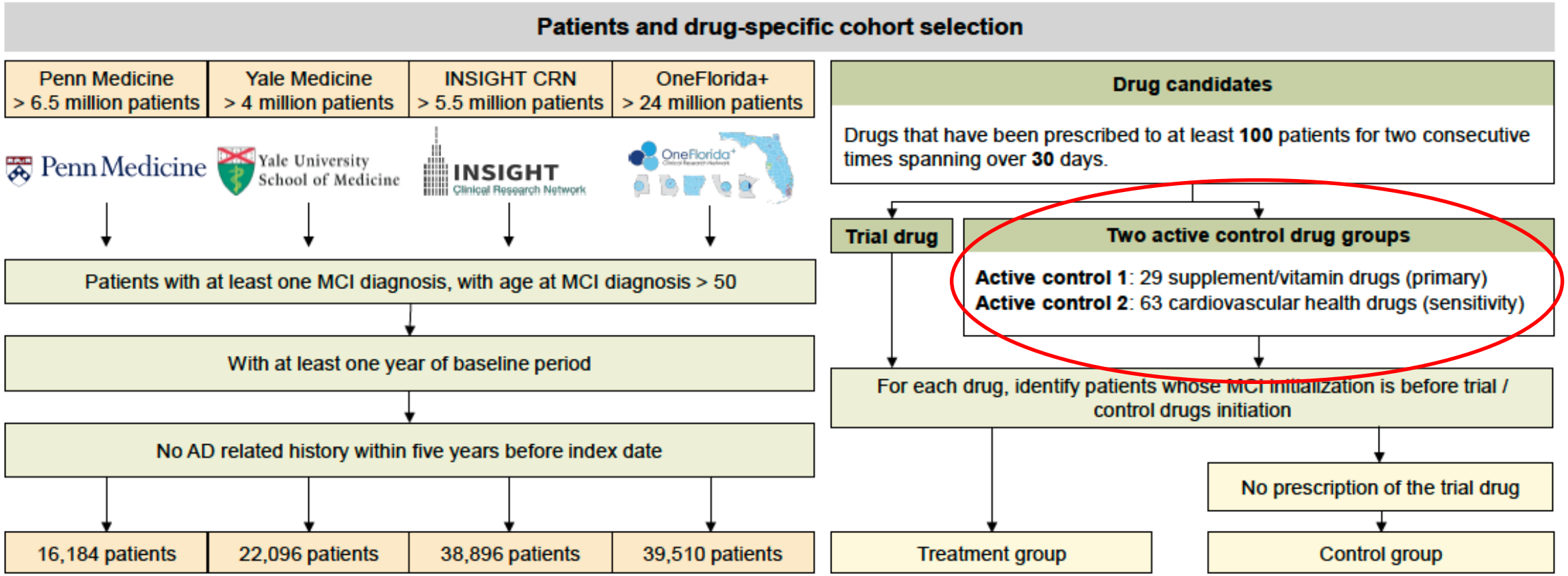
An ideal randomized controlled trial

- ❑ Eligibility criteria
- ❑ Treatment strategies
- ❑ Assignment procedure
- ❑ Follow-up period
- ❑ Outcome
- ❑ Causal contrast of interest
- ❑ Analysis plan

Emulation analysis

- ❑ Variable encodings
- ❑ Cohort query
- ❑ Emulate blind assignment
  - Weighting including IPTW, normalized and stabilized weights
  - Stratification and matching
- ❑ Analysis
  - Binary, count, time-to-event data
- ❑ Negative control experiments

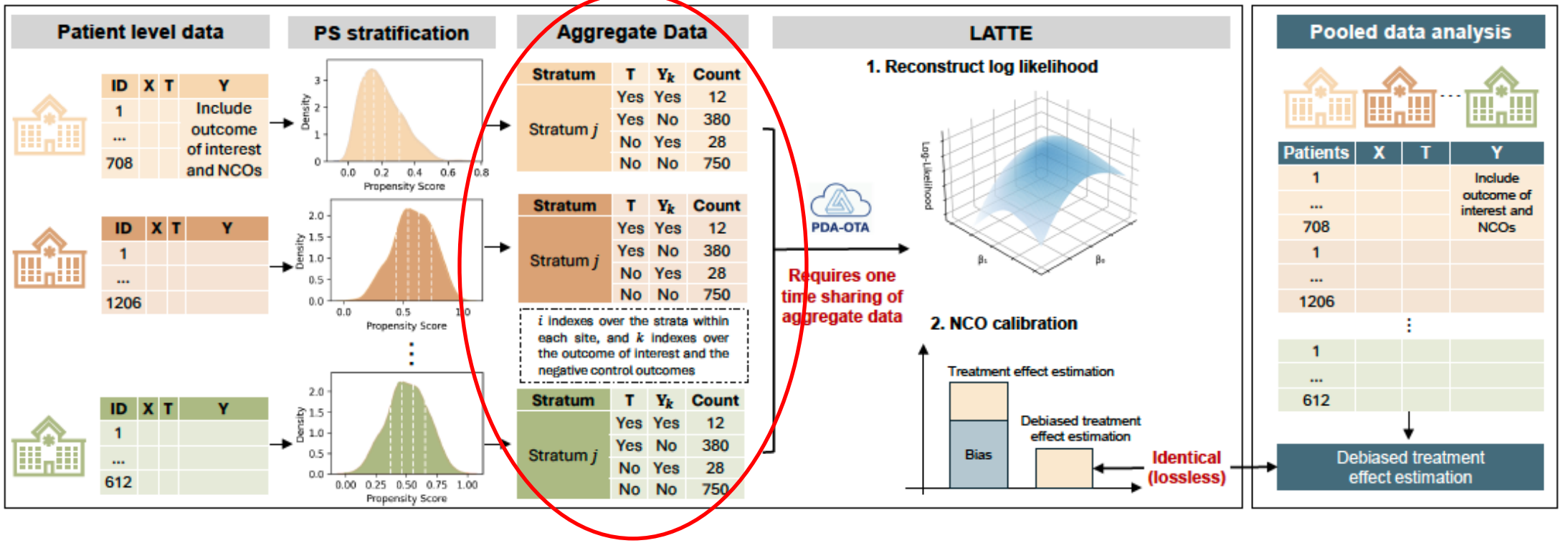
Federated drug repurposing (LATTE)



Target trial emulation for Alzheimer's disease and related dementia (ADRD) drug repurposing utilizing multi-center data.

LATTE: One-shot Lossless Algorithm for Federated Target Trial Emulation.

ADRD drug repurposing (LATTE)



propensity-score

Use AD (2x2 contingency tables) to **reconstruct** a global log-likelihood.

NCOs to calibrate for systematic biases.

LATTE is lossless.

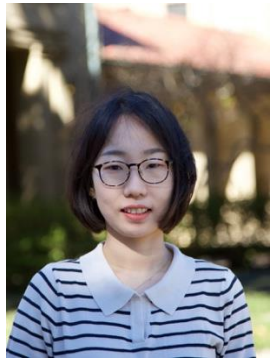
Lancet Digital Health

Alzheimer's disease drug repurposing using decentralised electronic health records: a one-shot lossless algorithm for federated target trial emulation (LATTE) and a multi-site observational study

Lu Li, Qiong Wu, Yiwen Lu, Kyra S O'Brien, Bingyu Zhang, Ting Zhou, Jiayi Tong, Dazheng Zhang, Yuqing Lei, Huilin Tang, Haoyang Li, Zhenxing Xu, Yu Huang, Yu Hu, Yujia Zhou, Fongci Lin, Na Hong, Ying Jiang, Fei Wang, Jiang Bian, Hua Xu, Yong Chen[1], Jeff D Williamson, David A Wolk, Yun Lu, David A Asch, Yong Chen[2]



Lu Li



Qiong Wu



Jiayi Tong



Yong Chen

eBioMedicine
Part of THE LANCET *Discovery Science*

This journal · Journals · Publish · Clinical · Global health · Multimedia · Events · About

ARTICLES · Volume 132, 106466, October 2026 · *Open Access* · Download Full Issue

Discovering repurposable drugs for Alzheimer's disease and related dementias: target trial emulation using decentralised real-world data

Qiong Wu [a,b,c] · Lu Li [b,d] · Yuqing Lei [b,e] · Ting Zhou [b,e] · Huilin Tang [b,e] · Bingyu Zhang [b,d] · Yiwen Lu [b,d] · Maxwell Salvatore [b,e,f,g] · Dazheng Zhang [b,d] · Jiayi Tong [h] · Tingyin Wang [b,d] · Siqi Chen [b,d] · Haoyang Li [i] · Zhenxing Xu [i] · Yu Huang [j] · Yu Hu [k] · Na Hong [l] · Yujia Zhou [l] · Fongci Lin [l] · Kyra S. O'Brien [m,n] · Yong Chen [o,p] · Ruowang Li [q] · Hua Xu [l] · Fei Wang [i] · Jiang Bian [j] · David A. Wolk [m] · Yong Chen [b,d,e,g,n,r] · Show less

NPJ Health System

FLAT-WHITE: From Target Trial Emulation to Federated High-Throughput Causal Machine Learning using Decentralized Real-World Data

Jiayi Tong[1], Lu Li[2,3], Dazheng Zhang[2,7], Kyra S.O'Brien[4,5], Yiwen Lu[2,3], Qiong Wu[6], Ting Zhou[2,7], Bingyu Zhang[2,3], Yuqing Lei[2,7], Huilin Tang[2,7], Naimin Jing[8], Haoyang Li[9], Zhenxing Xu[9], Yu Huang[10], Yu Hu[10], Yujia Zhou[11], Fongci Lin[11], Ying Jiang[12], Haitao Chu[13], Yong Chen[13], Fei Wang[9], Jiang Bian[10], Hua Xu[11], David A. Wolk[4], Yong Chen[2,3,7,14,15]

PDA

https://pdamethods.org/

Penn Medicine

Yale University School of Medicine

INSIGHT Clinical Research Network

OneFlorida+ Clinical Research Network



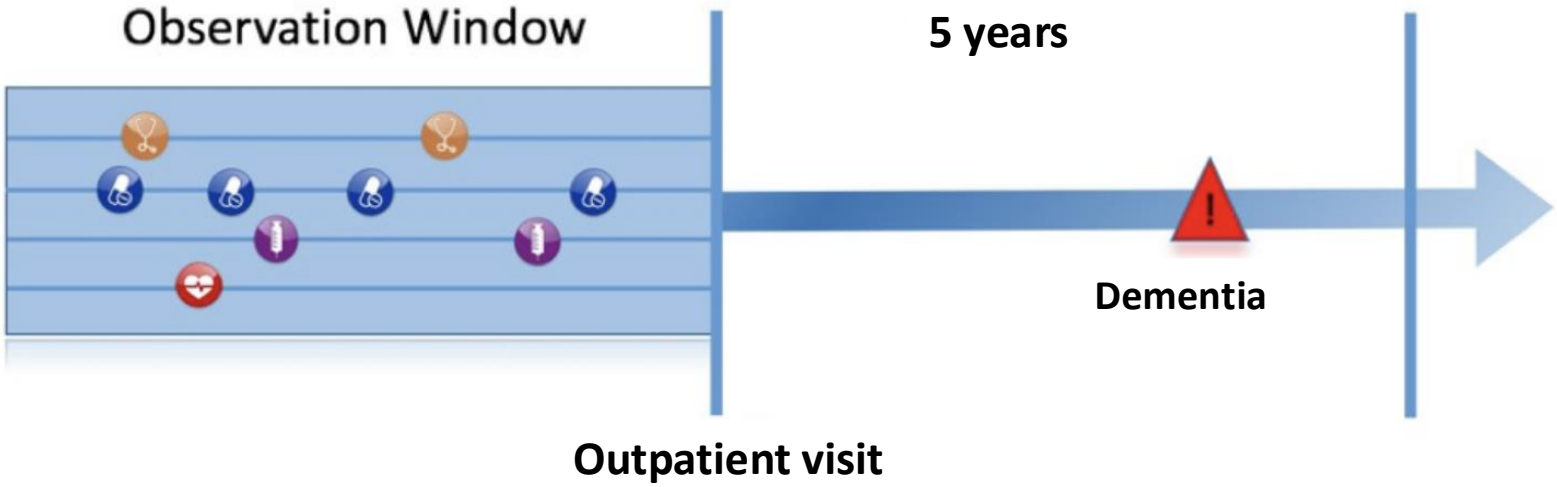
Model stability in federated learning

Fleur Vereijken (Erasmus MC)

What is a clinical Prediction model?

They answer questions like: "For patients with an outpatient visit in 2017,

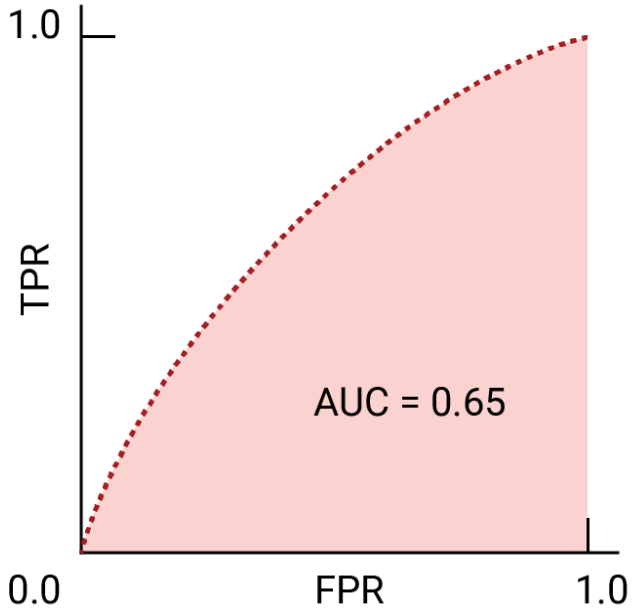
What is their risk of dementia within 5 years"



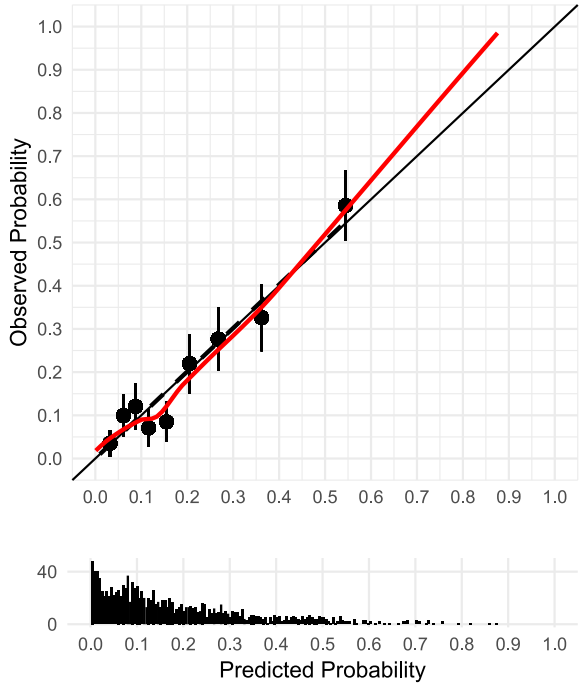
Prediction models can aid medical decision making and improve healthcare.

How to evaluate a prediction model?

AUROC



Calibration - Eavg



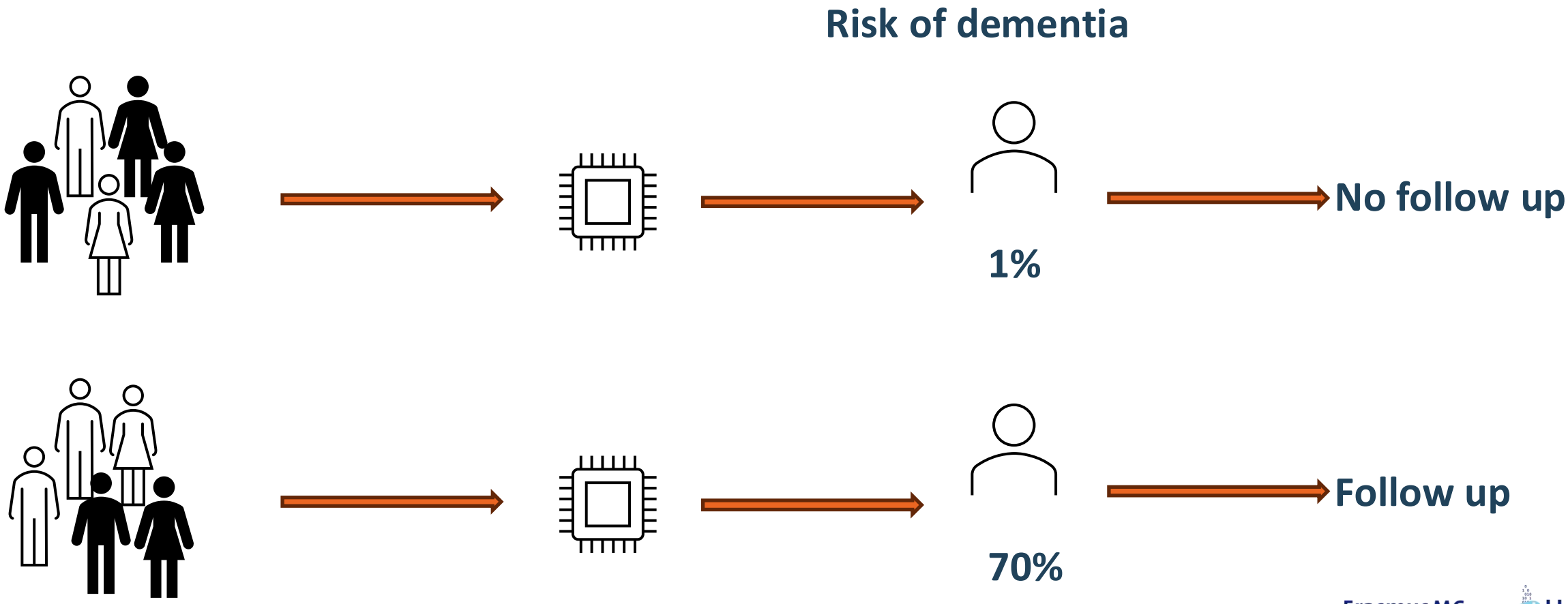
Model stability



Erasmus MC University Medical Center Rotterdam

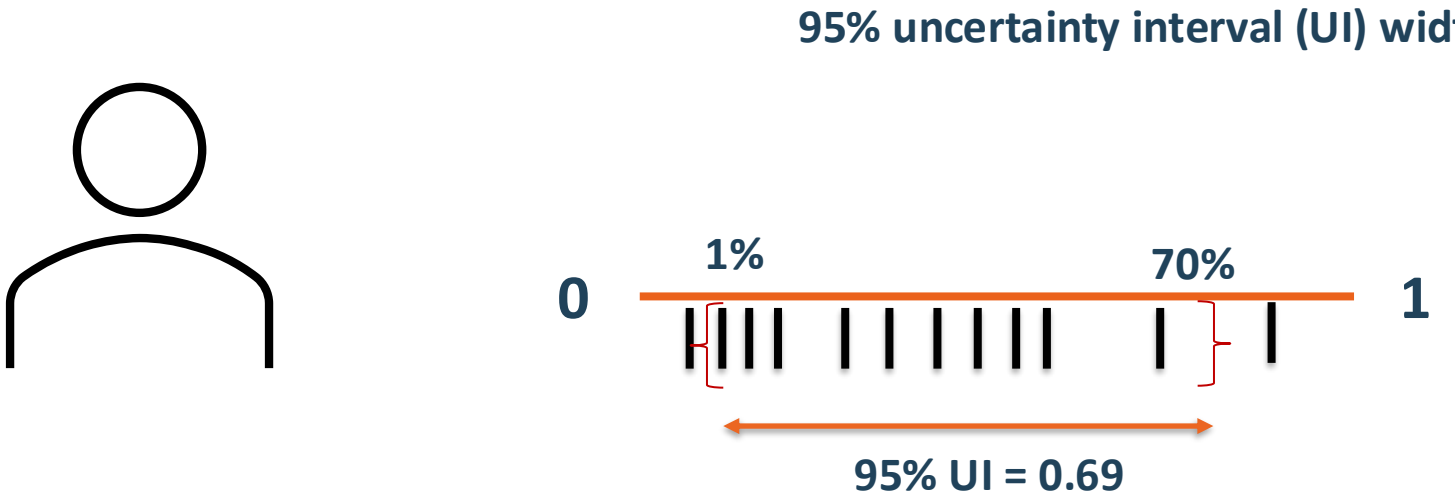
Health Data Science

You have an outpatient visit to the doctor



Model instability

The issue with different predictions…



Ideally, you have both good performance on population level and stability in predictions for reliable evidence

How does federated learning come to play in terms of model stability?

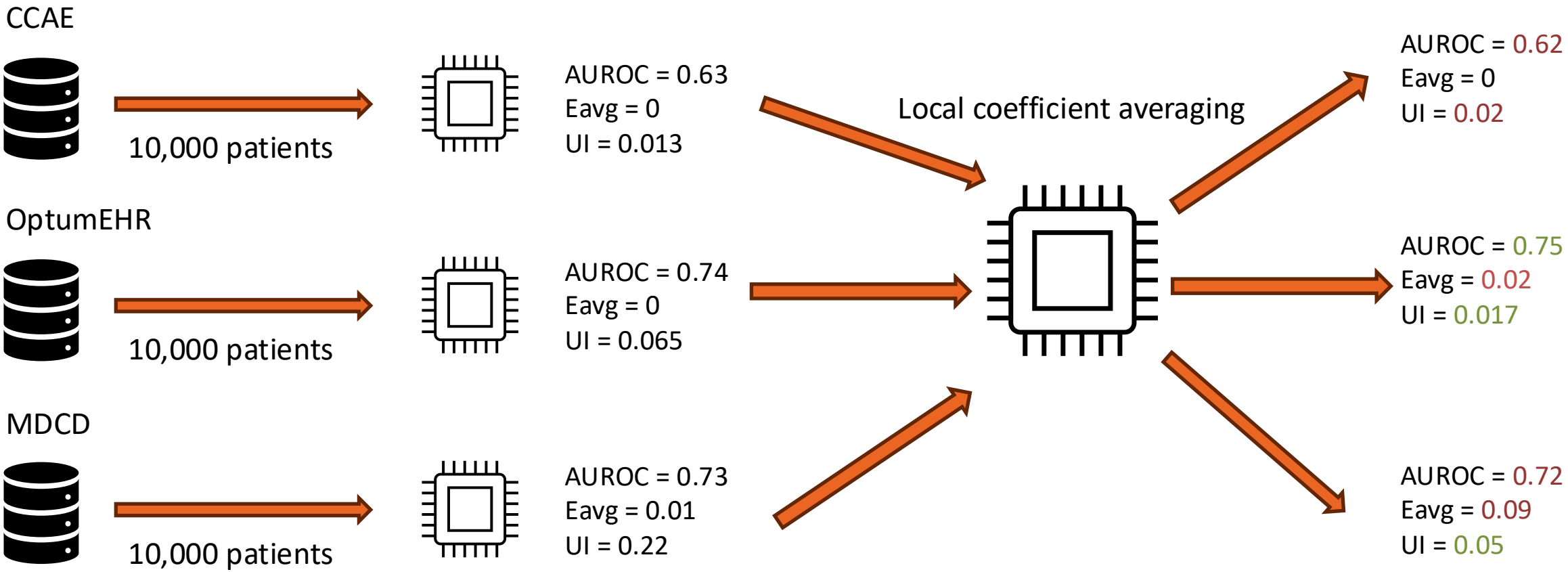
More data generally leads to more stability[1], [2]

Can federated learning, where we have more heterogenous data, also increase model stability?

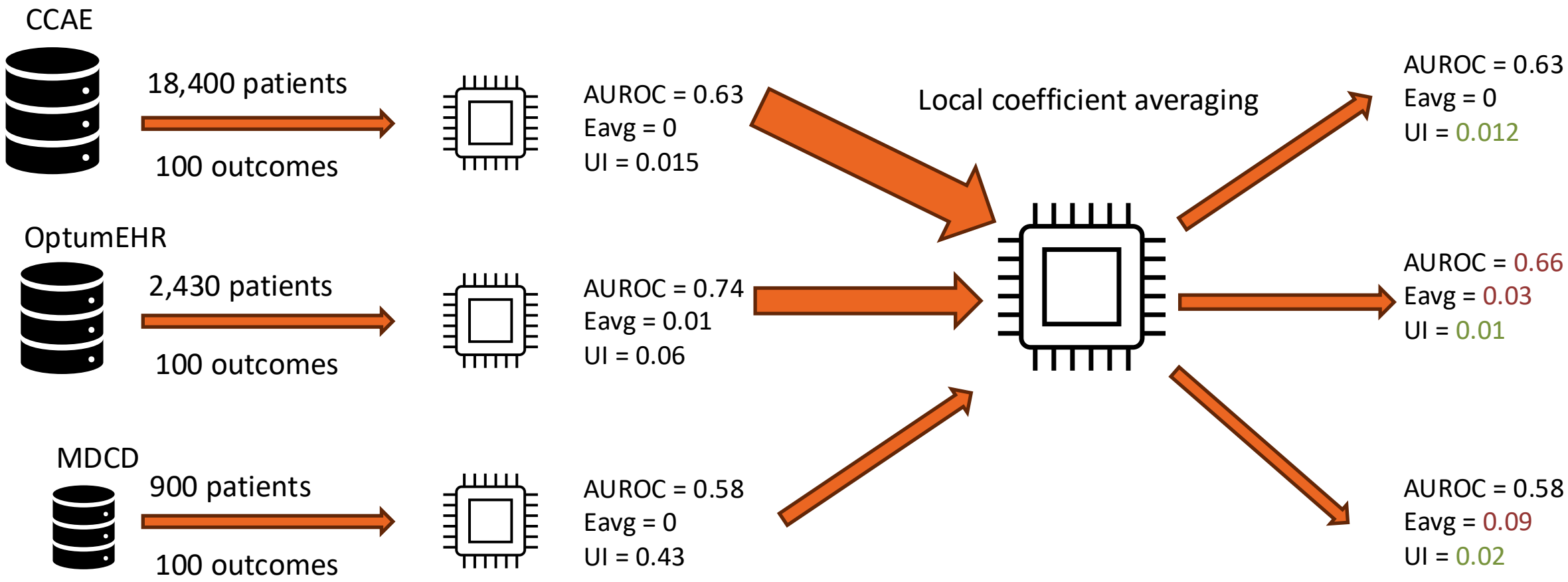
[1] Vereijken F, et al. *Proc. ACM FAccT* 2026:3302–3321. doi:10.1145/3805689.3812316
[2] Legha A, et al. *J Clin Epidemiol* 2025;112117. doi:10.1016/j.jclinepi.2025.112117

Erasmus MC University Medical Center Rotterdam
Health Data Science

Preliminary results



Preliminary results



Do you want to contribute?

How heterogenous can data be for federated learning to improve model stability?

f.vereijken@erasmusmc.nl

Fleur Vereijken



Joint evidence synthesis and bias calibration

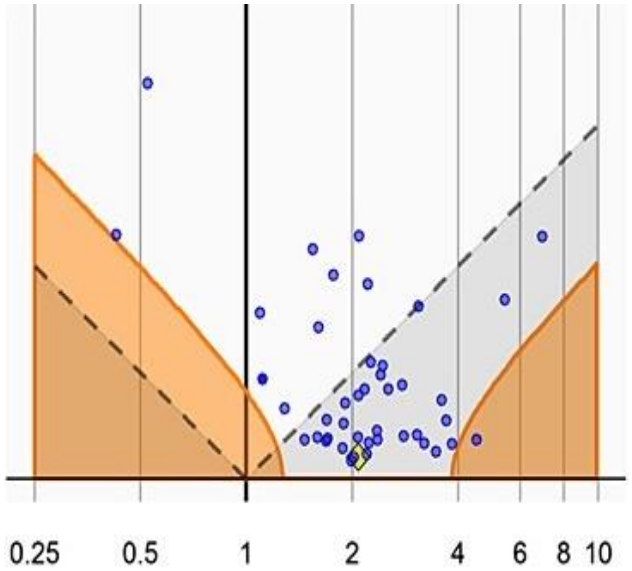
Using profile likelihoods and negative controls across OHDSI data networks

Fan Bu
Department of Biostatistics, University of Michigan

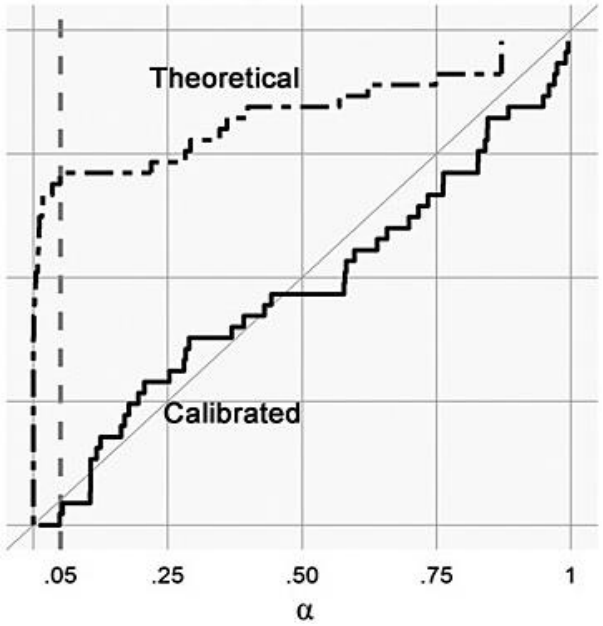
Ongoing work with OHDSI Methods Workgroup

Background: Empirical Calibration

- Negative control outcome (NCO): true effect size = 0
- 50 – 100 NCOs: represents *empirical null*
- Calibrate p-values (or CI) so *empirical Type 1 rate ≈ nominal* (e.g., 0.05)



Blue dots: NCO effect size & SE
Gray: before calibration
Orange: after calibration

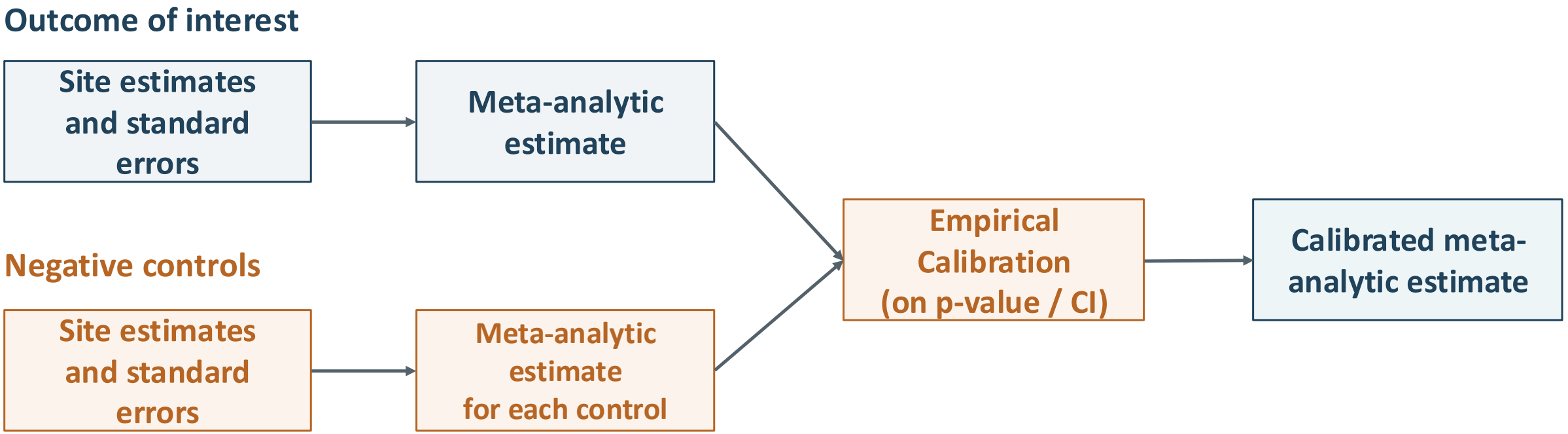


% of NCO with p < α
Before & after calibration

Source: Schuemie et al. (2014).

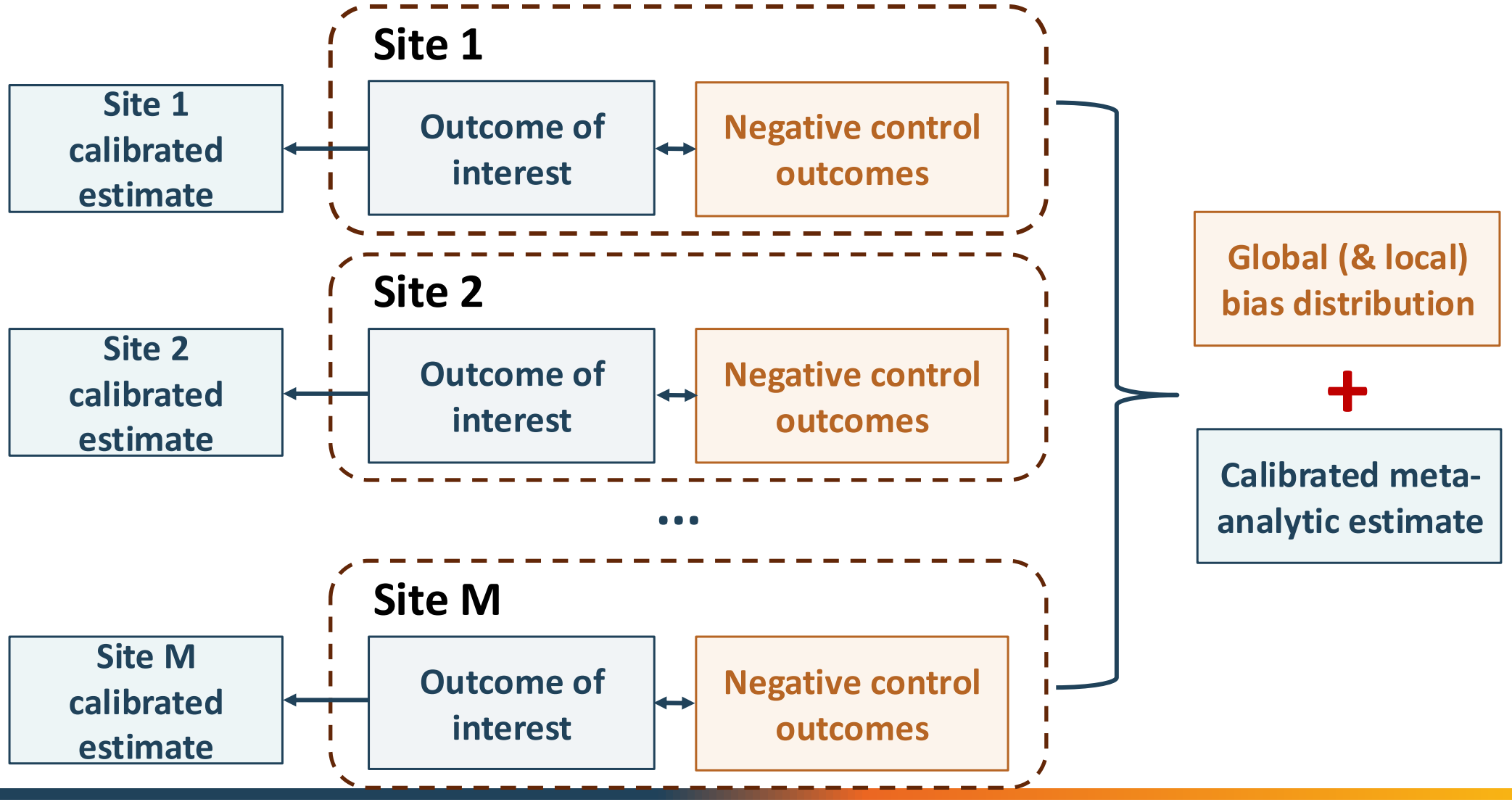
Schuemie et al (2014). Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine.*

Current workflow: synthesis then calibration

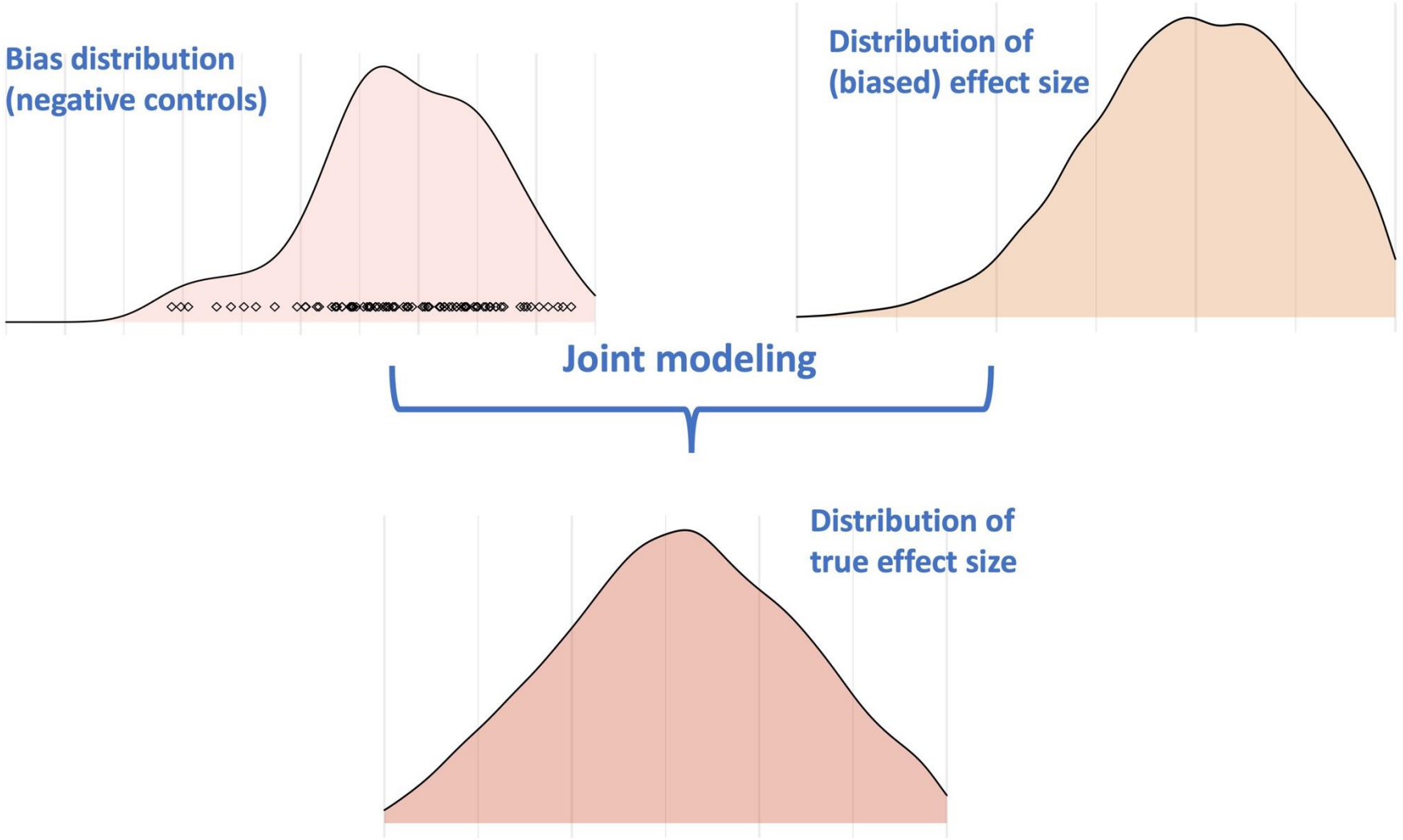


Issue: *separate* global and local calibration

Desired workflow: synthesis AND calibration

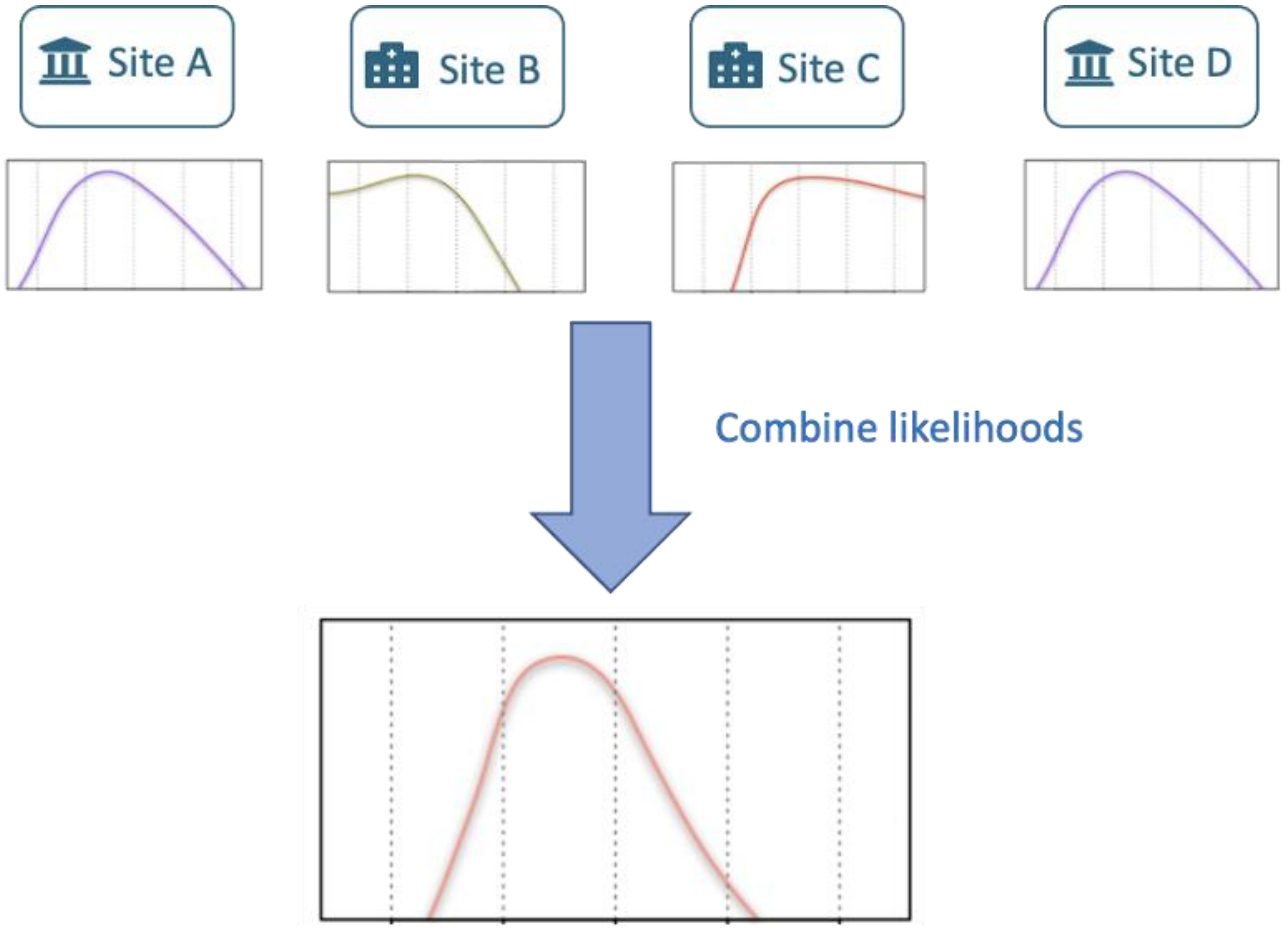


Calibration in a distributional manner (Bayesian)



Bu et al (2024). Bayesian Safety Surveillance with Adaptive Bias Correction. *Statistics in Medicine.*

Synthesis via likelihoods from each site

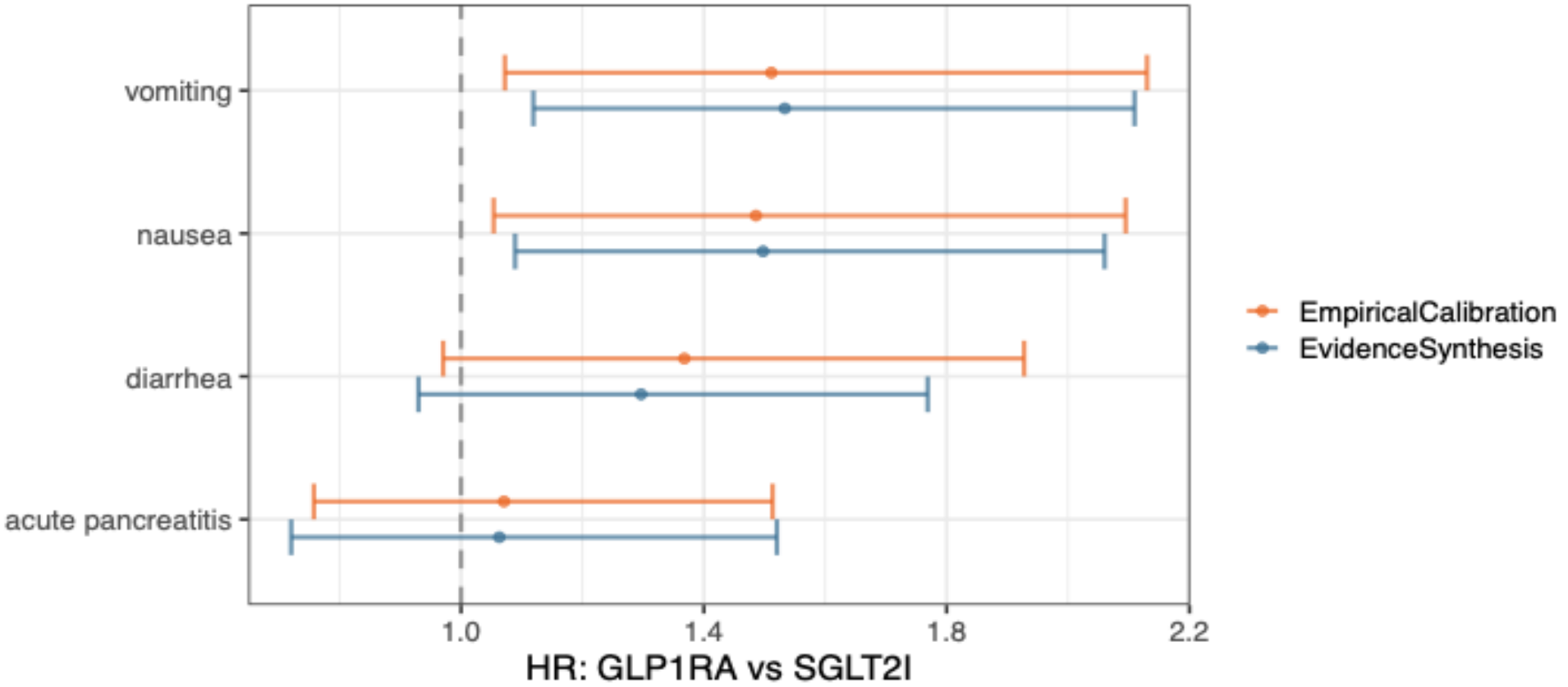


- Fit outcome models locally
- Extract profile likelihood curve for target outcome and negative controls
- Share profile likelihood objects (eval. over discrete grids)
- Retains *more info* & preserves *privacy*
- Profile likelihoods used for synthesis *and* bias learning

Profile likelihood implementation details: Schuemie et al (2022); *Stat Methods Med Research.*

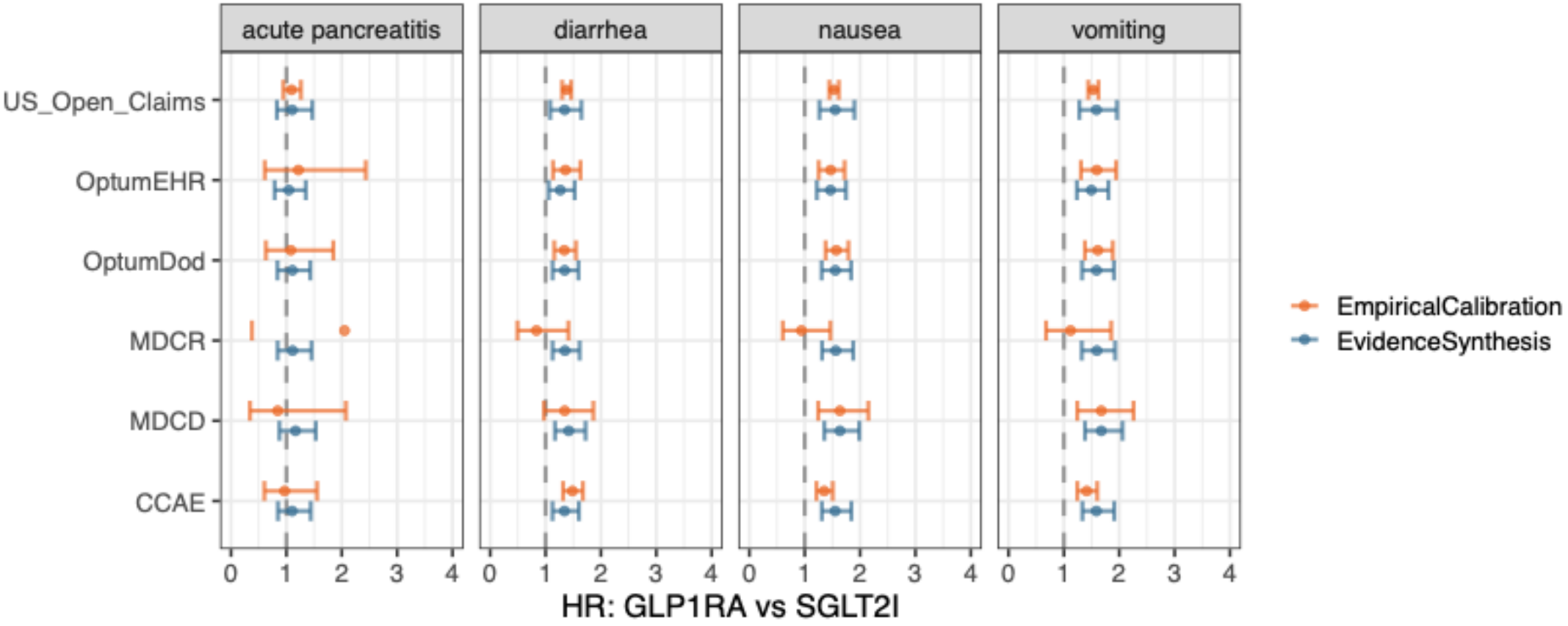
Preliminary example from LEGEND-T2DM

Gastrointestinal safety outcomes between GLP-1RA and SGLT2I new-users.

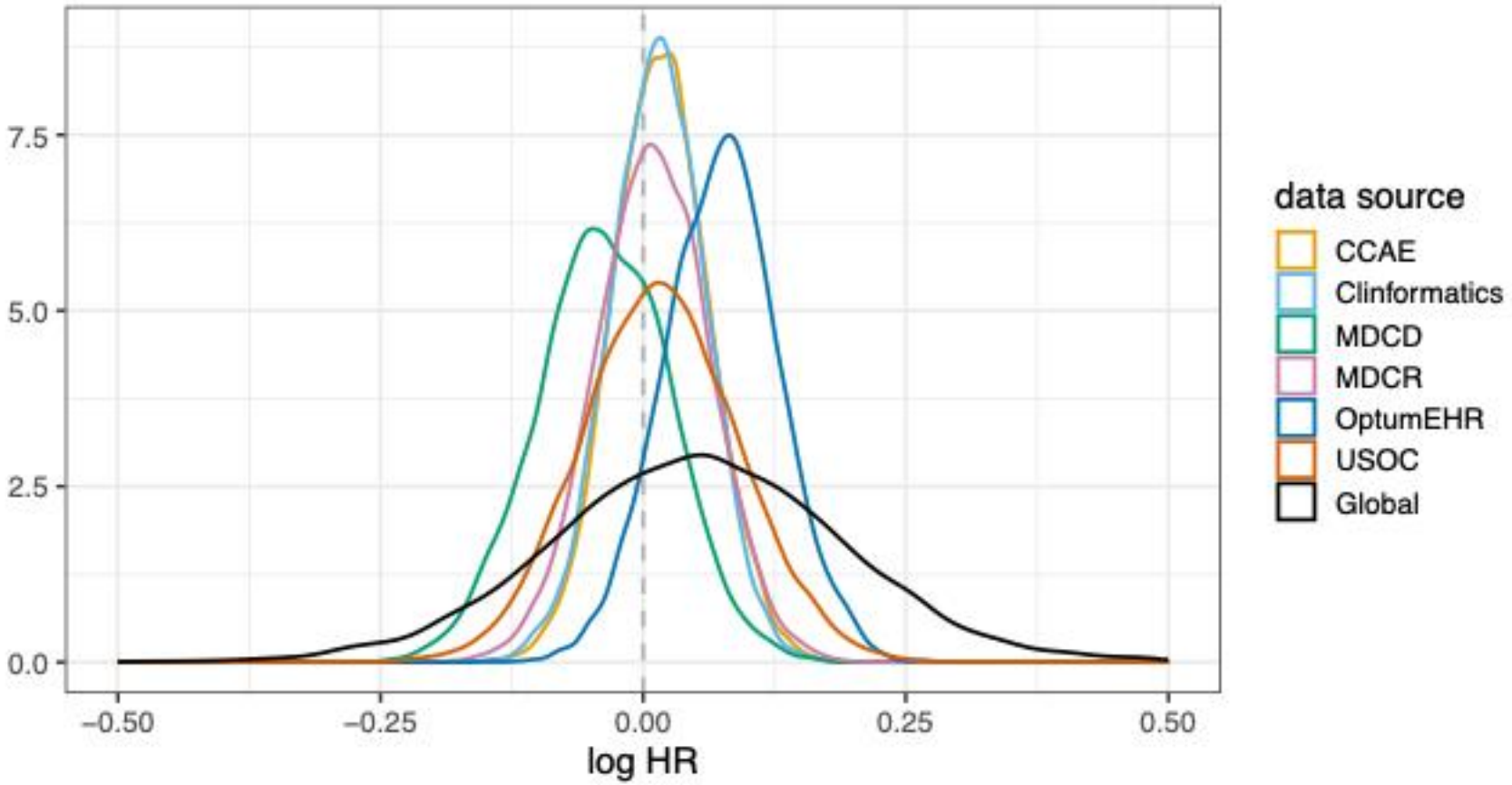


Preliminary example from LEGEND-T2DM

Source-level calibrated HR estimates.



Local & global empirical bias distributions



- Each site has *local bias* (informed by local negative controls)
- Site-level negative controls jointly inform a *shared global* bias distribution

To wrap up …

- Future directions:
  - Higher-dim likelihood object, e.g., tensor-train (Chen et al. 2026)
  - Between negative control correlation

- R package (under development): EvidenceSynthesis

- Thank you to all Methods workgroup & LEGEND collaborators!



Using Systematic Error Reference Outcomes (SERO) to Reduce Bias in Observational Studies

Yumou Qiu (Peking University)

Bias correction and literature reviews

- Outcome that treatments do not affect its values
- Exchangeability: learn bias of primary outcome: Schuemie et al. (2014)
- DAG-based approaches: Shi et al. (2020)
- Historical names: unaffected outcome, with known effects, falsification end points, NCO
- Many challenges, Rosenbaum (1992):

A limitation is evident. If a test for hidden bias based on r_{ij}^* fails to detect a difference, this cannot be equated with the absence of bias, for the performance of the test depends on the magnitude of the bias, the sample size, and the power of the test. Conversely, finding evidence of bias based on r_{ij}^* does not necessarily imply that the bias is large enough to substantially alter conclusions about the outcome r_{ij} of interest. What is needed is a confidence statement linking the observed magnitude of the difference in r_{ij}^* with the uncertainty about biases affecting r_{ij}. This paper develops such a statement.

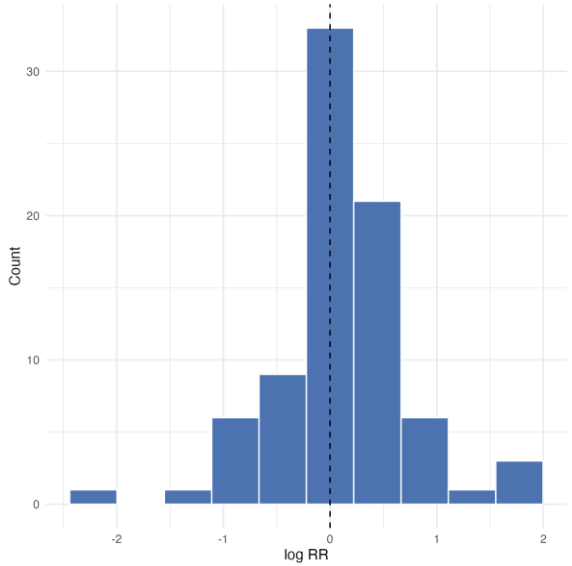
Systematic Error Reference Outcomes (SERO)

- We study distribution of SEROs: understand sources of biases
- Strong definition: $\mathbf{NCO}_i(1) = \mathbf{NCO}_i(0)$
- A weaker condition: $\text{Ave of SERO}_i(1) = \text{Ave of SERO}_i(0)$
- Don't mix the definition with exchangeability with primary outcome

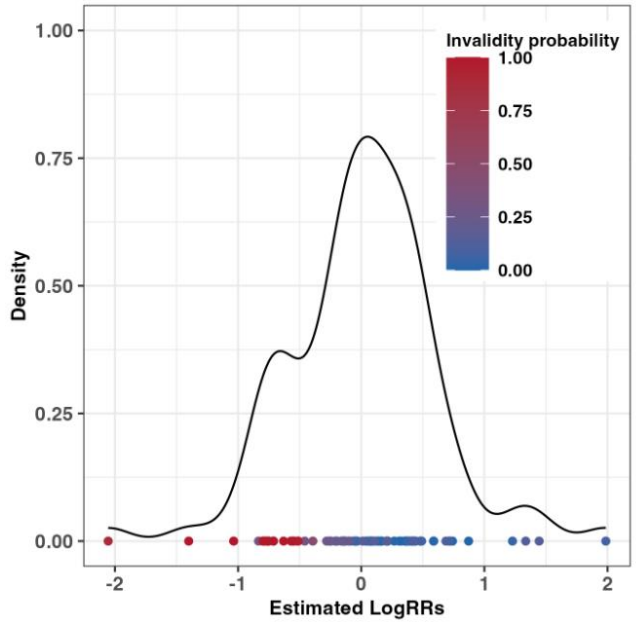
SERO, SERO, Bring bias closer to zero

Empirical distribution of estimated effects

- TriNetX Research Network
- GLP-1RAs vs SGLT2is
- Outcome of interest: acute pulmonary embolism
- SERO: 95 outcomes should not be affected by treatments: injury of knee, foreign body in ear, ingrowing nail, etc.

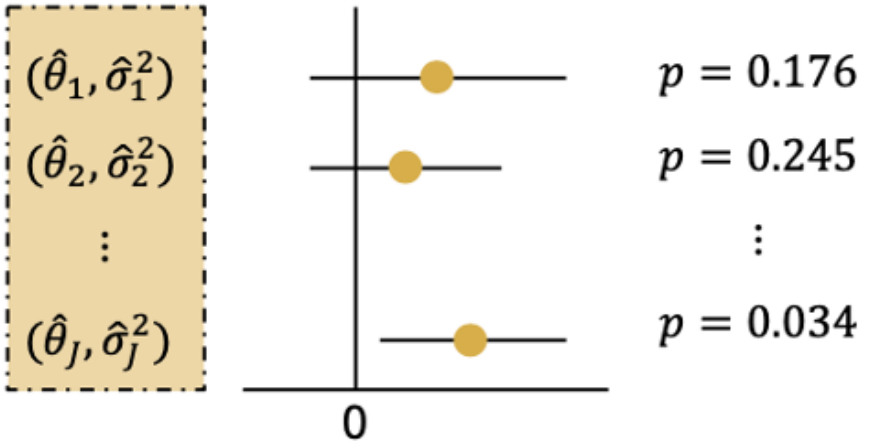


- Same network, same treatment
- Different patients
- Outcome of interest: cardiac arrest and ventricular fibrillation
- Empirical data show mixture normal

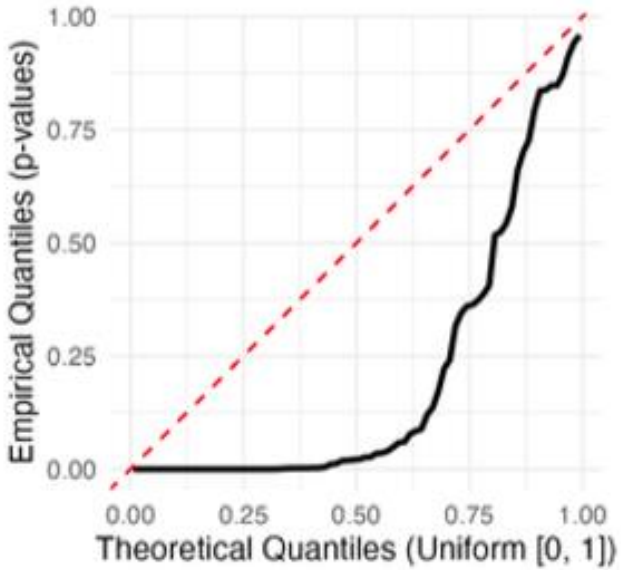


Weaker definition of SERO makes sense !

Diagnosis for zero bias



Empirical results



- Uniform p-values (diagonal line) means no bias
- Deviate from diagonal: systematic bias
- Estimated effect of SERO significantly $\neq 0 \Rightarrow$ Bias !

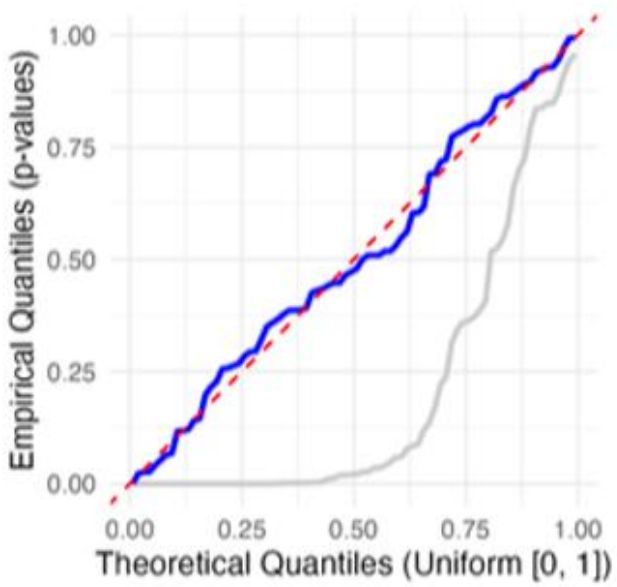
Where biases and variations come from?

- From data collection (sampling bias, …)
- From treatment receipt (unmeasured confounding bias)
- From estimation (improper modeling and methods, …)
- Variation of estimates (inevitable)
- Some SEROs may be invalid

Diagnosis for normal distribution of biases

- Mean, variances of biases of all SEROs
- p-value under this fitted normal
- Deviation from diagonal line:

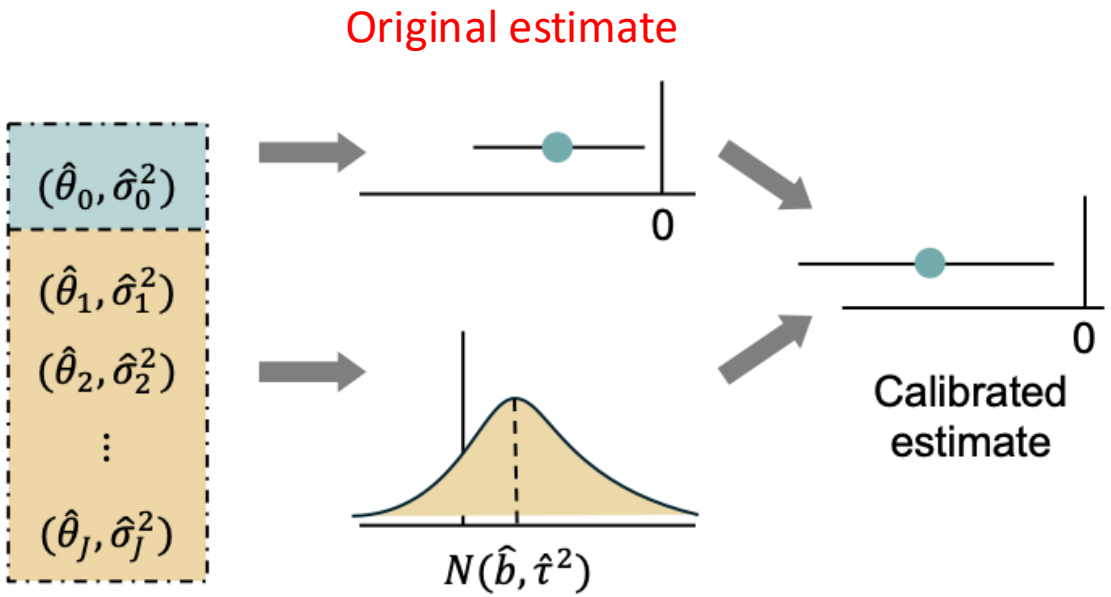
Empirical results



Biases of SEROs NOT follow a Normal

- ➢ Invalid SEROs
- ➢ SEROs have different source of bias

Calibrating lump-sum bias: normal distribution



1. Calibrated estimate

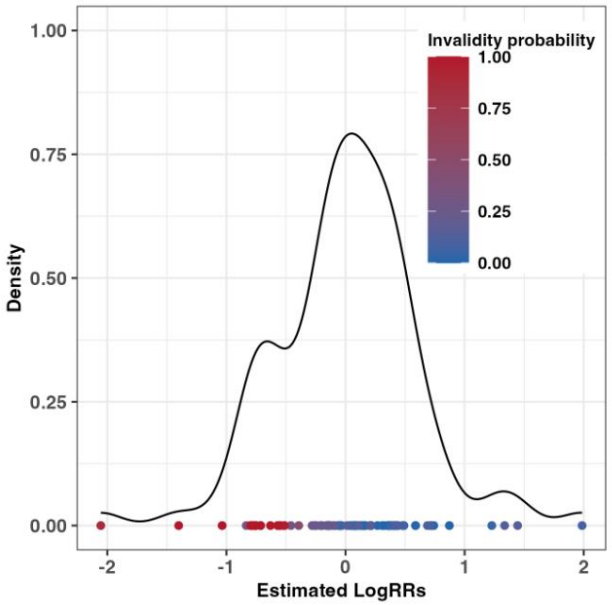
$$= \text{Original estimate} - \text{Ave biases of SEROs}$$

2. Confidence interval:

$$\text{Calibrated estimate} \pm 2\sqrt{\hat{\tau}^2 + \hat{\sigma}_0^2}$$

Var of SERO biases Est. var

Identifying invalid SEROs (convolution structure)



Bias distribution is NOT normal

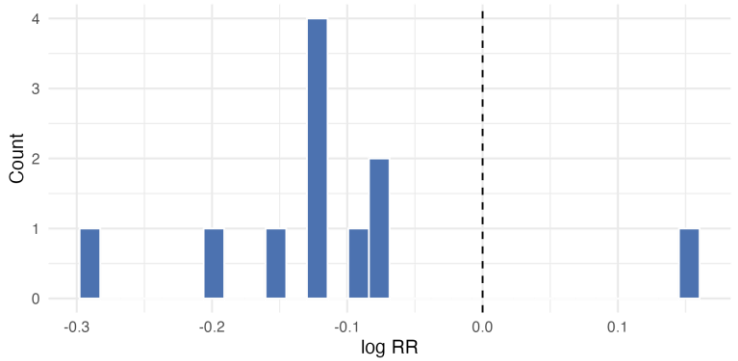
- Est. of SERO = causal effect + bias
- True causal effects β_j: zero for valid BROs;
- Invalid SERO: effect not zero, distribution G
- Biases b_j follow distribution F
- $\beta_j^* = \beta_j + b_j \sim H := \pi F + (1-\pi)(F * G)$
- $F, G,$ and π can be consistently estimated

Zhu, Y., Hu, W., Zhang, B., Hripcsak, G., Qiu, Y.*, and Chen, Y*. (2026) Robust Causal Inference for Observational Data under Residual Bias. *Journal of the American Statistical Association*. (under review)

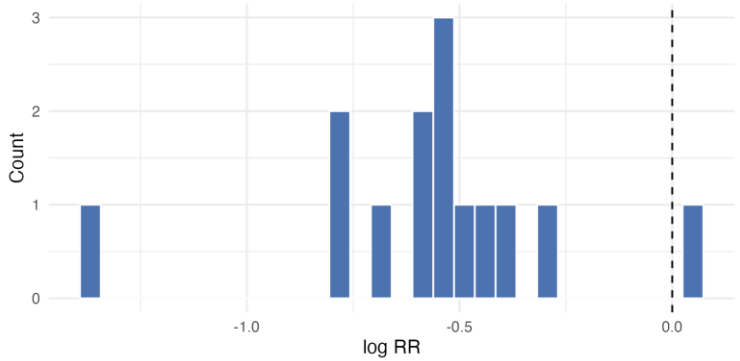
Bias distributions over multiple sites

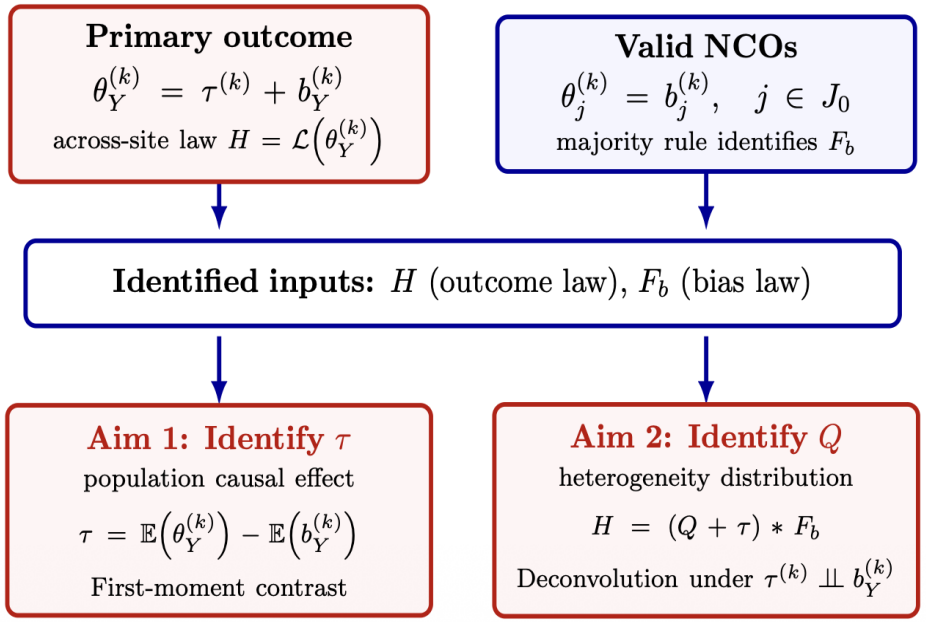
LEGENG-T2DM: 14 sites, DPP4is vs SUs

Histogram across sites for a SERO



Histogram for target outcome (heart failure)





- More than half SEROs follow bias distribution F_b (across sites)
- ATE τ and across-site variation Q can be consistently estimated

Hu, W., Zhu, Y., Zhang, B., Schuemie, M., Qiu, Y.*, and Chen, Y*. (2026) Multi-site Inference of Causal Effect in the Presence of Unmeasured Confounding. *Biometrika*. (under review)

RWE Studies - Causal inference/multi-site studies

JAMA Network Open

Research Letter | Diabetes and Endocrinology

Incretin Receptor Agonists and CPAP Use in Adults With Diabetes, Obesity, and Obstructive Sleep Apnea

Huilin Tang, PhD, MS; Bingyu Zhang, MS; Yiwen Lu, BS; Dazheng Zhang, PhD; Yuan Lu, ScD; David A. Asch, MD, MBA; Yong Chen, PhD

THE LANCET
Infectious Diseases

ARTICLES · Online first, September 30, 2025

Long COVID associated with SARS-CoV-2 reinfection among children and adolescents in the omicron era (RECOVER-EHR): a retrospective cohort study

Bingyu Zhang, MS [a,b] · Qiong Wu, PhD [a,c,d] · Prof Ravi Jhaveri, MD [f] · Ting Zhou, MD PhD [a,c] · Prof Michael J Becich, MD PhD [e] · Yuriy Bisyuk, MD PhD [g] · Frank Blancero, BA [h] · Prof Elizabeth A Chrischilles, PhD [i] · Prof Cynthia H Chuang, MD MSc [j] · ll, MS PhD [k] · Daniel Fort, PhD MPH [l] · Prof Carol R Horowitz, MD MPH [m] · Susan Kim, MD MMSc [n] ·

nature communications

Explore content · About the journal · Publish with us

nature › nature communications › articles › article

Article | Open access | Published: 24 July 2025

Risk of neuropsychiatric and related conditions associated with SARS-CoV-2 infection: a difference-in-differences analysis

Yiwen Lu, Jiayi Tong, Dazheng Zhang, Jiajie Chen, Lu Li, Yuqing Lei, Ting Zhou, Leyna V. Aragon, Michael J. Becich, Saul Blecker, Nathan J. Blum, Dimitri A. Christakis, Mady Hornig, Maxwell M. Hornig-Rohan, Ravi Jhaveri, W. Schuyler Jones, Amber Brown Keebler, Kelly Kelleher, Susan Kim, Abu Saleh Mohammad Mosa, Kathleen Pajer, Jonathan Platt, Hayden T. Schwenk, Bradley W. Taylor, ... Yong Chen

nature communications

Explore content · About the journal · Publish with us

nature › nature communications › articles › article

Article | Open access | Published: 11 April 2025

Cardiovascular post-acute sequelae of SARS-CoV-2 in children and adolescents: cohort study using electronic health records

Bingyu Zhang, Deepika Thacker, Ting Zhou, Dazheng Zhang, Yuqing Lei, Jiajie Chen, Elizabeth A. Chrischilles, Dimitri A. Christakis, Soledad Fernandez, Vidu Garg, Susan Kim, Abu S. M. Mosa, Marion R. Sills, Bradley W. Taylor, David A. Williams, Qiong Wu, Christopher B. Forrest & Yong Chen

The New York Times

Long Covid Risk for Children Doubles After a Second Infection, Study Finds

The new research provides evidence that Covid reinfections can increase the risk of long-term health consequences.

Annals of Internal Medicine

ORIGINAL RESEARCH

Real-World Effectiveness of BNT162b2 Against Infection and Severe Diseases in Children and Adolescents

Qiong Wu, PhD*; Jiayi Tong, MS*; Bingyu Zhang, MS; Dazheng Zhang, MS; Jiajie Chen, PhD; Yuqing Lei, MS; Yiwen Lu, BS; Yudong Wang, PhD; Lu Li, BA; Yishan Shen, MS; Jie Xu, PhD; L. Charles Bailey, MD, PhD; Jiang Bian, PhD; Dimitri A. Christakis, MD, MPH; Megan L. Fitzgerald, PhD; Kathryn Hirabayashi, MPH; Ravi Jhaveri, MD; Alka Khaitan, MD; Tianchen Lyu, MS; Suchitra Rao, MBBS, MSCS; Hanieh Razzaghi, PhD, MPH; Hayden T. Schwenk, MD, MPH; Fei Wang, PhD; Margot I. Gage Witvliet, PhD; Eric J. Tchetgen Tchetgen, PhD; Jeffrey S. Morris, PhD†; Christopher B. Forrest, MD, PhD†; and Yong Chen, PhD†

Annals of Internal Medicine®

Search this publica

LATEST ISSUES IN THE CLINIC FOR HOSPITALISTS MULTIMEDIA COLLECTIONS ACP JOURNAL CLUB NEWS CME /

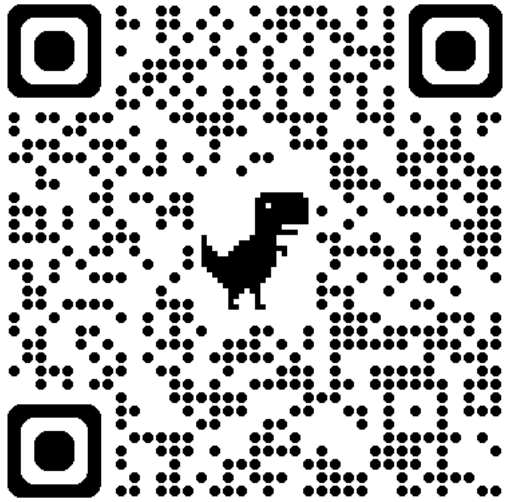
Original Research | 25 February 2025

Glucagon-Like Peptide-1 Receptor Agonists and Risk for Depression in Older Adults With Type 2 Diabetes: A Target Trial Emulation Study

Authors: Huilin Tang, PhD , Ying Lu, BA, William T. Donahoo, MD , Sarah C. Westen, PhD , Yong Chen, PhD , Jiang Bian, PhD , and Jingchuan Guo, MD, PhD AUTHOR, ARTICLE, & DISCLOSURE INFORMATION

Join the collaboration!

debiasedTrialEmulation



Module A: Preprocessing				
Time zero	Eligibility	Treatment strategy	Follow-up	...

Module B.1: Data input				
Treatment/ Exposure	Covariate	Outcome	Negative control outcomes	...

Module C.1 : Main function (Assume perfect data)	
PS method	Matching
	Stratification
	Weighting
Outcome measures	RR
	OR
	HR
Outcome model	GLM

Module C.2: Main function (Data with issues)	
Immortal time bias	Sequential TTE
	Clone-censoring-weight
Unmeasured confounder	Distributional NCO calibration
Information bias	Misclassification in treatment
	Misclassification in covariates
	Misclassification in outcomes
Selection bias	
Informative censoring	Loss to follow-up
	Treatment deviation
Estimands selection	Intention-to-treat/as-treated/per-protocol

Module D: Result output
Cohort after PS matching/stratification/weighting
Results (outcomes, uncalibrated)
Results (NCOs)
Results (outcomes, calibrated)

Module E: Diagnosis and sensitivity analysis			
Standard mean difference of covariates	Equipoise plot of covariates	EASE plot of NCOs	...

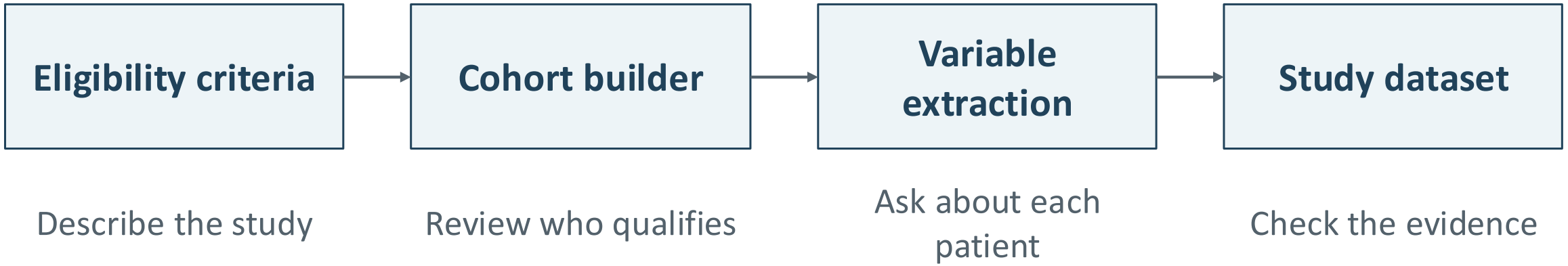


AI for reliable evidence in distributed networks

Bingyu Zhang (University of Pennsylvania)

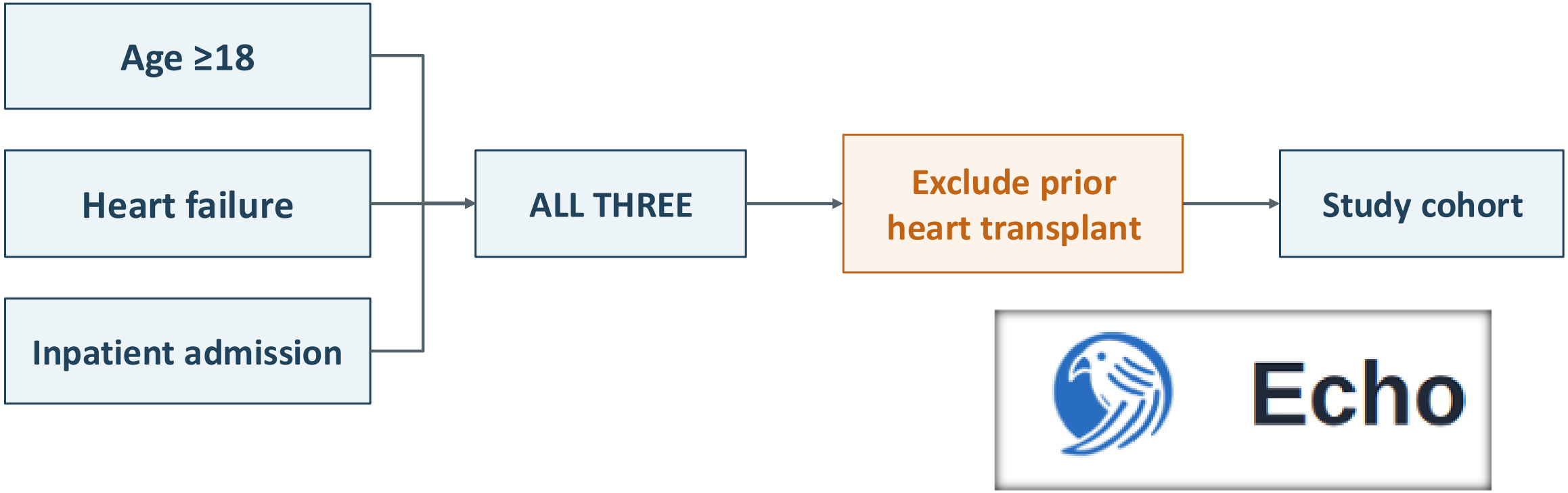
From eligibility criteria to analysis-ready cohorts is not straightforward

How much time does your team spend building the cohort?

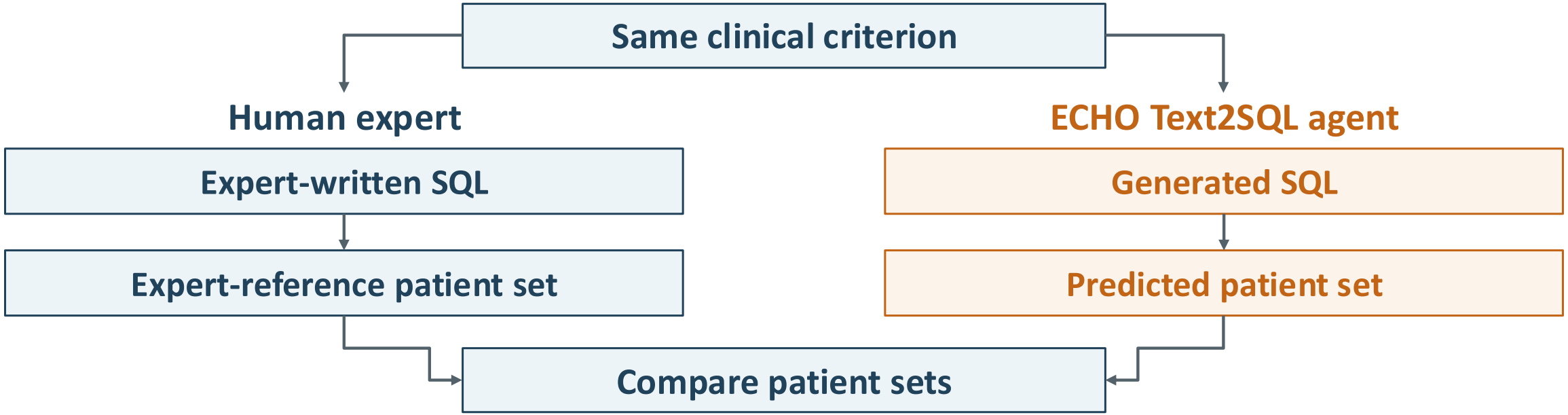


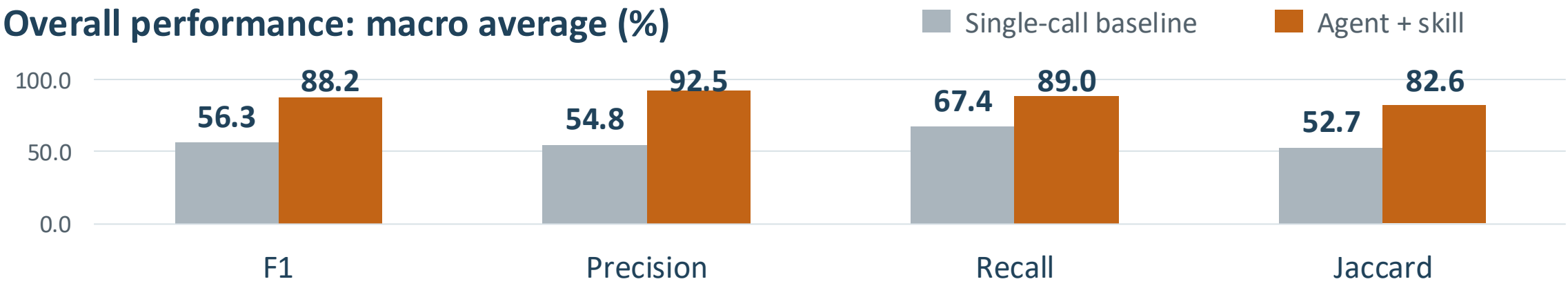
How do clinical criteria become a patient cohort?

“Adults hospitalized with heart failure, excluding prior heart transplant”



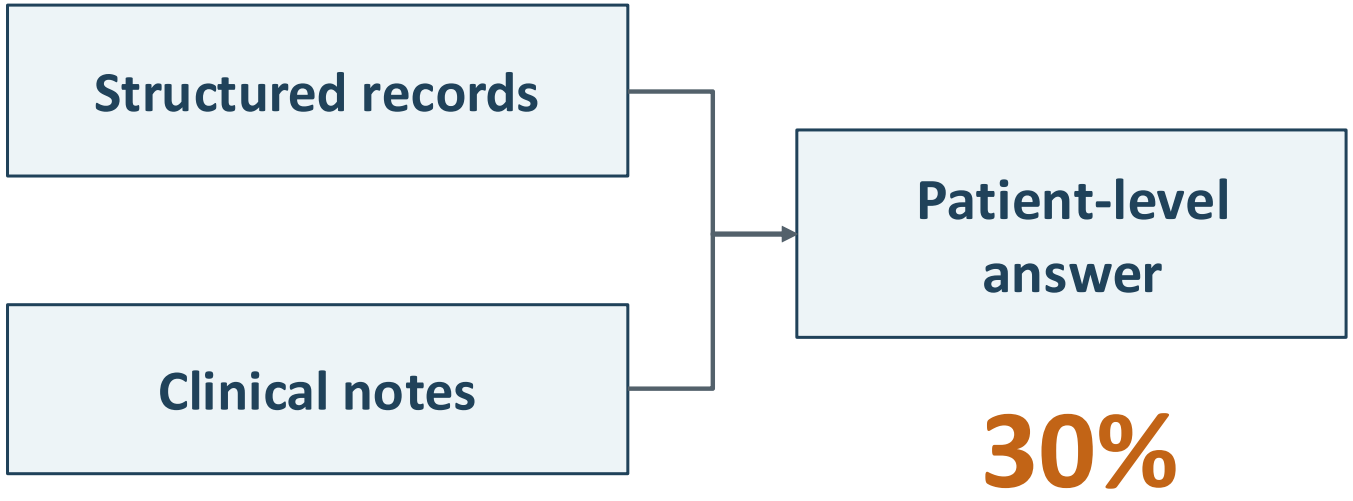
Does AI find the patients an expert would select?





Can we trace each extracted answer to the record?

“What was the lowest documented ejection fraction during this admission?”



Supporting source records

CONFIDENCE High

EVIDENCE

NOTE note N-100201-01

Admission echocardiogram: LVEF **30%**.

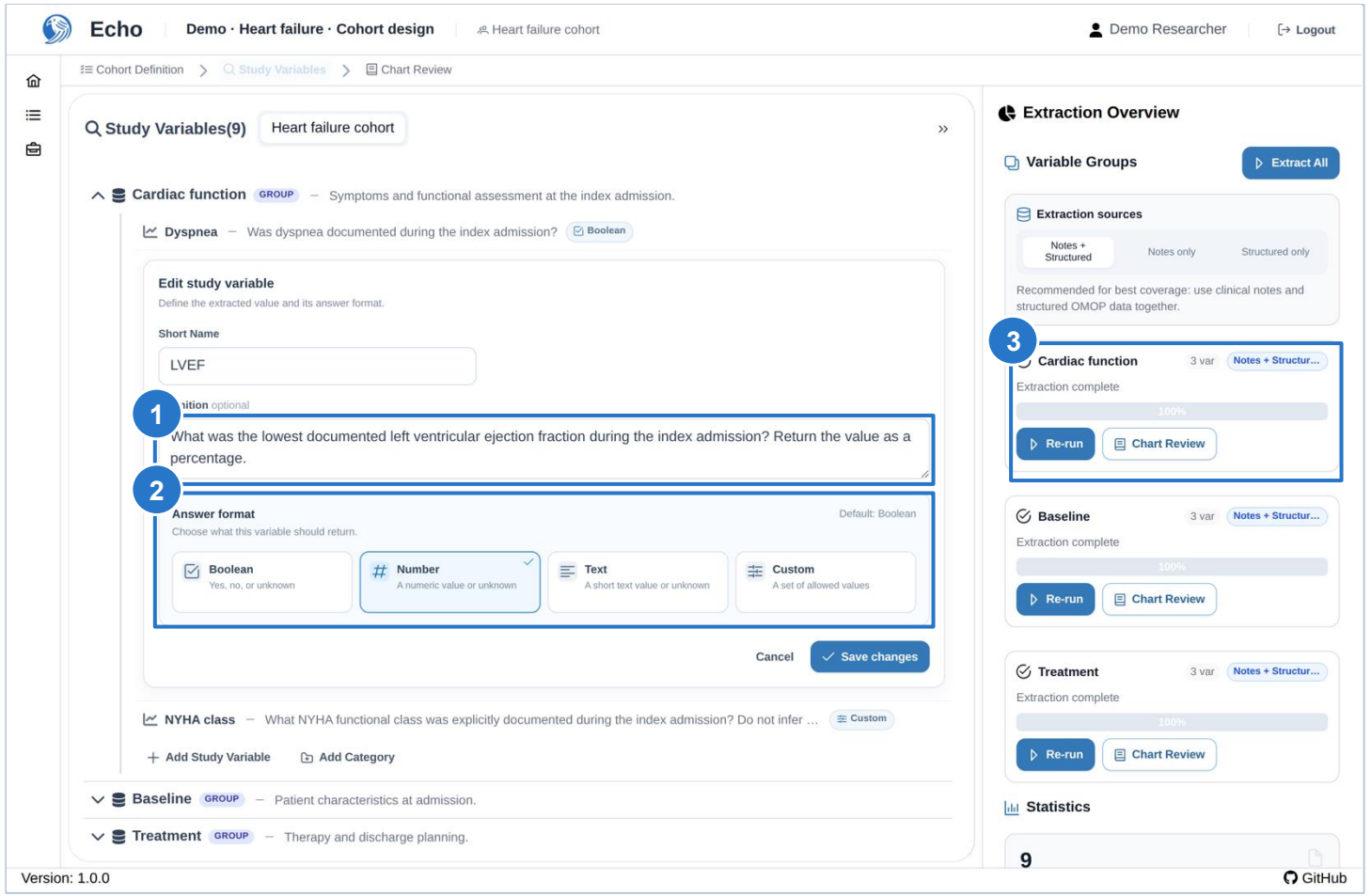
NOTE note N-100201-02

Cardiology follow-up: LVEF **35–40%**.

REASONING

Both measurements fall within the index admission. The question requests the lowest documented LVEF; 30% is lower than 35–40%. No value was inferred.

Explore it!



1 Define a study variable

Describe study variable in natural language.

2 Specify the answer

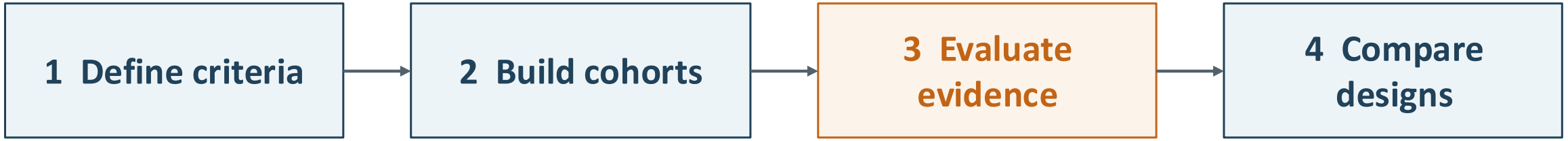
Choose a number, yes/no, free text, or a custom set of allowed values.

3 Extract and review

Run agent-assisted extraction by group, then open Chart Review to inspect answers and evidence.

TrialMap: an end-to-end solution for clinical trial design

What happens to the evidence when we change eligibility?



Translate protocol language into computable rules

Apply original and expanded eligibility

Estimate outcomes and calibrate with SEROs

Number of eligible patients, benefit, safety, and reliability

One example throughout: expanding eligibility from ECOG 0-1 to ECOG 0-2

What if we included ECOG 2?

ECOG performance status describes a patient's daily functioning

0	1	2
Fully active	**Some activity limits**	**Self-care remains possible**
No restriction in usual activity	Walking and light work remain possible	Unable to work, but up more than half the day

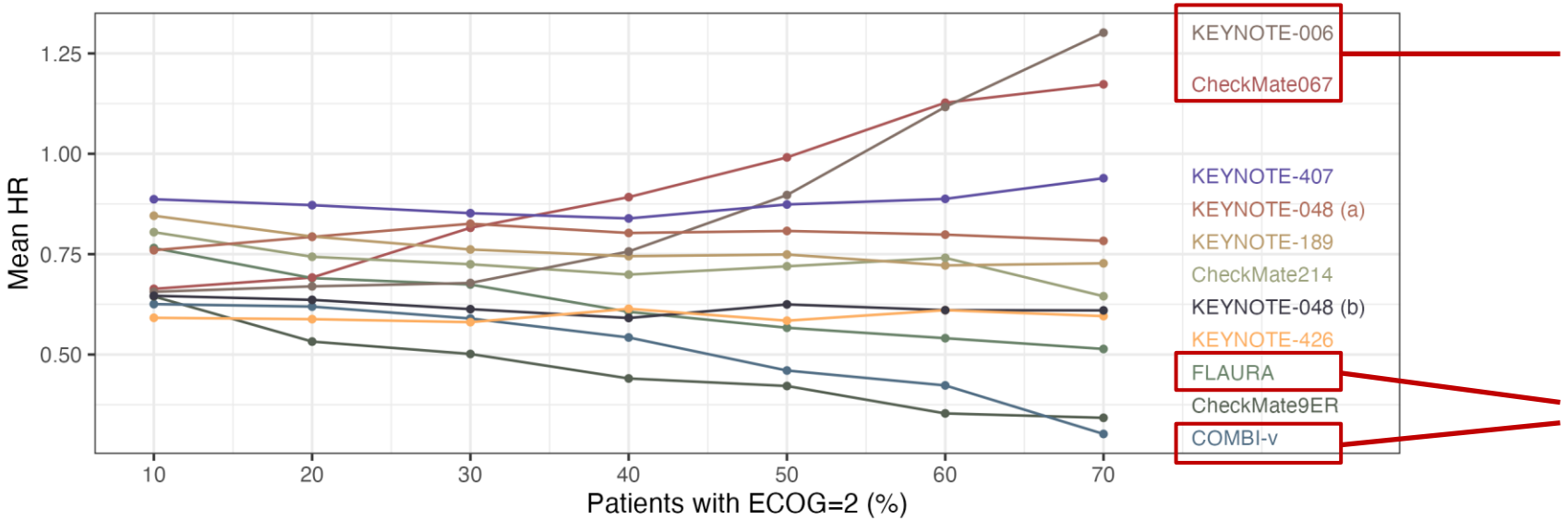
Original eligibility: 0-1

Expanded eligibility: 0-2

Does broader eligibility change estimated benefit?

Real-world emulations of 10 oncology trials across 4 cancer types

What happens as the proportion with ECOG 2 increases?



Melanoma immunotherapy

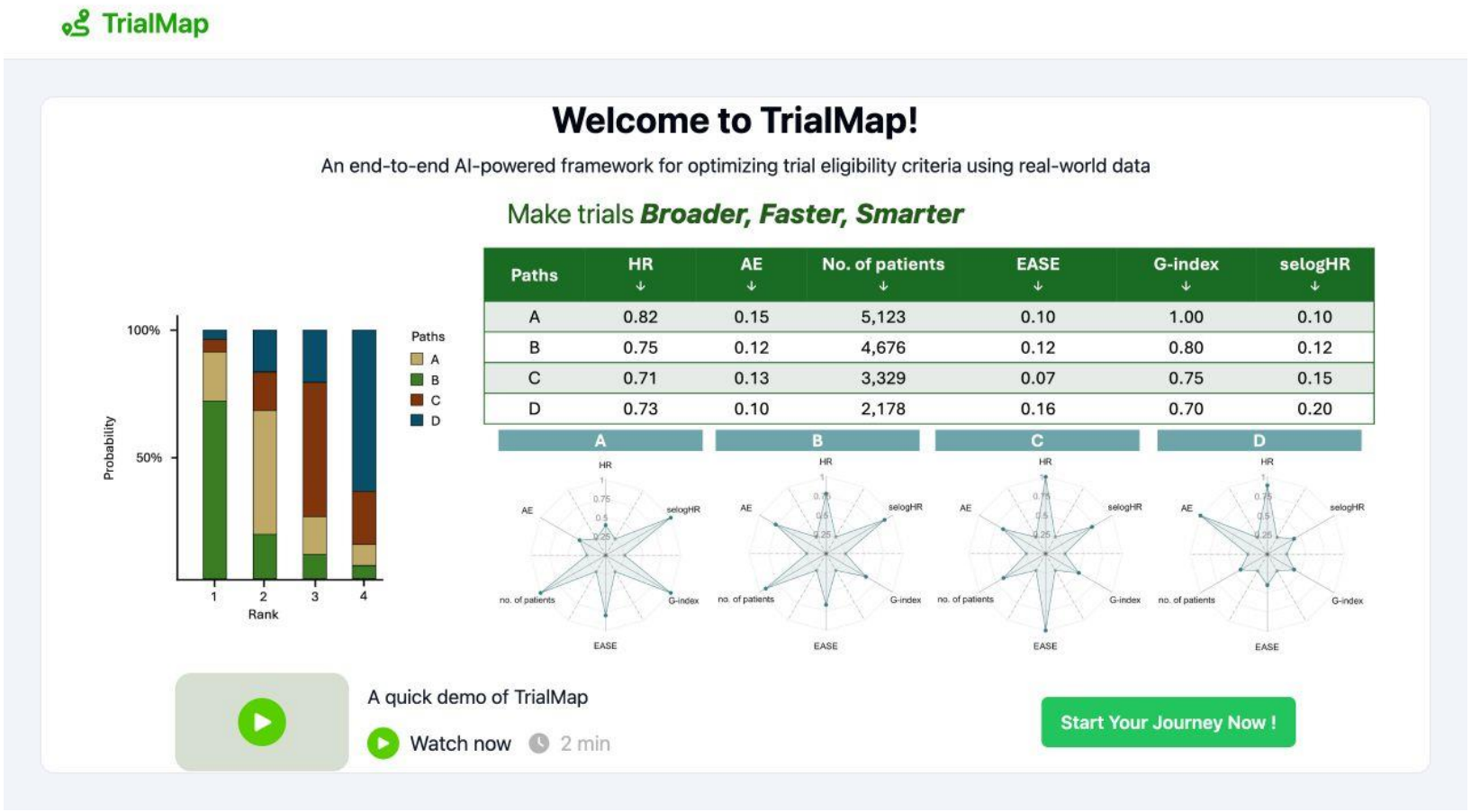
Estimated benefit attenuates

Lung cancer targeted therapy

Relatively stable estimates

Broader eligibility needs evidence on benefit, safety, and bias

Explore the design choices



Try Now!

clinicalnlp.org/TrialMap/

Reliable evidence is a network effort

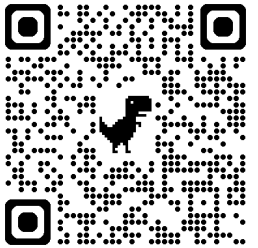
PDA



EvidenceSynthesis



debiasedTrialEmulation



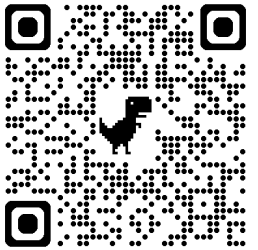
Today's plenary connects the pieces

Yong	SOTA Methods for learning across sites
Chongliang	Use cases that bring the methods into practice
Fleur	Stability of model predictions
Fan	Joint evidence synthesis and bias calibration
Yumou	SEROs that help reveal bias
Bingyu	AI for learning and study implementation

ECHO



TrialMap



What evidence question could we answer better together?

Thank you to our collaborators and the OHDSI community



Demo 1: MOSAiC for Federated logistic regression

JOURNAL ARTICLE

Grid Binary LOgistic REgression (GLORE): building shared models without sharing data

Yuan Wu ✉, Xiaoqian Jiang, Jihoon Kim, Lucila Ohno-Machado

Journal of the American Medical Informatics Association, Volume 19, Issue 5, September 2012, Pages 758–764, https://doi.org/10.1136/amiajnl-2012-000862

Published: 17 April 2012 **Article history** ▾

JOURNAL ARTICLE

Learning from electronic health records across multiple sites: A communication-efficient and privacy-preserving distributed algorithm

Get access

Rui Duan, Mary Regina Boland, Zixuan Liu, Yue Liu, Howard H Chang, Hua Xu, Haitao Chu, Christopher H Schmid, Christopher B Forrest, John H Holmes, Martijn J Schuemie, Jesse A Berlin, Jason H Moore, Yong Chen ✉

Journal of the American Medical Informatics Association, Volume 27, Issue 3, March 2020, Pages 376–385, https://doi.org/10.1093/jamia/ocz199

Published: 09 December 2019 **Article history** ▾

Article | Open access | Published: 15 July 2025

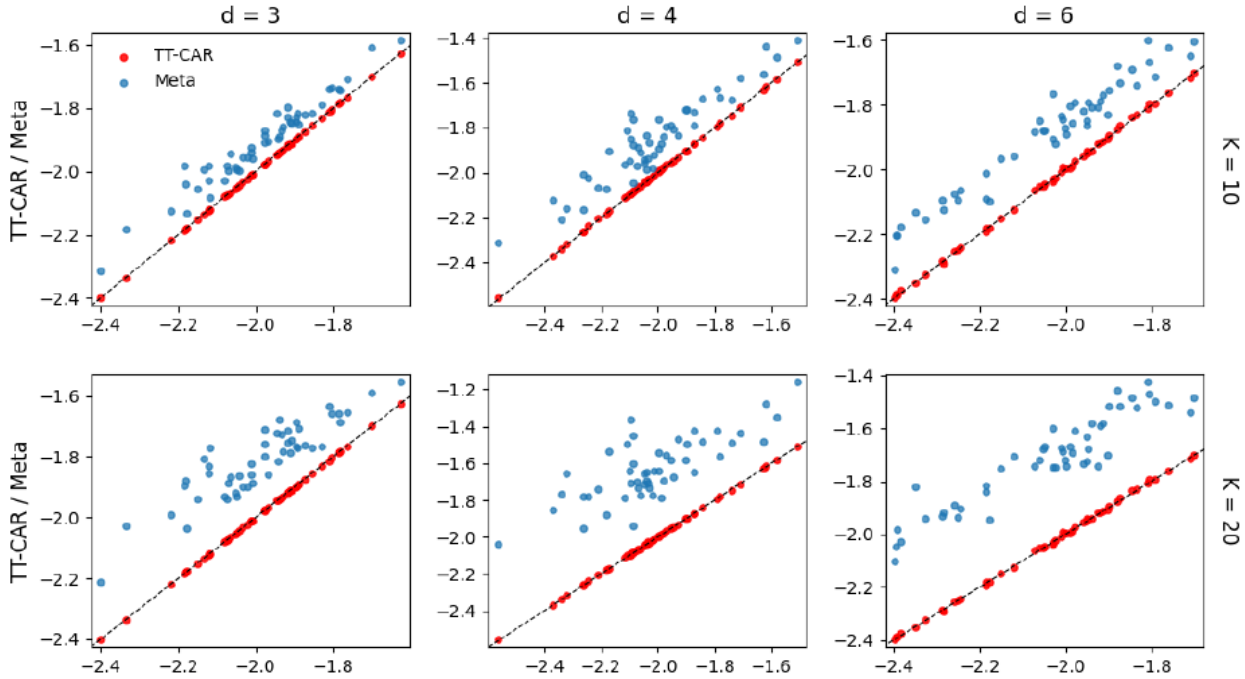
COLA-GLM: collaborative one-shot and lossless algorithms of generalized linear models for decentralized observational healthcare data

Qiong Wu, Jenna M. Reps, Lu Li, Bingyu Zhang, Yiwen Lu, Jiayi Tong, Dazheng Zhang, Thomas Lumley, Milou T. Brand, Mui Van Zandt, Thomas Falconer, Xing He, Yu Huang, Haoyang Li, Chao Yan, Guojun Tang, Andrew E. Williams, Fei Wang, Jiang Bian, Bradley Malin, George Hripcsak, Martijn J. Schuemie, Yun Lu, Steve Drew, ... Yong Chen ✉ + Show authors

npj Digital Medicine **8**, Article number: 442 (2025) | Cite this article

Chen, Lu, Qiu, Wang, Wang, 2026. MOSAiC JASA-ACS (discussion paper)

Demo 1: MOSAiC for Federated GLM



MOSAiC provides a one-shot and lossless solution for GLMs, allowing site-specific nuisance parameters.

Approach	One-shot?	Lossless?	Support model heterogeneity?	Independent of initial estimators?	Model type	Covariate type
GLORE	✗	✓	✗	✓	Logistic	All
ODAL	✓	✗	✗	✗	Logistic	All
COLA-GLM	✓	✓	✗	✓	All	Categorical
MOSAiC	✓	✓	✓	✓	All	All

Chen, Lu, Qiu, Wang, Wang, 2026. MOSAiC JASA-ACS (discussion paper)

Demo 2: MOSAiC for Stratified Cox Models

JOURNAL ARTICLE

WebDISCO: a web service for distributed cox model learning without patient-level data sharing

Get access

Chia-Lun Lu, Shuang Wang ✉, Zhanglong Ji, Yuan Wu, Li Xiong, Xiaoqian Jiang, Lucila Ohno-Machado

Journal of the American Medical Informatics Association, Volume 22, Issue 6, November 2015, Pages 1212–1219, https://doi.org/10.1093/jamia/ocv083

Published: 09 July 2015 **Article history** ▾

JAMIA
A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

AMIA
INFORMATICS PROFESSIONALS. LEADING THE WAY.

Issues | More Content ▾ | Submit ▾ | Purchase | Alerts | About ▾ | AI Discovery Assistant



Volume 27, Issue 7
July 2020

Article Contents

JOURNAL ARTICLE

Learning from local to global: An efficient distributed algorithm for modeling time-to-event data

Get access

Rui Duan, Chongliang Luo, Martijn J Schuemie, Jiayi Tong, C Jason Liang, Howard H Chang, Mary Regina Boland, Jiang Bian, Hua Xu, John H Holmes, Christopher B Forrest, Sally C Morton, Jesse A Berlin, Jason H Moore, Kevin B Mahoney, Yong Chen ✉

Journal of the American Medical Informatics Association, Volume 27, Issue 7, July 2020, Pages 1028–1036, https://doi.org/10.1093/jamia/ocaa044

Published: 06 July 2020 **Article history** ▾

PDF

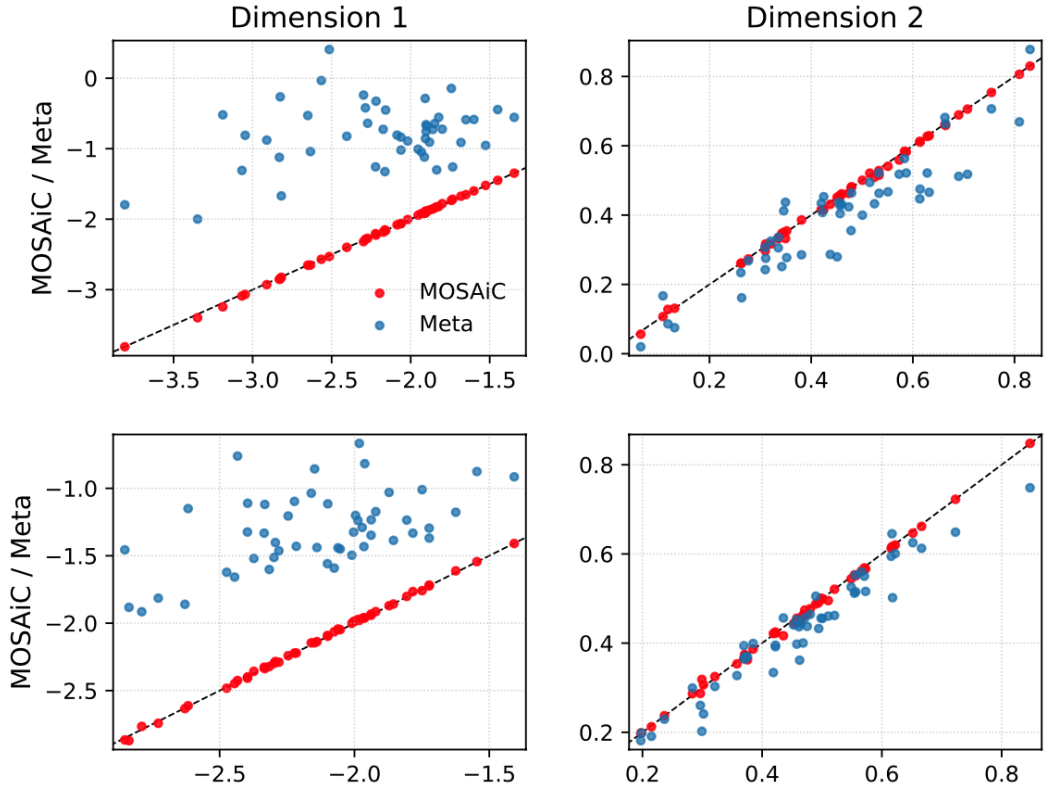
Article | Open access | Published: 22 April 2022

ODACH: a one-shot distributed algorithm for Cox model with heterogeneous multi-center data

Chongliang Luo, Rui Duan, Adam C. Naj, Henry R. Kranzler, Jiang Bian & Yong Chen ✉

Scientific Reports **12**, Article number: 6627 (2022) | Cite this article

Demo 2: MOSAiC provides a one-shot and lossless solution for stratified Cox models.



Approach	One-shot?	Lossless?	Independent of initial estimators?
WebDISCO	✗	✓	✓
ODAC/ODACH	✓	✗	✗
MOSAiC	✓	✓	✓

LATTE (lossless algorithms for target trial emulation) with time-to-event outcomes



Hua Xu

Yale SCHOOL OF MEDICINE



Cui Tao

MAYO CLINIC Mayo Clinic



Kai Zheng

UC Irvine



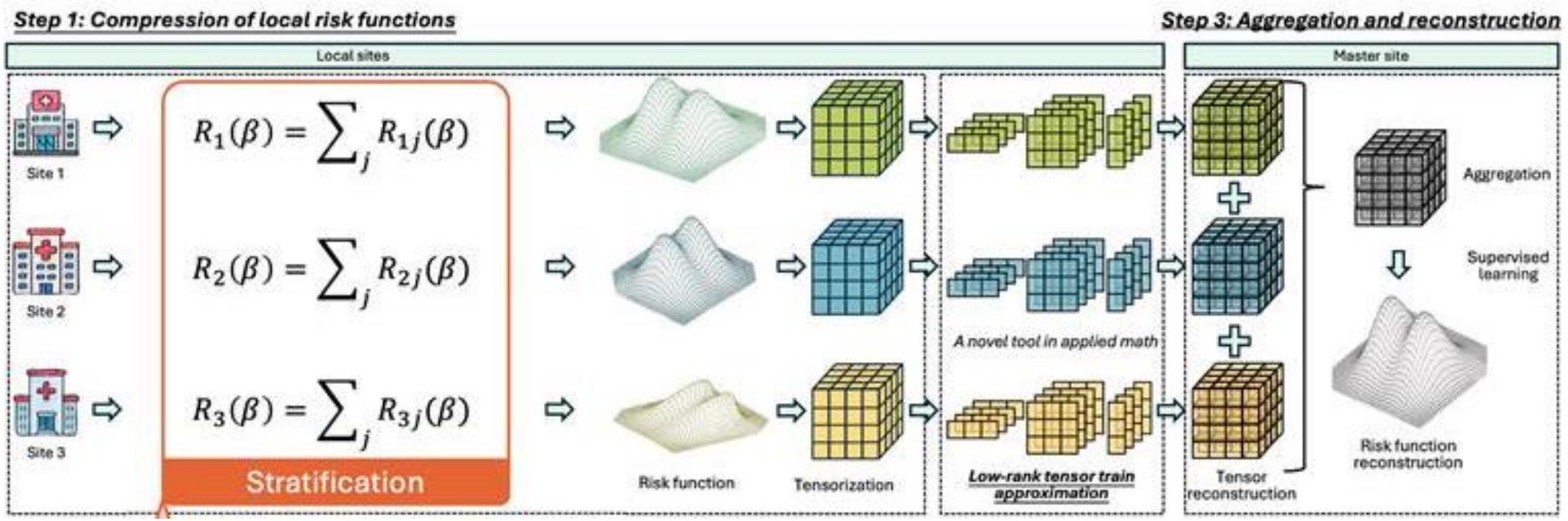
Fei Wang

Weill Cornell Medicine



Jiang Bian

INDIANA UNIVERSITY SCHOOL OF MEDICINE
Regenstrief Institute



Chen S, et al. MOTTO: a lossless tensor-train framework for multi-site time-to-event analysis of GLP-1 receptor agonists and psychiatric outcome.

Wang T, et al. Federated target trial emulation for time-to-event outcomes via POLARIS: Pooled-equivalent One-shot Likelihood Aggregation for Real-world Inference in Survival. Invited revision at npj Metabolic Health and Disease.

Lei Y, et al. MOCHA: A Multi-site One-shot Lossless Cox Hazard Analysis.

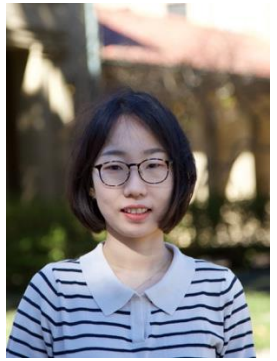
Lancet Digital Health

Alzheimer's disease drug repurposing using decentralised electronic health records: a one-shot lossless algorithm for federated target trial emulation (LATTE) and a multi-site observational study

Lu Li, Qiong Wu, Yiwen Lu, Kyra S O'Brien, Bingyu Zhang, Ting Zhou, Jiayi Tong, Dazheng Zhang, Yuqing Lei, Huilin Tang, Haoyang Li, Zhenxing Xu, Yu Huang, Yu Hu, Yujia Zhou, Fongci Lin, Na Hong, Ying Jiang, Fei Wang, Jiang Bian, Hua Xu, Yong Chen[1], Jeff D Williamson, David A Wolk, Yun Lu, David A Asch, Yong Chen[2]



Lu Li



Qiong Wu



Jiayi Tong



Yong Chen

eBioMedicine
Part of THE LANCET *Discovery Science*

This journal · Journals · Publish · Clinical · Global health · Multimedia · Events · About

ARTICLES · Volume 132, 106466, October 2026 · *Open Access* · Download Full Issue

Discovering repurposable drugs for Alzheimer's disease and related dementias: target trial emulation using decentralised real-world data

Qiong Wu [a,b,c] · Lu Li [b,d] · Yuqing Lei [b,e] · Ting Zhou [b,e] · Huilin Tang [b,e] · Bingyu Zhang [b,d] · Yiwen Lu [b,d] · Maxwell Salvatore [b,e,f,g] · Dazheng Zhang [b,d] · Jiayi Tong [h] · Tingyin Wang [b,d] · Siqi Chen [b,d] · Haoyang Li [i] · Zhenxing Xu [i] · Yu Huang [j] · Yu Hu [k] · Na Hong [l] · Yujia Zhou [l] · Fongci Lin [l] · Kyra S. O'Brien [m,n] · Yong Chen [o,p] · Ruowang Li [q] · Hua Xu [l] · Fei Wang [i] · Jiang Bian [j] · David A. Wolk [m] · Yong Chen [b,d,e,g,n,r] · Show less

NPJ Health System

FLAT-WHITE: From Target Trial Emulation to Federated High-Throughput Causal Machine Learning using Decentralized Real-World Data

Jiayi Tong[1], Lu Li[2,3], Dazheng Zhang[2,7], Kyra S.O'Brien[4,5], Yiwen Lu[2,3], Qiong Wu[6], Ting Zhou[2,7], Bingyu Zhang[2,3], Yuqing Lei[2,7], Huilin Tang[2,7], Naimin Jing[8], Haoyang Li[9], Zhenxing Xu[9], Yu Huang[10], Yu Hu[10], Yujia Zhou[11], Fongci Lin[11], Ying Jiang[12], Haitao Chu[13], Yong Chen[13], Fei Wang[9], Jiang Bian[10], Hua Xu[11], David A. Wolk[4], Yong Chen[2,3,7,14,15]

PDA

https://pdamethods.org/

Penn Medicine

Yale University School of Medicine

INSIGHT Clinical Research Network

OneFlorida+ Clinical Research Network

