Biomedical Informatics

discovery and impact

Dean's Distinguished Lecture:

Drawing reproducible conclusions from observational clinical data

George Hripcsak, MD, MS

Biomedical Informatics, Columbia University

Medical Informatics Services, NewYork-Presbyterian

COLUMBIA | COLUMBIA UNIVERSITY IRVING MEDICAL CENTER

NewYork-Presbyterian

Disclosures

- No financial disclosures

Team

PhD and Postdoc


Anna Ostropolets


Linying Zhang


Tara Anand


Lauren Richter


Ben Albert


Tiffany Callahan

OHDSI


Patrick Ryan


Karthik Natarajan


Soumitra Sengupta


Marc Suchard


Martijn Schuemie


David Madigan

Causal Inference


Elias Bareinboim


David Blei


David Albers

Observational research

- Subjects **observed** in their natural settings
  - Often using data collected for other purposes
  - **Real-world evidence**
- Versus experimental
  - Randomized clinical trials (RCTs)
- Administrative claims data
  - IBM Marketscan (10M's)
- Electronic health record (EHR) data
  - Columbia clinical data warehouse (6M)
- Other sources
  - Census, social media, mobile sensors, imaging

US National EHR data, per year

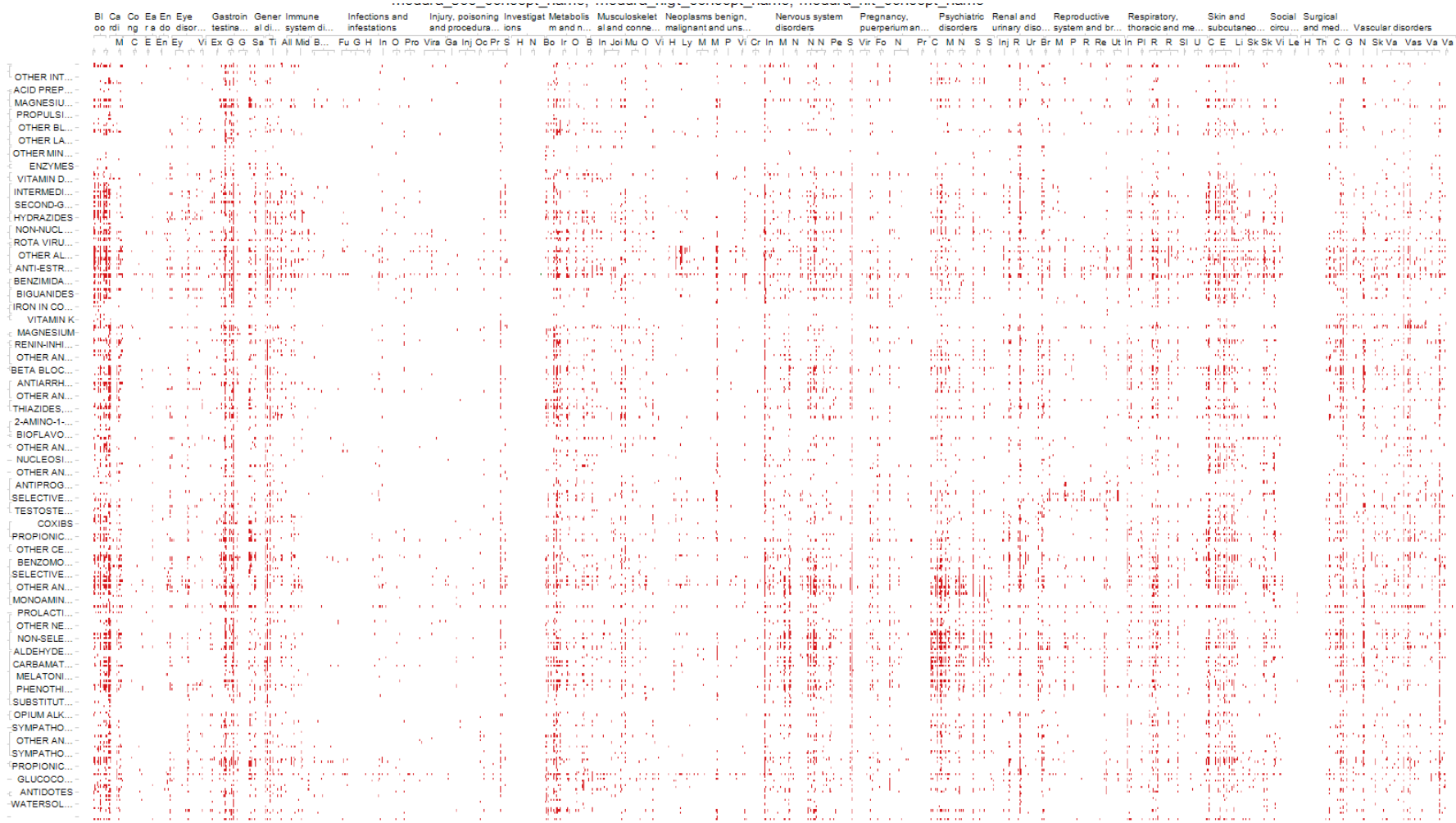
- Healthcare $4,000,000,000,000 industry in US
  - can't duplicate

1,000,000,000	visit notes
35,000,000	admit notes, discharge sum.
46,000,000	procedure notes
3,000,000,000	prescriptions
1,000,000,000	laboratory tests
>50,000,000,000	facts

Why large-scale analysis is needed in healthcare

All health outcomes of interest

All drugs



Patrick Ryan

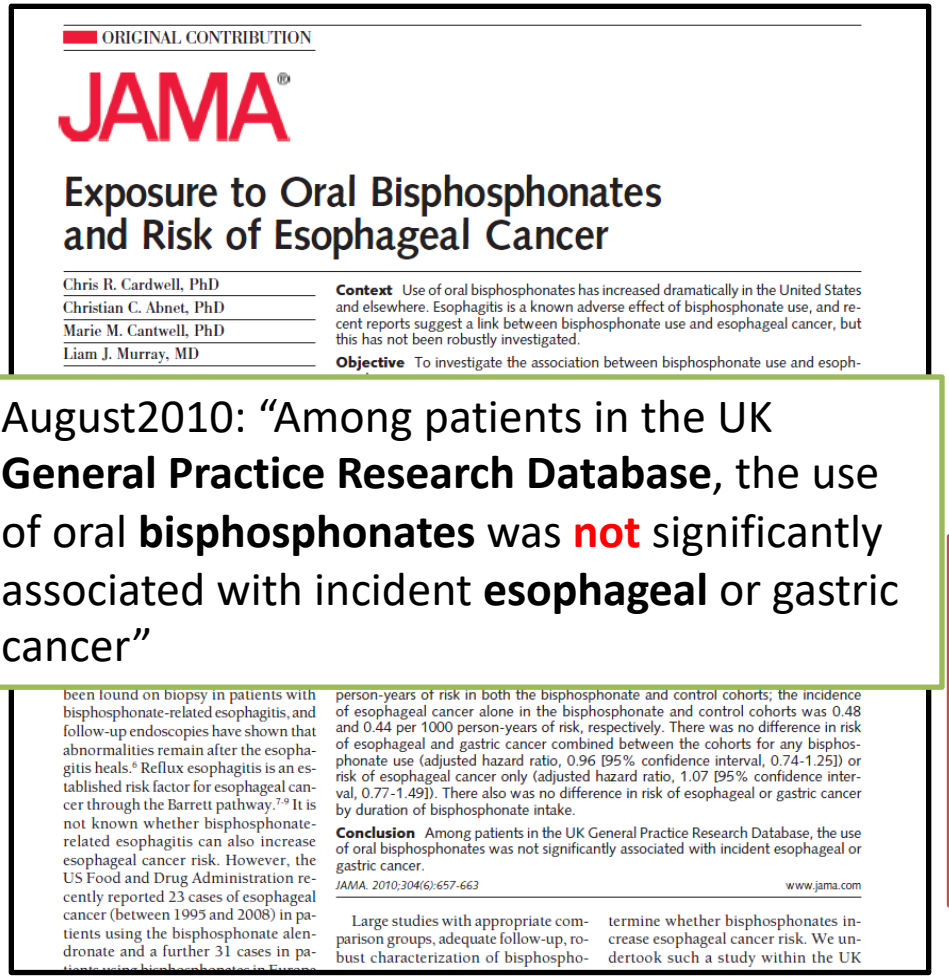
Patient-level predictions for personalized evidence requires big data

2 million patients seem excessive or unnecessary?

- Imagine a provider wants to compare her patient with other patients with the same gender (50%), in the same 10-year age group (10%), and with the same comorbidity of Type 2 diabetes (5%)
- Imagine the patient is concerned about the risk of ketoacidosis (0.5%) associated with two alternative treatments they are considering
- With 2 million patients, you'd only expect to observe 25 similar patients with the event, and would only be powered to observe a relative risk > 2.0

Aggregated data across a health system of 1,000 providers may contain 2,000,000 patients

But there is a catch



ORIGINAL CONTRIBUTION

JAMA®

Exposure to Oral Bisphosphonates and Risk of Esophageal Cancer

Chris R. Cardwell, PhD
Christian C. Abnet, PhD
Marie M. Cantwell, PhD
Liam J. Murray, MD

Context Use of oral bisphosphonates has increased dramatically in the United States and elsewhere. Esophagitis is a known adverse effect of bisphosphonate use, and recent reports suggest a link between bisphosphonate use and esophageal cancer, but this has not been robustly investigated.

Objective To investigate the association between bisphosphonate use and esoph-

been found on biopsy in patients with bisphosphonate-related esophagitis, and follow-up endoscopies have shown that abnormalities remain after the esophagitis heals.[6] Reflux esophagitis is an established risk factor for esophageal cancer through the Barrett pathway.[7-9] It is not known whether bisphosphonate-related esophagitis can also increase esophageal cancer risk. However, the US Food and Drug Administration recently reported 23 cases of esophageal cancer (between 1995 and 2008) in patients using the bisphosphonate alendronate and a further 31 cases in pa-

person-years of risk in both the bisphosphonate and control cohorts; the incidence of esophageal cancer alone in the bisphosphonate and control cohorts was 0.48 and 0.44 per 1000 person-years of risk, respectively. There was no difference in risk of esophageal and gastric cancer combined between the cohorts for any bisphosphonate use (adjusted hazard ratio, 0.96 [95% confidence interval, 0.74-1.25]) or risk of esophageal cancer only (adjusted hazard ratio, 1.07 [95% confidence interval, 0.77-1.49]). There also was no difference in risk of esophageal or gastric cancer by duration of bisphosphonate intake.

Conclusion Among patients in the UK General Practice Research Database, the use of oral bisphosphonates was not significantly associated with incident esophageal or gastric cancer.

JAMA. 2010;304(6):657-663 www.jama.com

Large studies with appropriate comparison groups, adequate follow-up, robust characterization of bisphospho- termine whether bisphosphonates increase esophageal cancer risk. We undertook such a study within the UK

August2010: "Among patients in the UK **General Practice Research Database**, the use of oral **bisphosphonates** was **not** significantly associated with incident **esophageal** or gastric cancer"

BMJ

RESEARCH

Oral bisphosphonates and risk of cancer of oesophagus, stomach, and colorectum: case-control analysis within a UK primary care cohort

Jane Green, clinical epidemiologist,[1] Gabriela Czanner, statistician,[1] Gillian Reeves, statistical epidemiologist,[1] Joanna Watson, epidemiologist,[1] Lesley Wise, manager, Pharmacoepidemiology Research and Intelligence Unit,[2] Valerie Beral, professor of cancer epidemiology[1]

[1]Cancer Epidemiology Unit, University of Oxford, Oxford OX3 7LF
[2]Medicines and Healthcare products Regulatory Agency, Pharmacoepidemiology Research Unit, London SW8 5NQ
Correspondence to: J Green jane.green@ceu.ox.ac.uk

Cite this as: BMJ 2010;341:c4444 doi:10.1136/bmj.c4444

ABSTRACT

Objective To examine the hypothesis that risk of oesophageal, but not of gastric or colorectal, cancer is increased in users of oral bisphosphonates.

Design Nested case-control analysis within a primary care cohort of about 6 million people in the UK, with prospectively recorded information on prescribing of bisphosphonates.

Setting UK General Practice Research Database cohort.

Participants Men and women aged 40 years or over—

Conclusions The risk of oesophageal cancer increased with 10 or more prescriptions for oral bisphosphonates and with prescriptions over about a five year period. In Europe and North America, the incidence of oesophageal cancer at age 60-79 is typically 1 per 1000 population over five years, and this is estimated to increase to about 2 per 1000 with five years' use of oral bisphosphonates.

INTRODUCTION

Adverse gastrointestinal effects are common among

corticosteroids. Cancers of the stomach and colorectum were not associated with prescription of bisphosphonate: relative risks for one or more versus no prescriptions were 0.87 (0.66 to 1.19) and 0.87 (0.77 to 1.00). The specificity

style data. General Practice Research Database prescription data have been shown to be virtually complete, and the data on incidence of cancer (based on hospital records) are around 95% valid and

Sept2010: "In this large nested case-control study within a UK cohort [**General Practice Research Database**], we found a significantly increased risk of **oesophageal** cancer in people with previous prescriptions for oral **bisphosphonates**"

Retracted COVID-19 papers

EDITORIAL

Expression of Concern: Mehra MR et al. Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. N Engl J Med. DOI: 10.1056/NEJMoa2007621.

Eric J. Rubin, M.D., Ph.D.

Article

3 References 12 Citing Articles

On May 1, 2020, we published "Cardiovascular Dis and Mortality in Covid-19,"[1] a study of the effect of preexistir angiotensin-converting enzyme (ACE) inhibitors and angiote (ARBs) on Covid-19. This retrospective study used data drawn from an i

Metrics June 18, 2020

THE LANCET
Volume 395, Issue 10240, 13–19 June 2020, Page e102

Comment

Expression of concern: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis

The Lancet Editors

Show more

Share Cite

https://doi.org/10.1016/S0140-6736(20)31290-3 Get rights and content

13 February 2020…

The NEW ENGLAND JOURNAL of MEDICINE

SOUNDING BOARD

The Magic of Randomization versus the Myth of Real-World Evidence

Rory Collins, F.R.S., Louise Bowman, M.D., F.R.C.P., Martin Landray, Ph.D., F.R.C.P., and Richard Peto, F.R.S.

Nonrandomized observational analyses of large electronic patient databases are being promoted as an alternative to randomized clinical trials as a source of "real-world evidence" about the efficacy and safety of new and existing treatments.[1-3] For drugs or procedures that are already being used widely, such observational studies may involve exposure of large numbers of patients. Consequently, they have the potential to detect rare adverse effects that cannot plausibly be attributed to bias, generally because the relative risk is large (e.g., Reye's syndrome associated with the use of aspirin, or rhabdomyolysis associated with the use of statin therapy).[4] Nonrandomized clinical observation may also suffice to detect large beneficial effects when good outcomes would not otherwise be expected (e.g., control of diabetic ketoacidosis with insulin treatment, or the rapid shrinking of tumors with chemotherapy).

However, because of the potential biases inherent in observational studies, such studies cannot generally be trusted when — as is often the case — the effects of the treatment of interest are actually null or only moderate (i.e., less than a twofold difference in the incidence of the health outcome between using and not using the treatment).[4-6] In those circumstances, large observational studies may yield misleading associations of a treatment with health outcomes that are statistically significant but noncausal, or that are mistakenly null when the treatment really does have clinically important effects. Instead, randomized, controlled trials of adequate size are generally required to ensure that any moderate benefits or moderate harms of a treatment are assessed reliably enough to guide patient care appropriately (Box 1).[5-7]

False comparison

- NOT observational versus RCT
  - Don’t have the money to do the RCTs we need
  - Point-of-care randomization will help
- Observational versus expert opinion and instinct
  - Guidelines are mostly expert opinion
- We need to optimize our observational research

Desired attributes for reliable evidence

Desired attribute	Question	Researcher	Data	Analysis		Result
Repeatable	Identical	Identical	Identical	Identical	=	Identical
Reproducible	Identical	Different	Identical	Identical	=	Identical
Replicable	Identical	Same or different	Similar	Identical	=	Similar
Generalizable	Identical	Same or different	Different	Identical	=	Similar
Robust	Identical	Same or different	Same or different	Different	=	Similar
Calibrated	Similar (controls)	Identical	Identical	Identical	=	Statistically consistent

Patrick Ryan

Reliable evidence requires a new tool: **the community**

Desired attribute	Question	Researcher	Data	Analysis		Result
R	ntical	Identical	Identical	Identical	=	Identical
	ntical	Different	Identical		=	Identical
R	Identical	Same or	Similar		=	Similar
Generalizable	Iden		Different	Identical	=	Similar
Robust	Identical	Same or differ	Same or	Different	=	Similar
Calibrated	Similar (controls)	Iden		Identical		

Community of researchers with important public health questions

Data network using community standards

Analyses sharing community open-source tools

Application of community best practices for evaluation

Evidence sharing across community

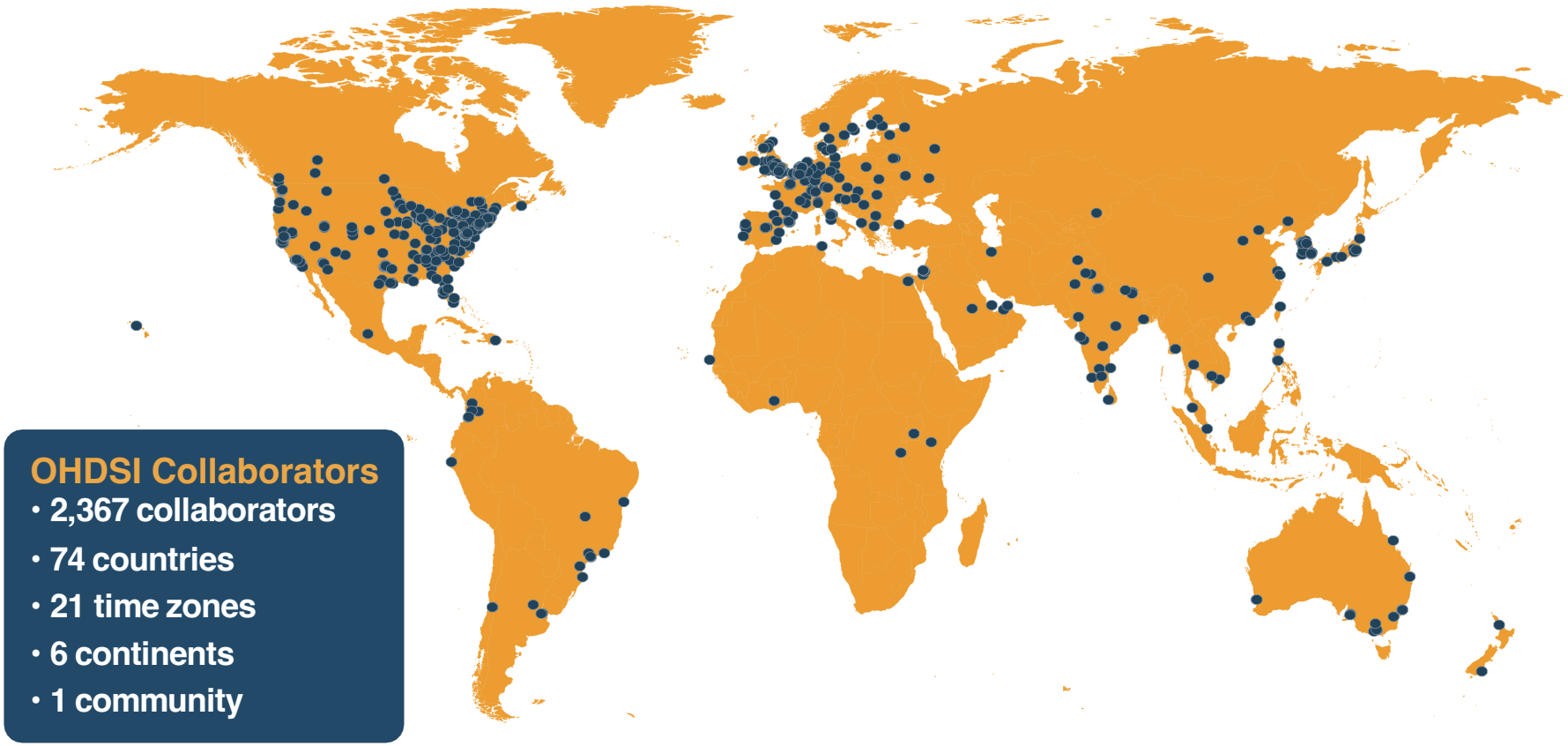
Observational Health Data Sciences and Informatics (OHDSI, as “Odyssey”)

Mission: To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care

A multi-stakeholder, interdisciplinary, international collaborative with a coordinating center at Columbia University

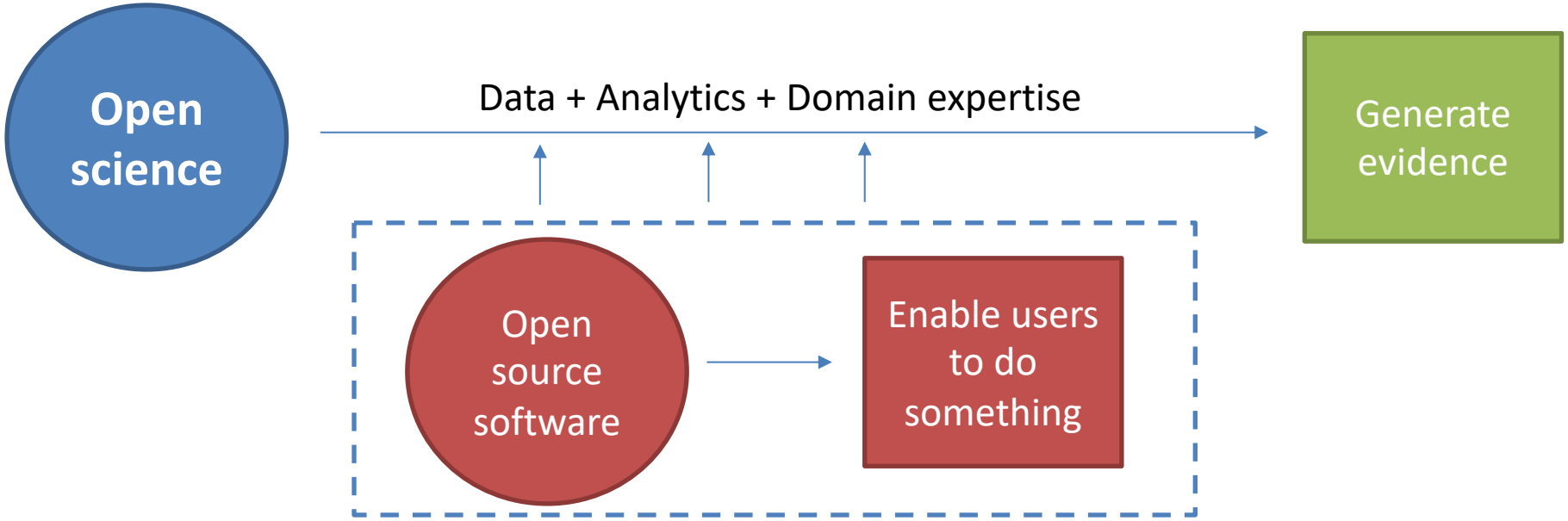
http://ohdsi.org

OHDSI

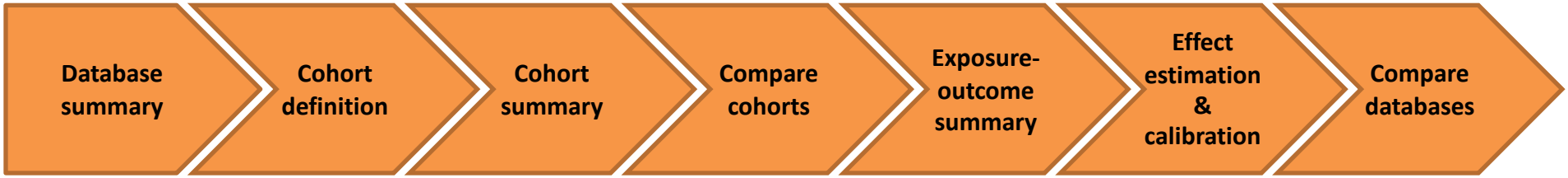


- Experts in informatics, statistics, epidemiology, clinical sciences
- Active participation from academia, government, industry, providers
- Currently records on about **800 million unique patients** in >300 databases
- 344 papers, specific influence on EMA and FDA for COVID-19

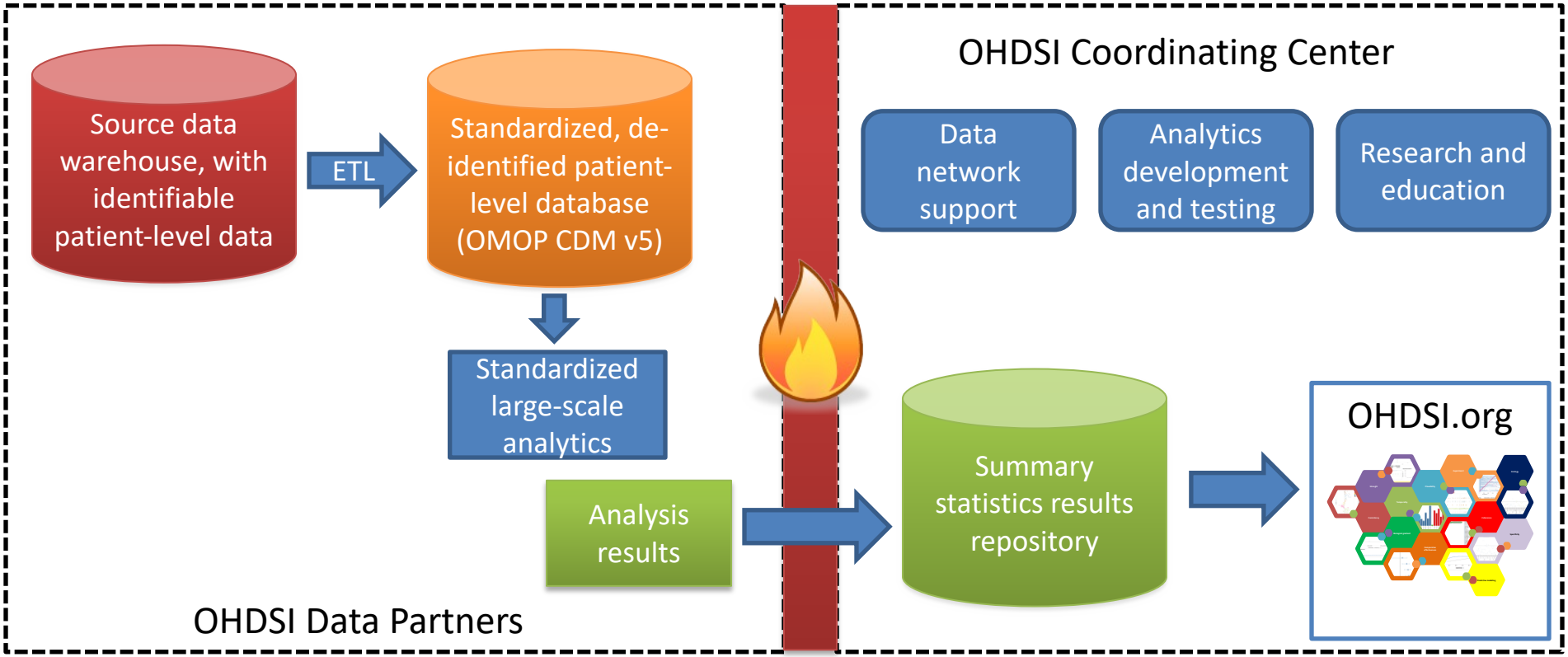
Open Science



Standardized, transparent workflows

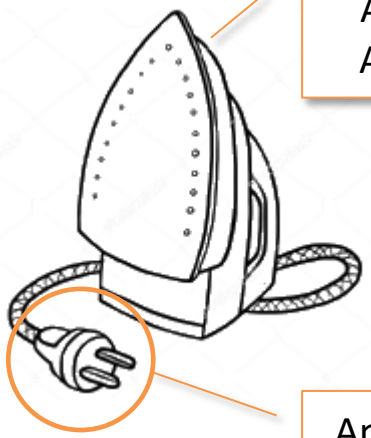


How OHDSI Works



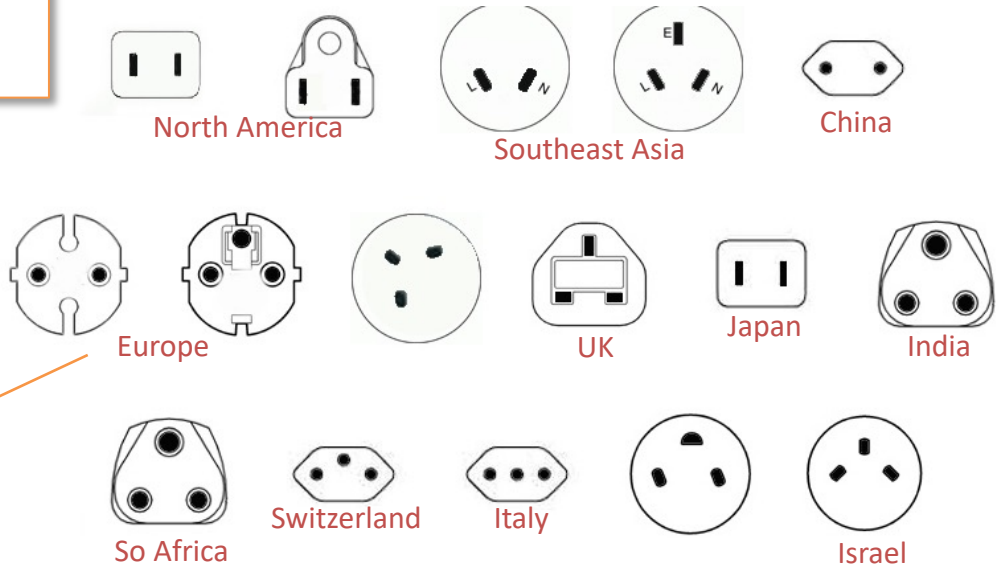
Current Approach: “One Study – One Script“

"What's the adherence to my drug in the data assets I own?"

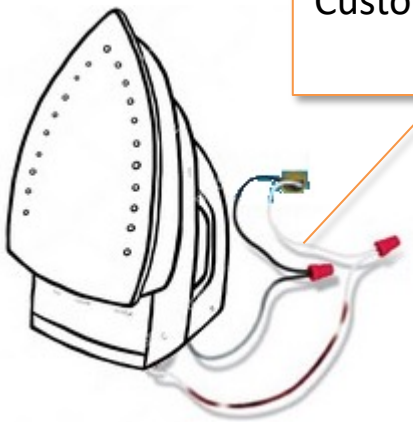


Analytical method: Adherence to Drug

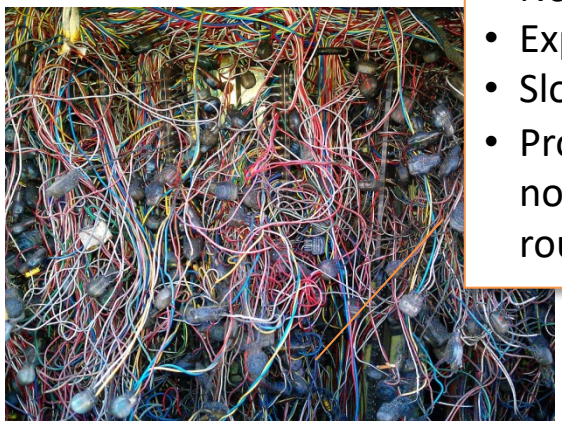
Application to data



Current solution:



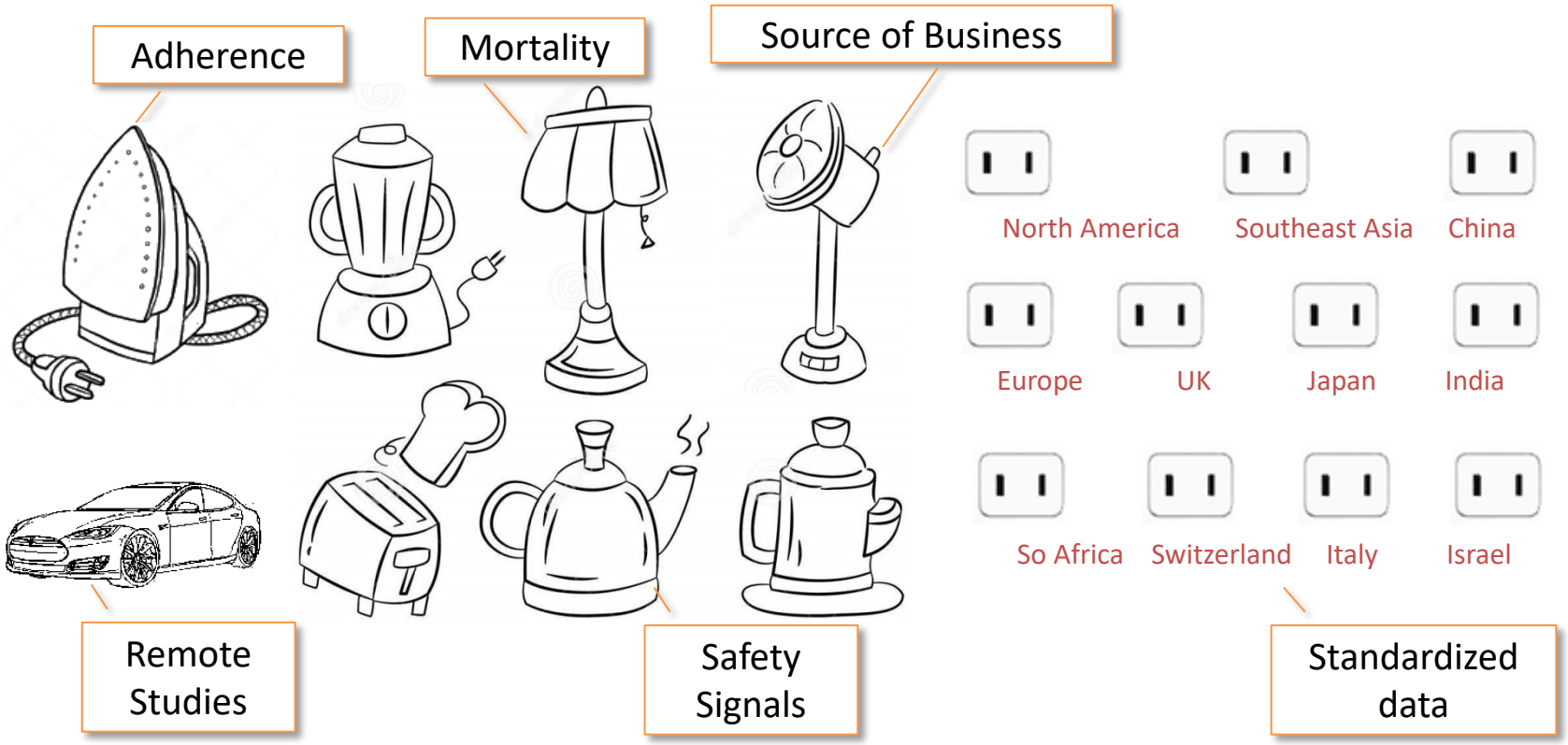
Custom script for each study



- Not scalable
- Expensive
- Slow
- Prohibitive to non-expert routine use

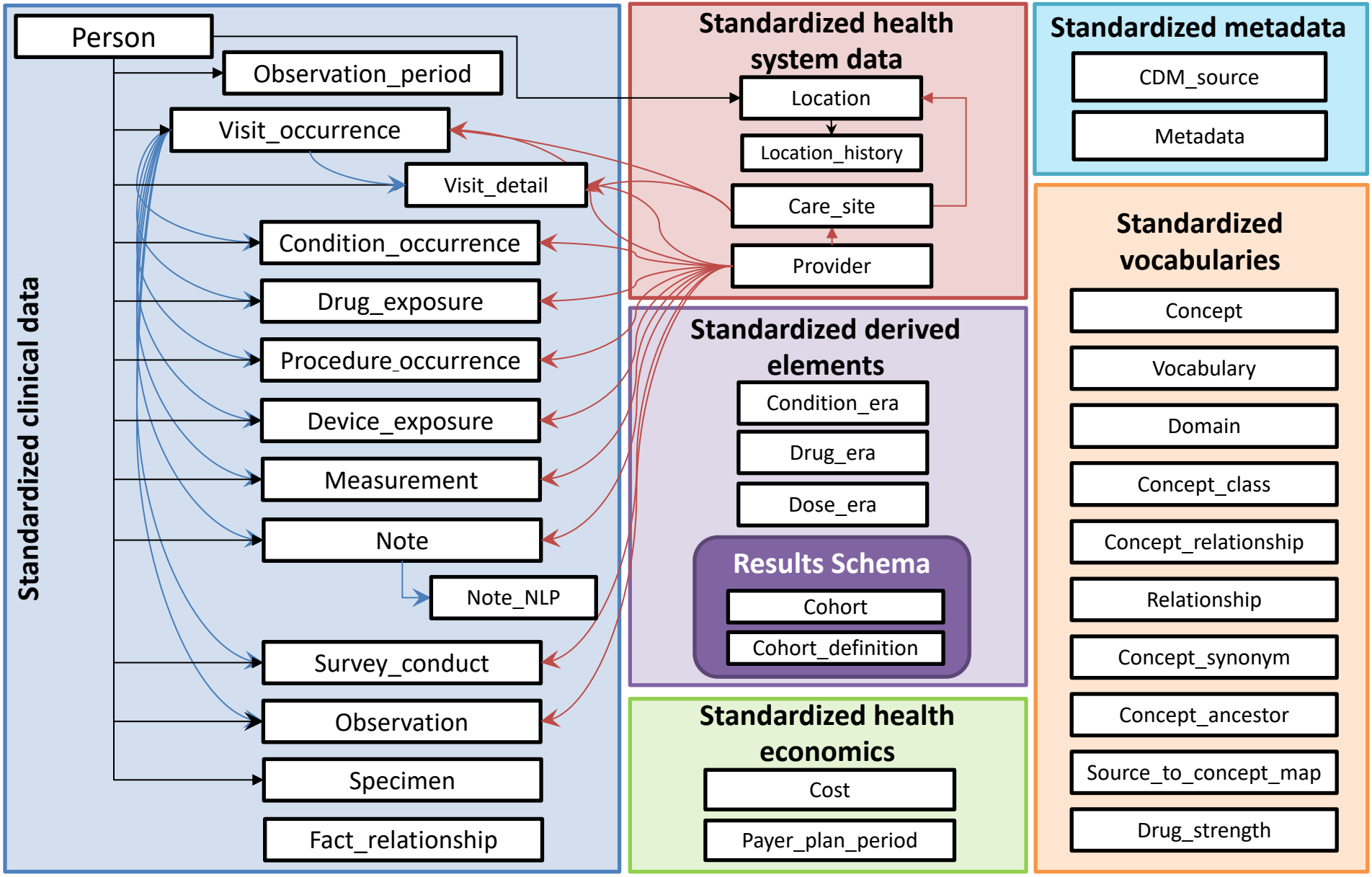
Christian Reich

Solution: Standardized Data and Analytics

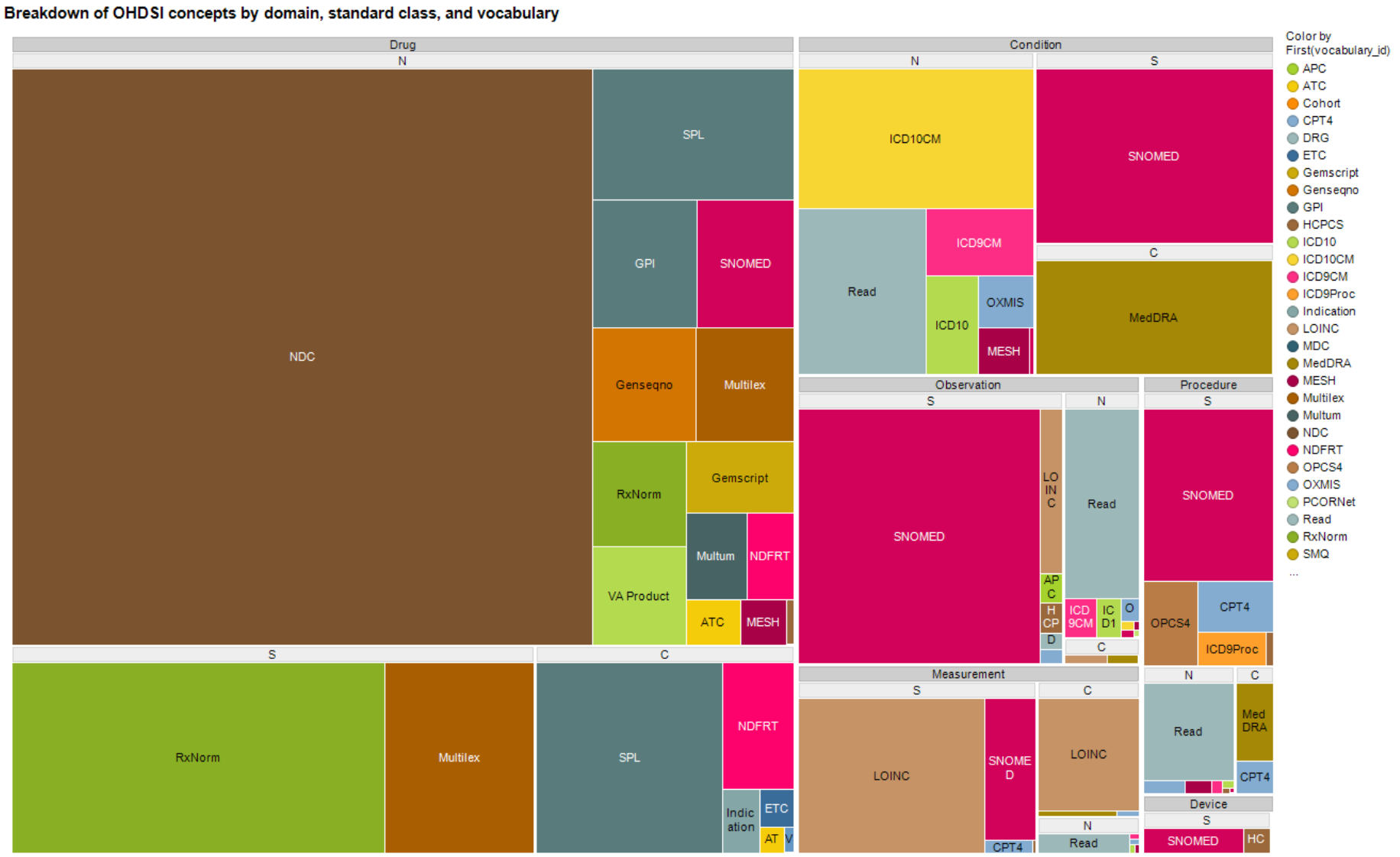


1. ATLAS, Remote Studies
   - Standard Cohorts
   - Standardized Analytics
2. OMOP CDM
   - Standardized Format
   - Standardized Coding

Deep information model
OMOP Common Data Model



Extensive vocabularies

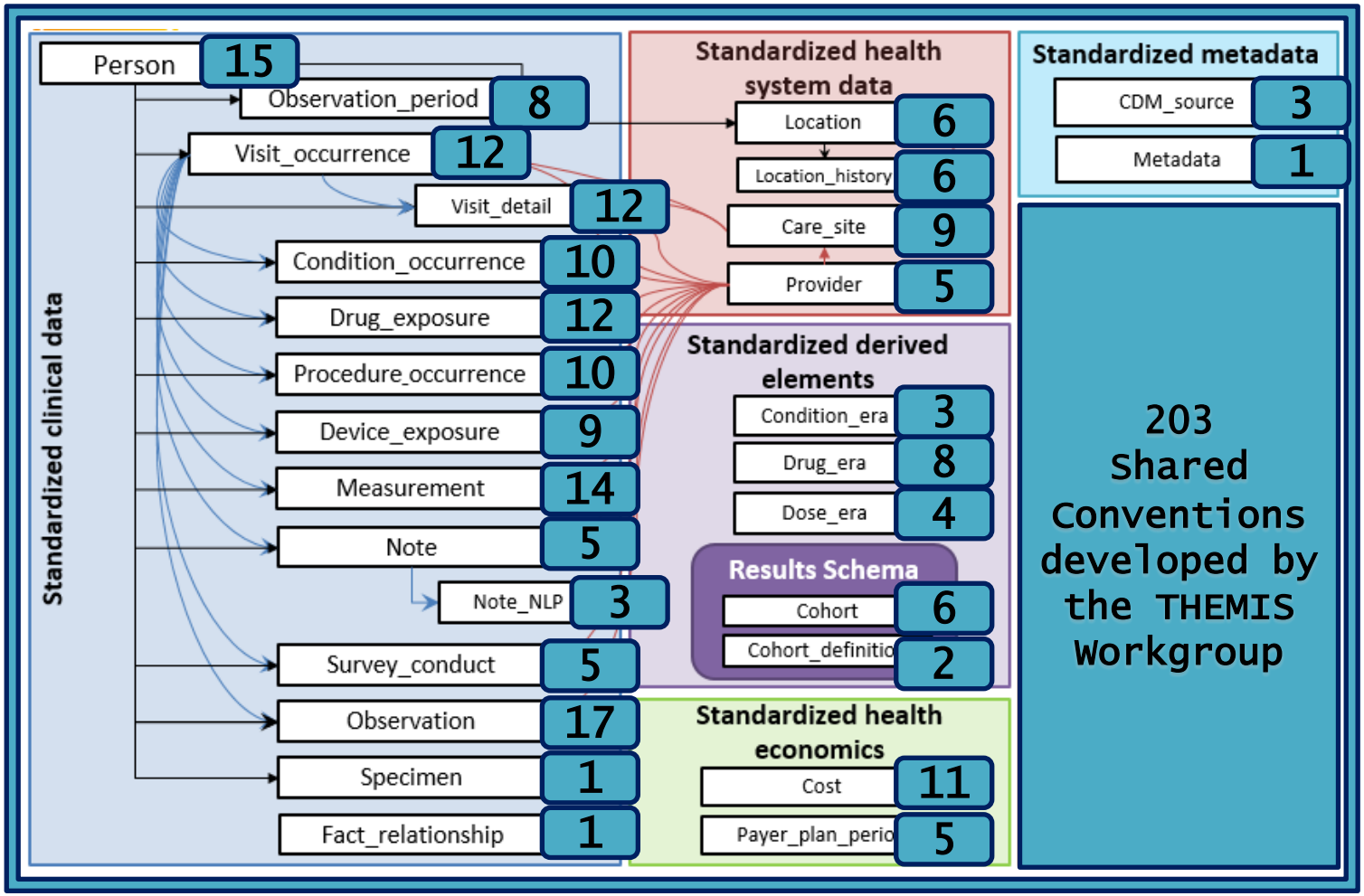


OHDSI's standardized vocabularies

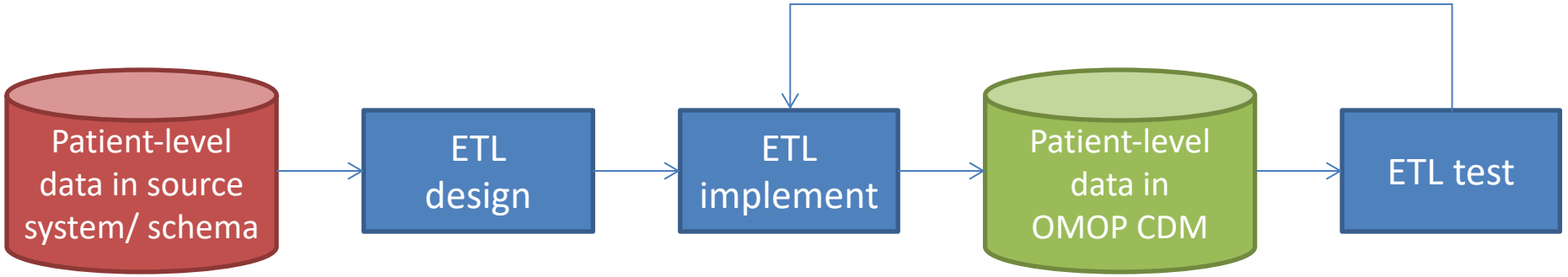
- 153 Vocabularies across 41 domains
  - MU3 standards: SNOMED, RxNorm, LOINC
  - Disparate sources: ICD9CM, ICD10(CM), Read, NDC, Gemscript, CPT4, HCPCS…
- **>9 million concepts**
  - >3.3 million standard concepts
  - >5.1 million source codes
  - >629,000 classification concepts
- >55 million concept relationships
- >84 million ancestral relationships

Publicly available for download at: http://athena.ohdsi.org/

Standardized conventions



Preparing your data



OHDSI tools built to help

WhiteRabbit: profile your source data

RabbitInAHat: map your source structure to CDM tables and fields

ATHENA: standardized vocabularies for all CDM domains

Usagi: map your source codes to CDM vocabulary

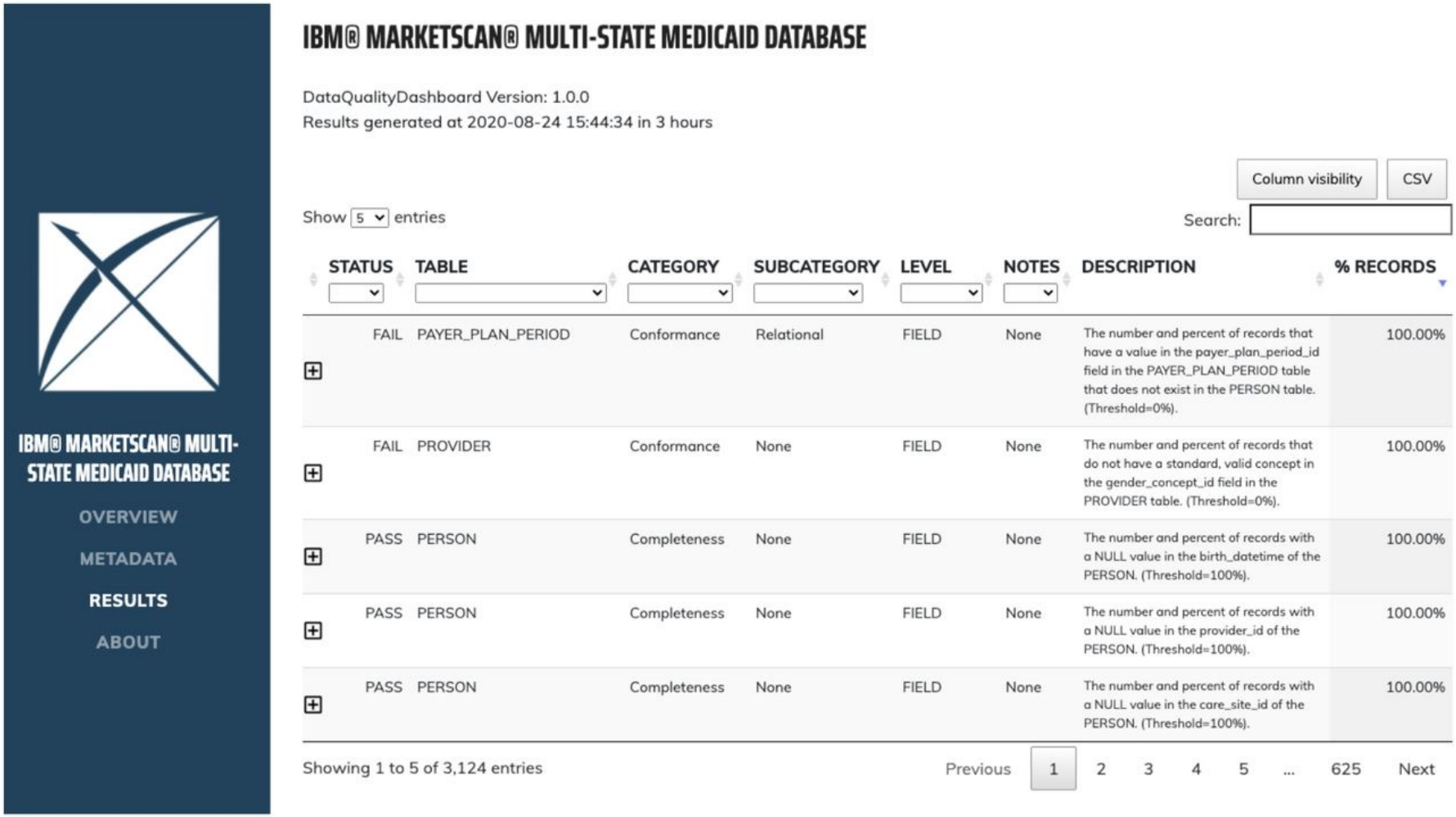
CDM: DDL, index, constraints for Oracle, SQL Server, PostgresQL; Vocabulary tables with loading scripts

ACHILLES: profile your CDM data; review data quality assessment; explore population-level summaries

OHDSI Forums:
Public discussions for OMOP CDM Implementers/developers

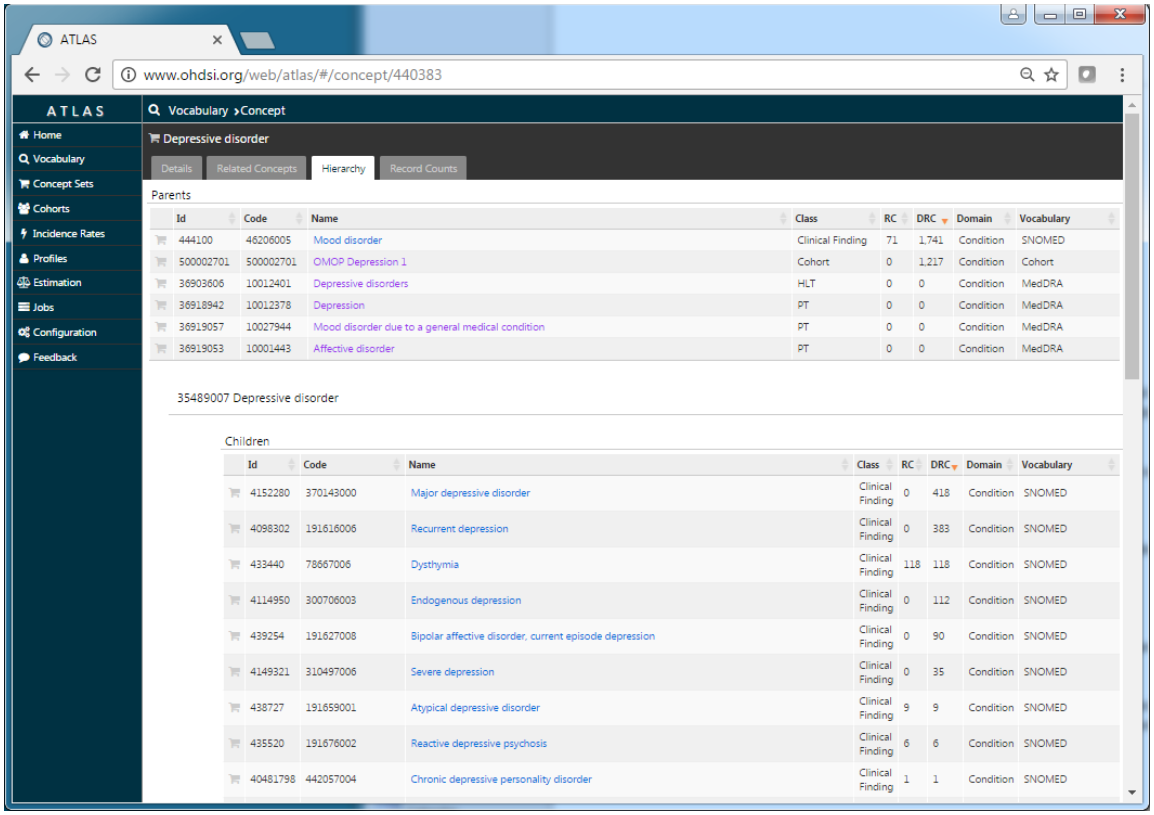
http://github.com/OHDSI

Data Quality Dashboard



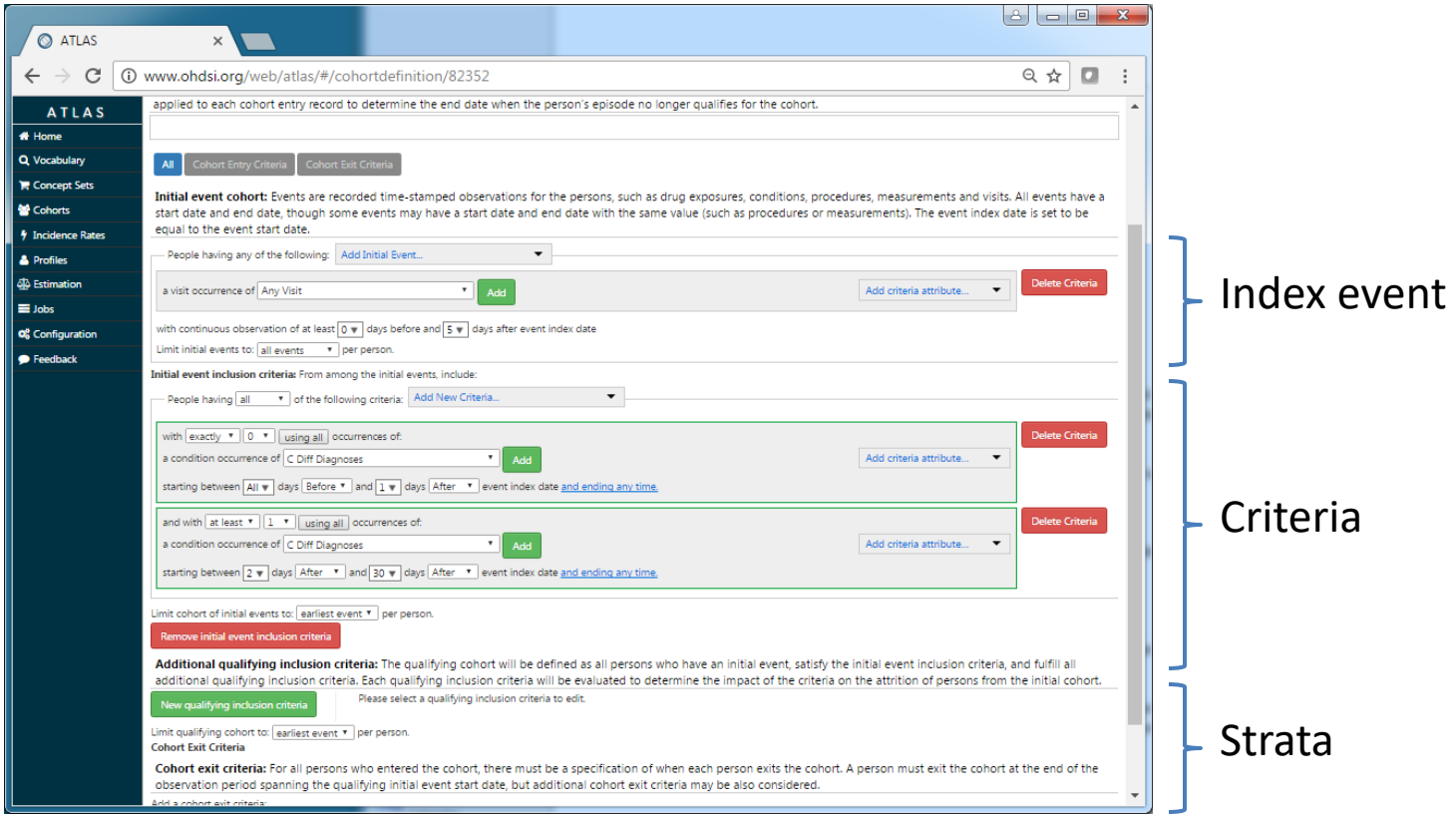
ATLAS: Ontology support

- What terms do I need to create a cohort
- Tied to the database: what terms are used
  - Especially important for someone *else's* database



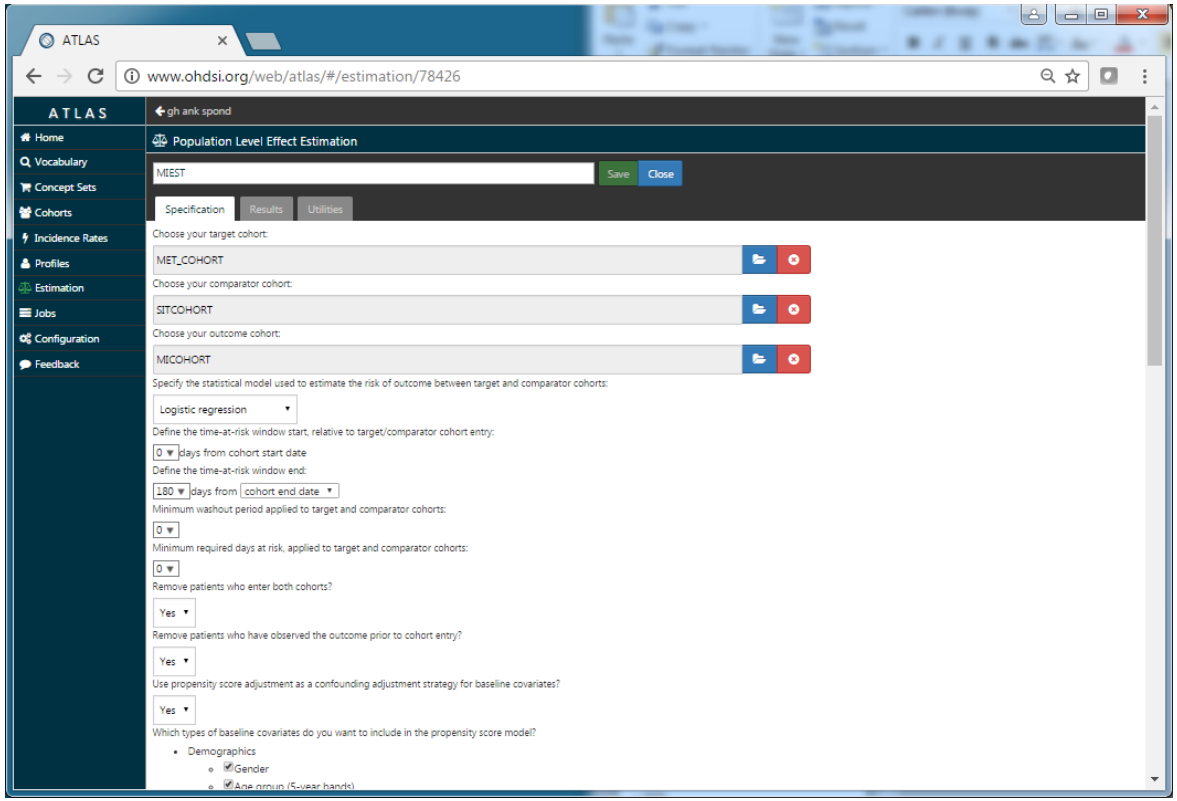
ATLAS: Cohort building

- Optimized for observational research
  - Time series: who and *when* (vs classification)
    - Observation period, event timing
  - Assume a complex definition – Linearized AND-OR group



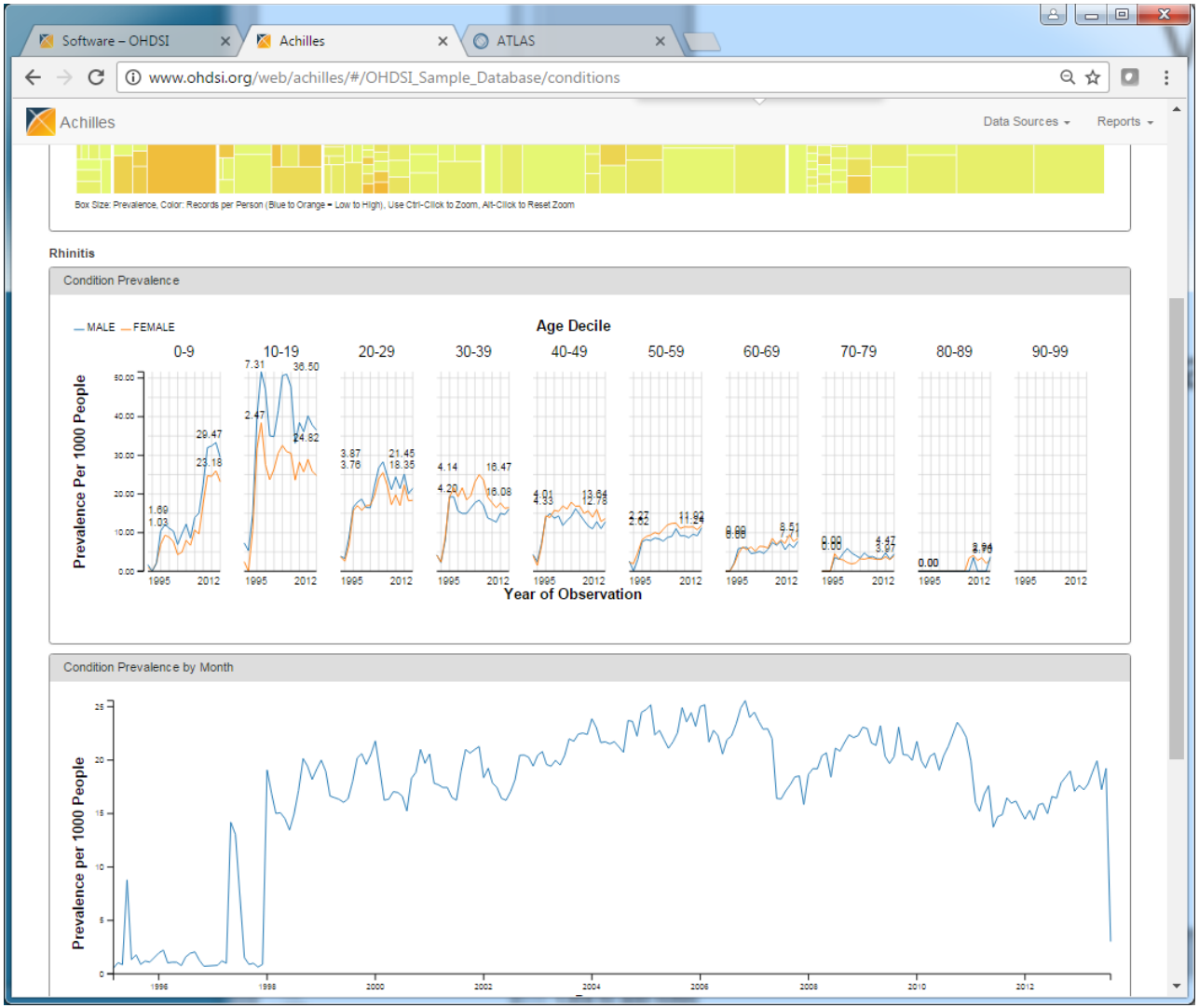
ATLAS: Analysis (observational)

- Approach: log regression, Poisson regression, survival
- Confounder: regularized-regression propensity score
- Residual confounding: calibration
- Diagnostics



ATLAS: Visualization

- Tables
- Graphs



Evidence OHDSI seeks to generate from observational data

- **Clinical characterization - tally**
  - Natural history: Who has diabetes, and who takes metformin?
  - Quality improvement: What proportion of patients with diabetes experience complications?
- **Population-level estimation - cause**
  - Safety surveillance: Does metformin cause lactic acidosis?
  - Comparative effectiveness: Does metformin cause lactic acidosis more than glyburide?
- **Patient-level prediction - predict**
  - Precision medicine: Given everything you know about me, if I take metformin, what is the chance I will get lactic acidosis?
  - Disease interception: Given everything you know about me, what is the chance I will develop diabetes?

Phenotyping

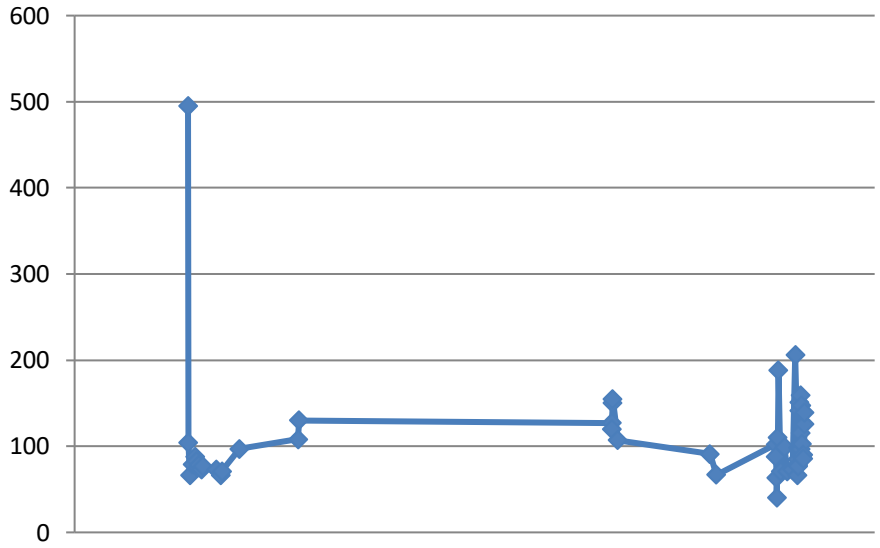
Meaning

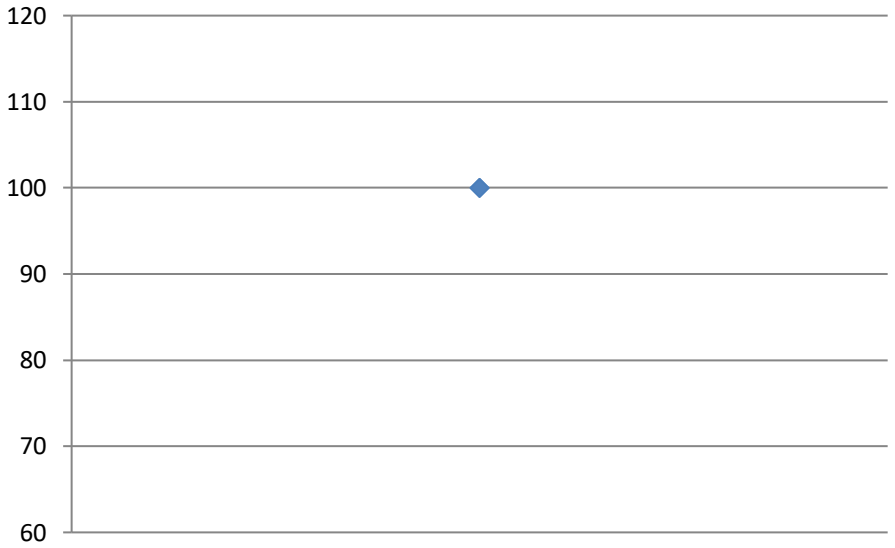
- PERRLA

Pupils equal, round, reactive to light and accommodation

Missing

- Data are mostly missing
  - Sampled when sick
- Implicit information
  - Pertinent negatives by attending vs CC3





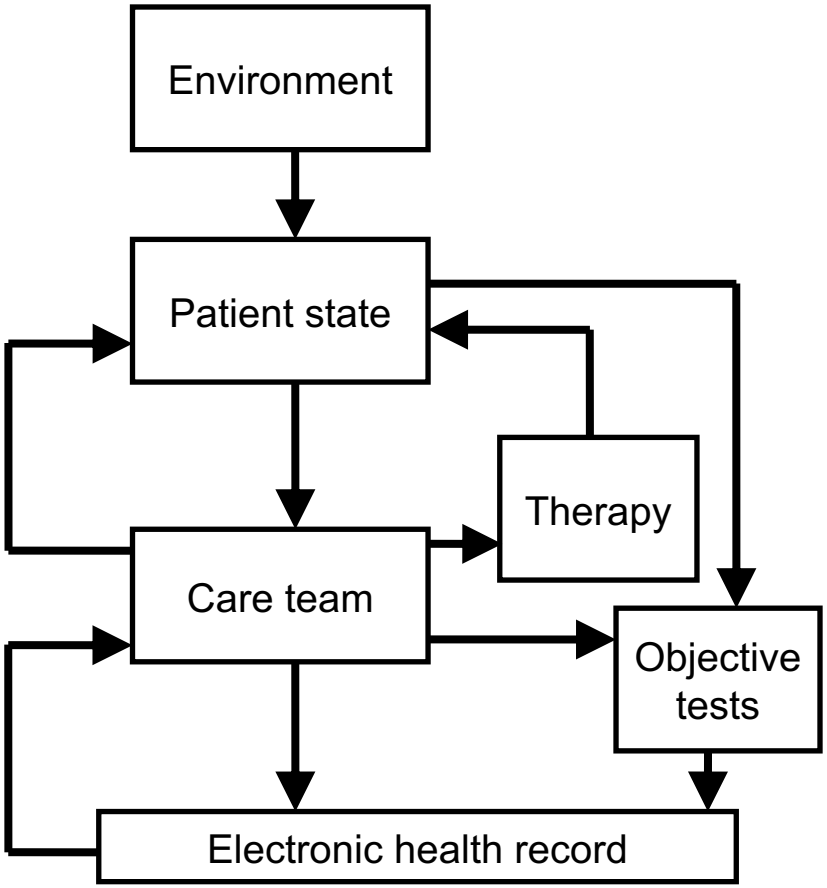
Noisy

- As low as 50% accuracy (Hogan JAMIA 1997)
- … 36 year old man … 27 year old woman …

Complex

- Which is the right time?
  - When specimen drawn
  - When specimen received
  - When test performed
  - When result updated
  - When result received by the patient
  - When patient told clinician
  - When clinician wrote the note

Health care process bias



Hripcsak JAMIA 2013

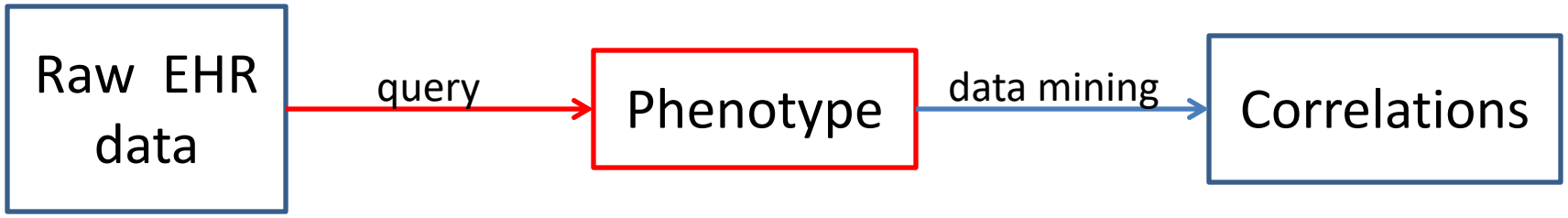
Good news

- Doctors successfully infer patient state from records

- We need to mimic the doctor's reasoning
  - Deconvolve the truth

EHR-derived **phenotype**

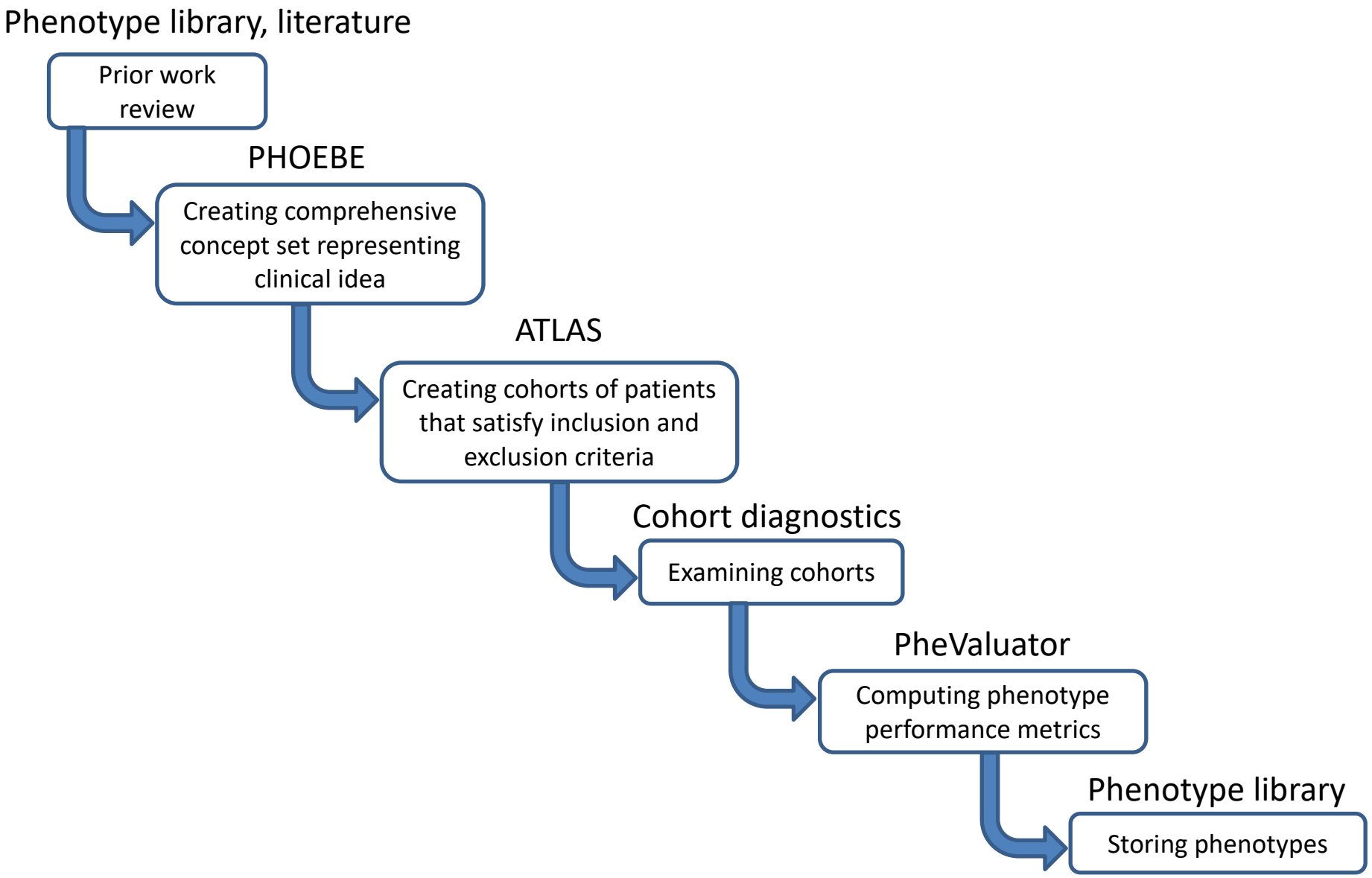
- Clinically relevant feature derived from EHR
  - Patient has (a diagnosis of) type II diabetes
  - Recent rash and fever
  - Drug-induced liver injury
- Then use phenotype in correlation studies, …
  - Which treatments associated with best outcomes



EHR-derived **phenotype**

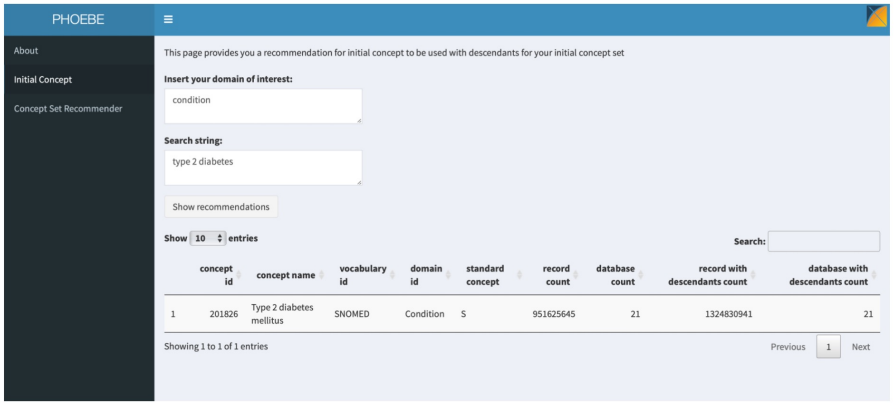
- Want to know if patient has type 2 diabetes
  - Don’t just look up the disease in the record
  - Yes, look for diagnosis codes
  - Diabetes medications
  - Glucose suggestive of diabetes
  - Special diabetes tests
  - Diabetes complications
  - Mentions in notes
  - Exclusions like type 1 diabetes
- Can take months to define and test

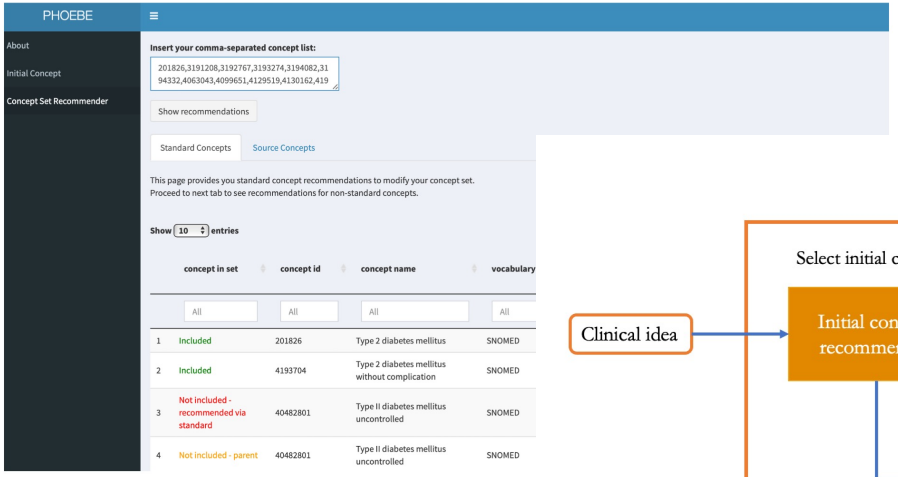
OHDSI phenotyping pipeline

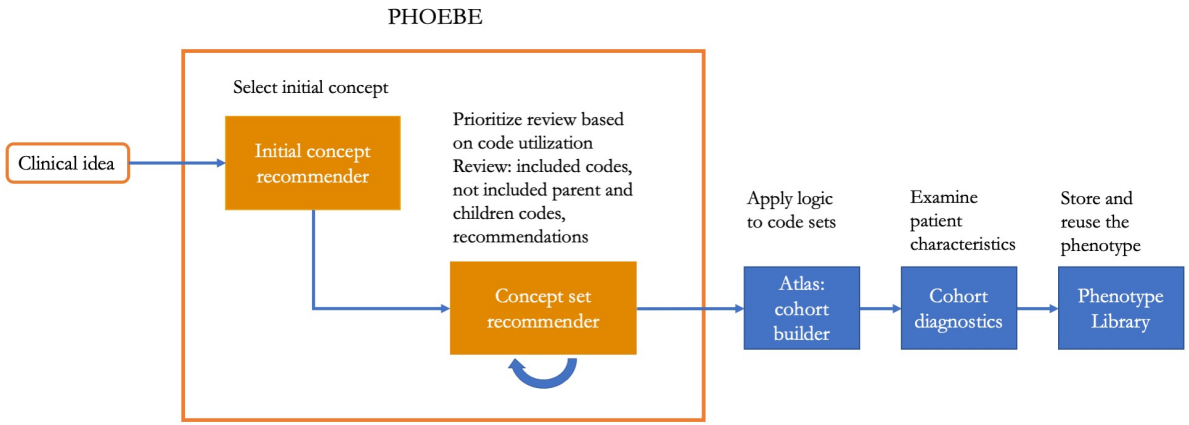


PHOEBE: get the right concepts

Exploit vast data network: rate of every code everywhere



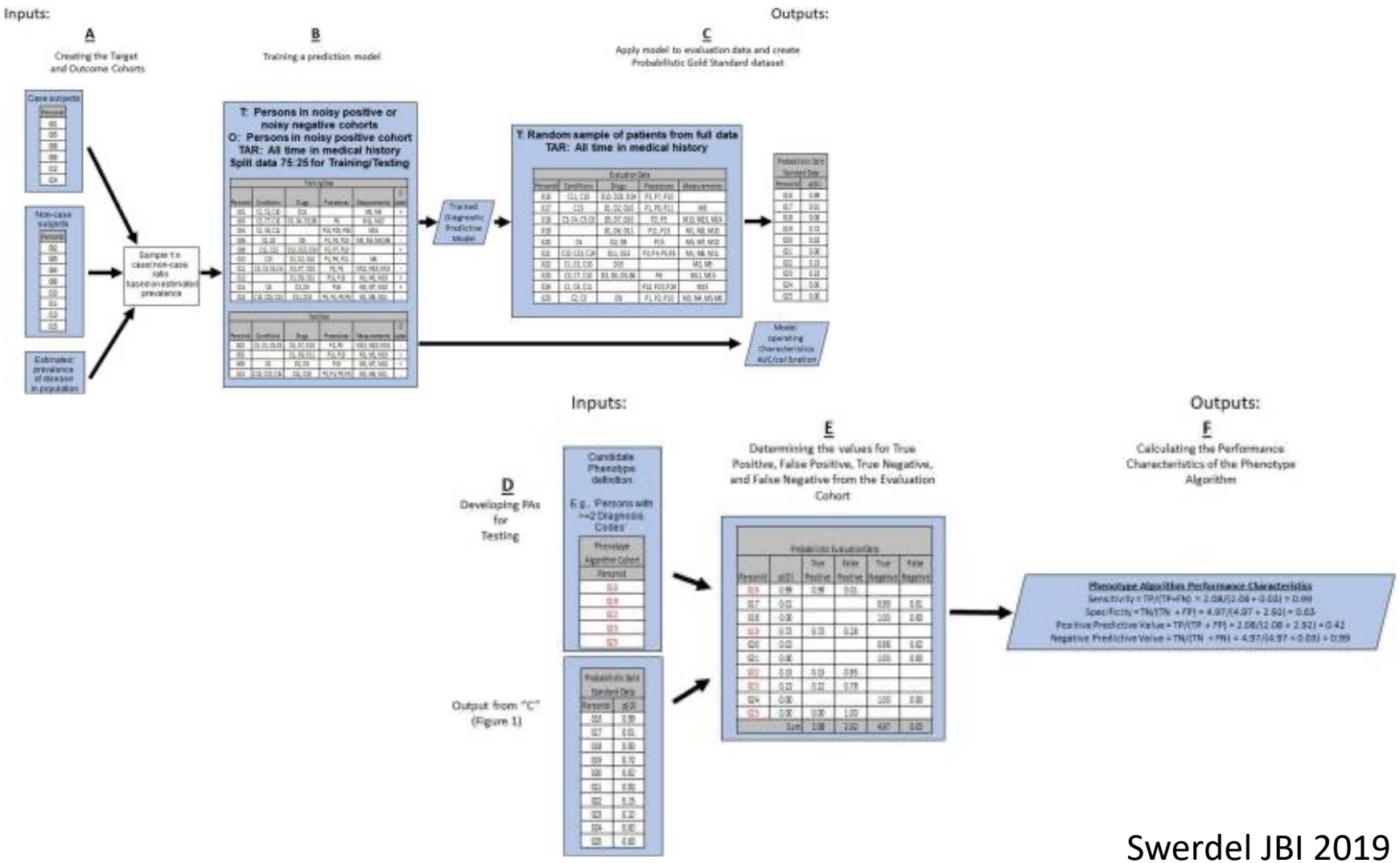




Ostropolets AMIA 2021

Phevaluator: automate the evaluation

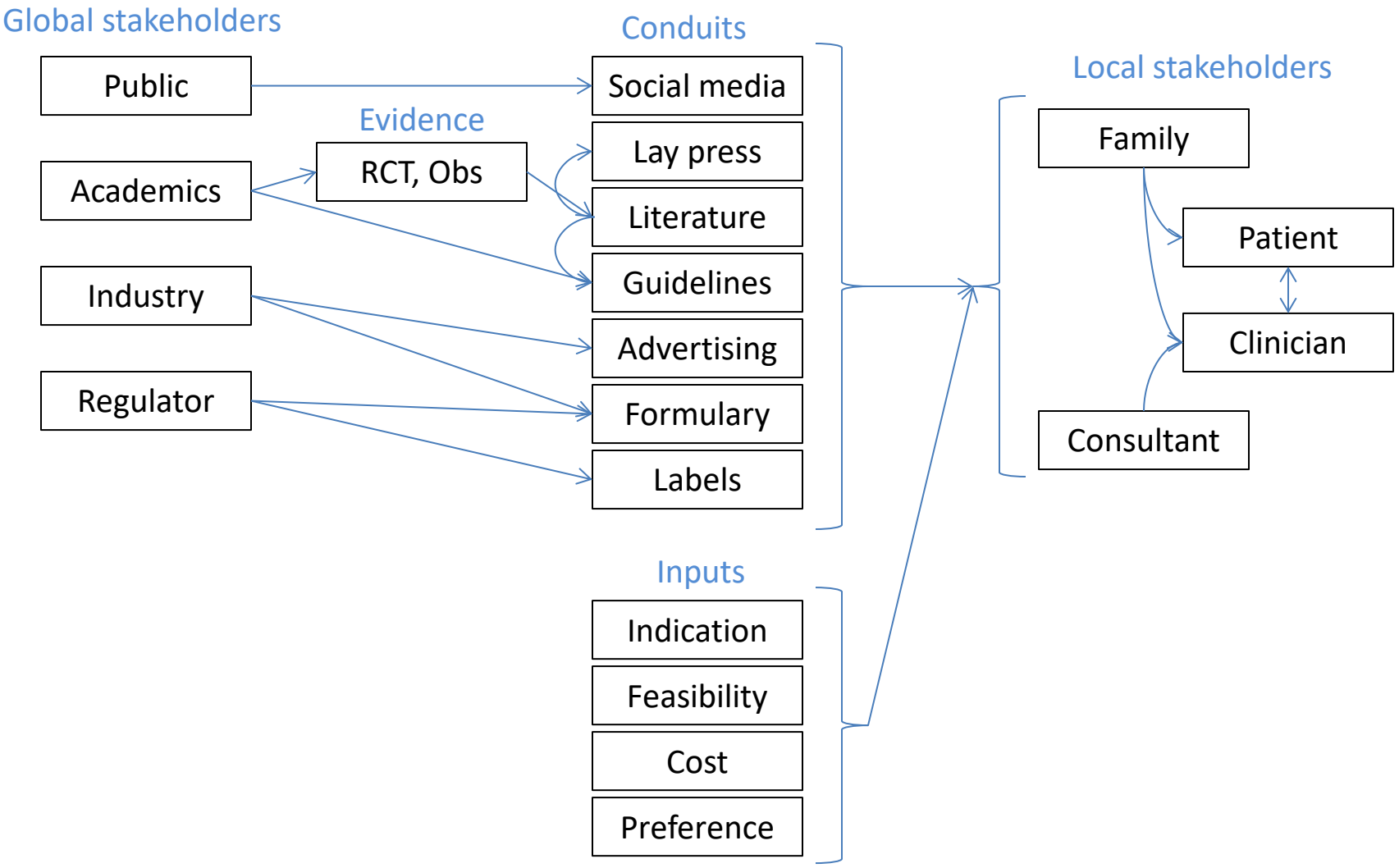
Proxy for manual chart review



Swerdel JBI 2019

OHDSI in Action: Characterization

Treatment Pathways



OHDSI participating data partners

Abbre-viation	Name	Description	Population, millions
AUSOM	Ajou University School of Medicine	South Korea; inpatient hospital EHR	2
CCAE	MarketScan Commercial Claims and Encounters	US private-payer claims	119
CPRD	UK Clinical Practice Research Datalink	UK; EHR from general practice	11
CUMC	Columbia University Medical Center	US; inpatient EHR	4
GE	GE Centricity	US; outpatient EHR	33
INPC	Regenstrief Institute, Indiana Network for Patient Care	US; integrated health exchange	15
JMDC	Japan Medical Data Center	Japan; private-payer claims	3
MDCD	MarketScan Medicaid Multi-State	US; public-payer claims	17
MDCR	MarketScan Medicare Supplemental and Coordination of Benefits	US; private and public-payer claims	9
OPTUM	Optum ClinFormatics	US; private-payer claims	40
STRIDE	Stanford Translational Research Integrated Database Environment	US; inpatient EHR	2
HKU	Hong Kong University	Hong Kong; EHR	1

Proceedings of the National Academy of Sciences, 2016

PNAS

COLLOQUIUM PAPER

Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak[a,b,c,1], Patrick B. Ryan[c,d], Jon D. Duke[c,e], Nigam H. Shah[c,f], Rae Woong Park[c,g], Vojtech Huser[c,h], Marc A. Suchard[c,i,j,k], Martijn J. Schuemie[c,d], Frank J. DeFalco[c,d], Adler Perotte[a,c], Juan M. Banda[c,f], Christian G. Reich[c,l], Lisa M. Schilling[c,m], Michael E. Matheny[c,n,o], Daniella Meeker[c,p,q], Nicole Pratt[c,r], and David Madigan[c,s]

[a]Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032; [b]Medical Informatics Services, NewYork-Presbyterian Hospital, New York, NY 10032; [c]Observational Health Data Sciences and Informatics, New York, NY 10032; [d]Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08560; [e]Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN 46205; [f]Center for Biomedical Informatics Research, Stanford University, CA 94305; [g]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea, 443-380; [h]Lister Hill National Center for Biomedical Communications (National Library of Medicine), National Institutes of Health, Bethesda, MD 20894; [i]Department of Biomathematics, University of California, Los Angeles, CA 90095; [j]Department of Biostatistics, University of California, Los Angeles, CA 90095; [k]Department of Human Genetics, University of California, Los Angeles, CA 90095; [l]Real World Evidence Solutions, IMS Health, Burlington, MA 01809; [m]Department of Medicine, University of Colorado School of Medicine, Aurora, CO 80045; [n]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN 37212; [o]Geriatric Research, Education and Clinical Center, VA Tennessee Valley Healthcare System, Nashville, TN 37212; [p]Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089; [q]Department of Pediatrics, University of Southern California, Los Angeles, CA 90089; [r]Division of Health Sciences, University of South Australia, Adelaide, SA, Australia 5001; and [s]Department of Statistics, Columbia University, New York, NY 10027

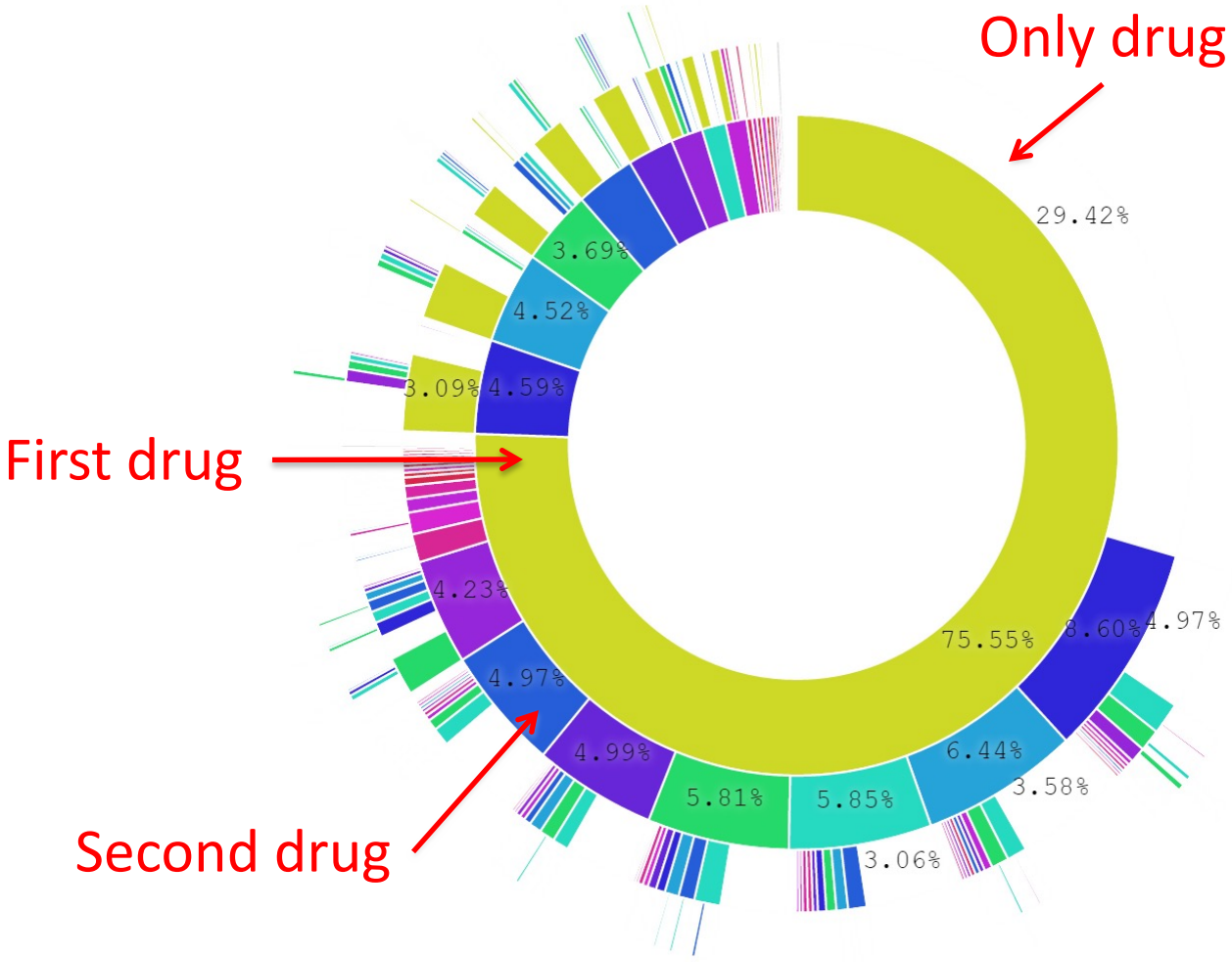
Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved April 5, 2016 (received for review June 14, 2015)

Observational research promises to complement experimental research by providing large, diverse populations that would be infeasible for an experiment. Observational research can test its own clinical hypotheses, and observational studies also can contribute to the design of experiments and inform the generalizability of experimental research. Understanding the diversity of populations

Without sufficiently broad databases available in the first stage, randomized trials are designed without explicit knowledge of actual disease status and treatment practice. Literature reviews are restricted to the population choices of previous investigations, and pilot studies usually are limited in scope. By exploiting the ClinicalTrials.gov national trial registry (9) and electronic health

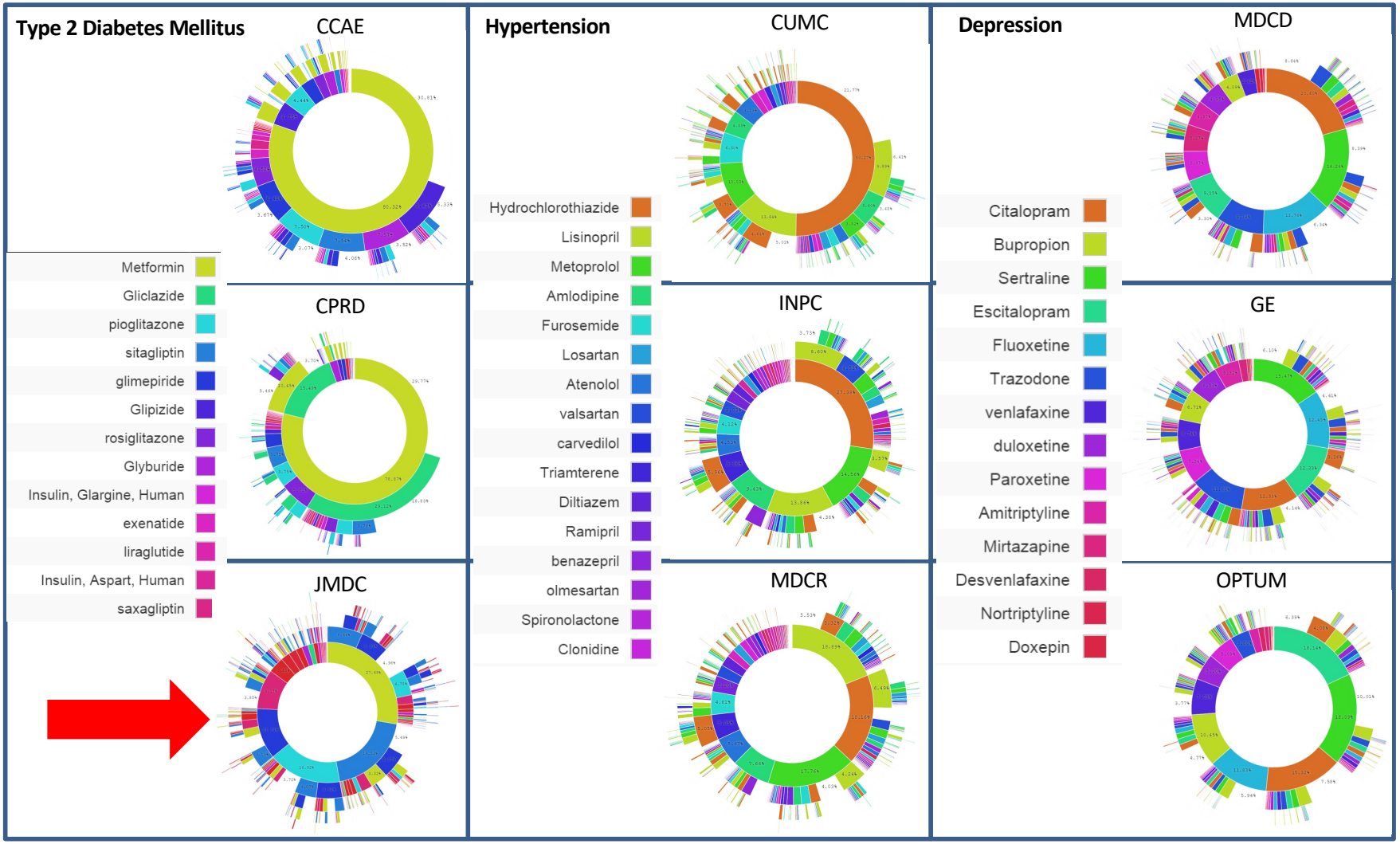
Treatment pathways for diabetes

T2DM : All databases



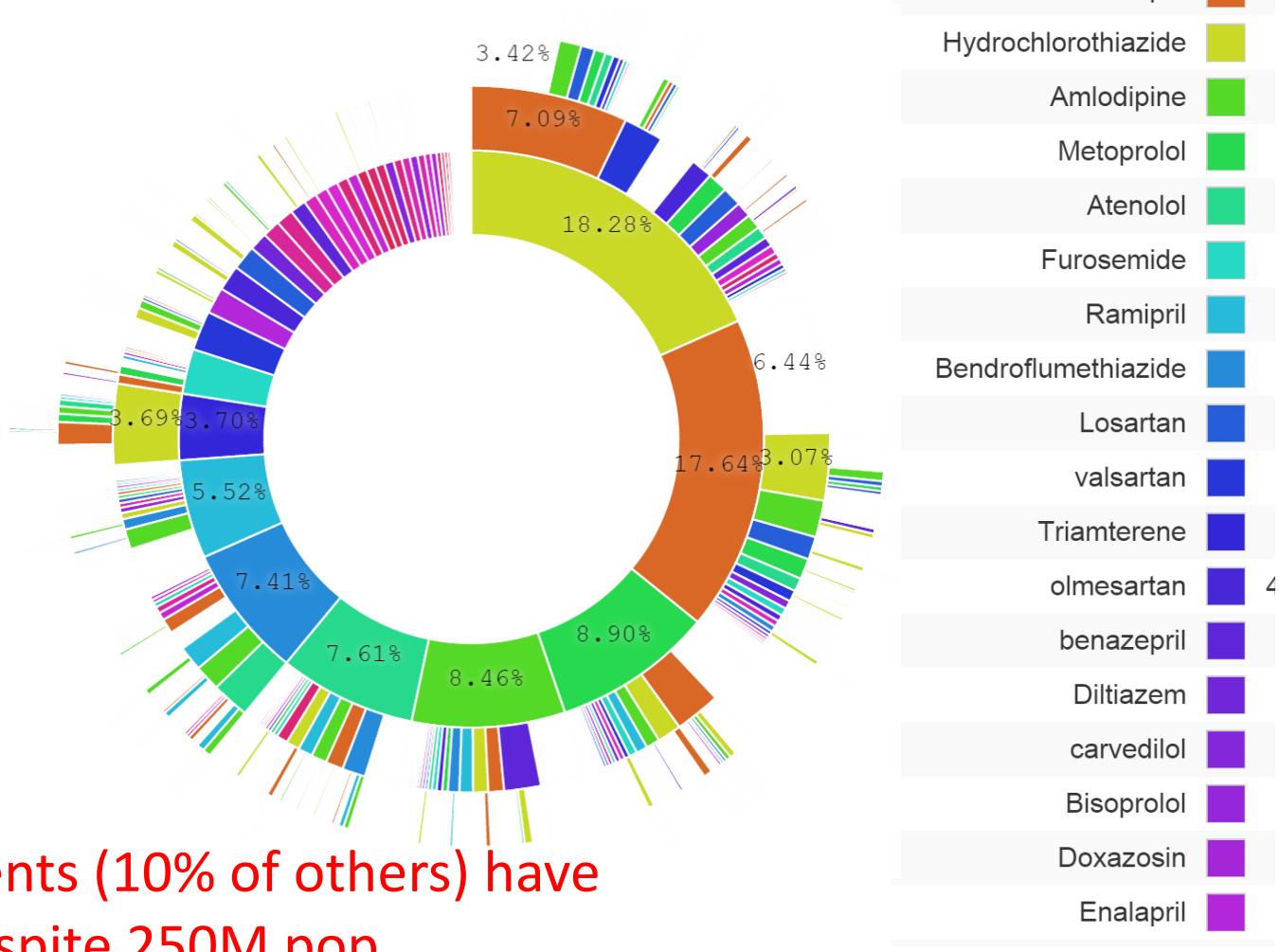
- Metformin
- pioglitazone
- sitagliptin
- Glipizide
- glimepiride
- Gliclazide
- Glyburide
- rosiglitazone
- Insulin, Glargine, Human
- exenatide
- Insulin, Aspart, Human
- liraglutide
- saxagliptin
- Insulin, Lispro, Human
- Glucose
- Insulin, Isophane, Human

Population-level heterogeneity across systems, and patient-level heterogeneity within systems



Patient-level heterogeneity

HTN: All databases



25% of HTN patients (10% of others) have a unique path despite 250M pop

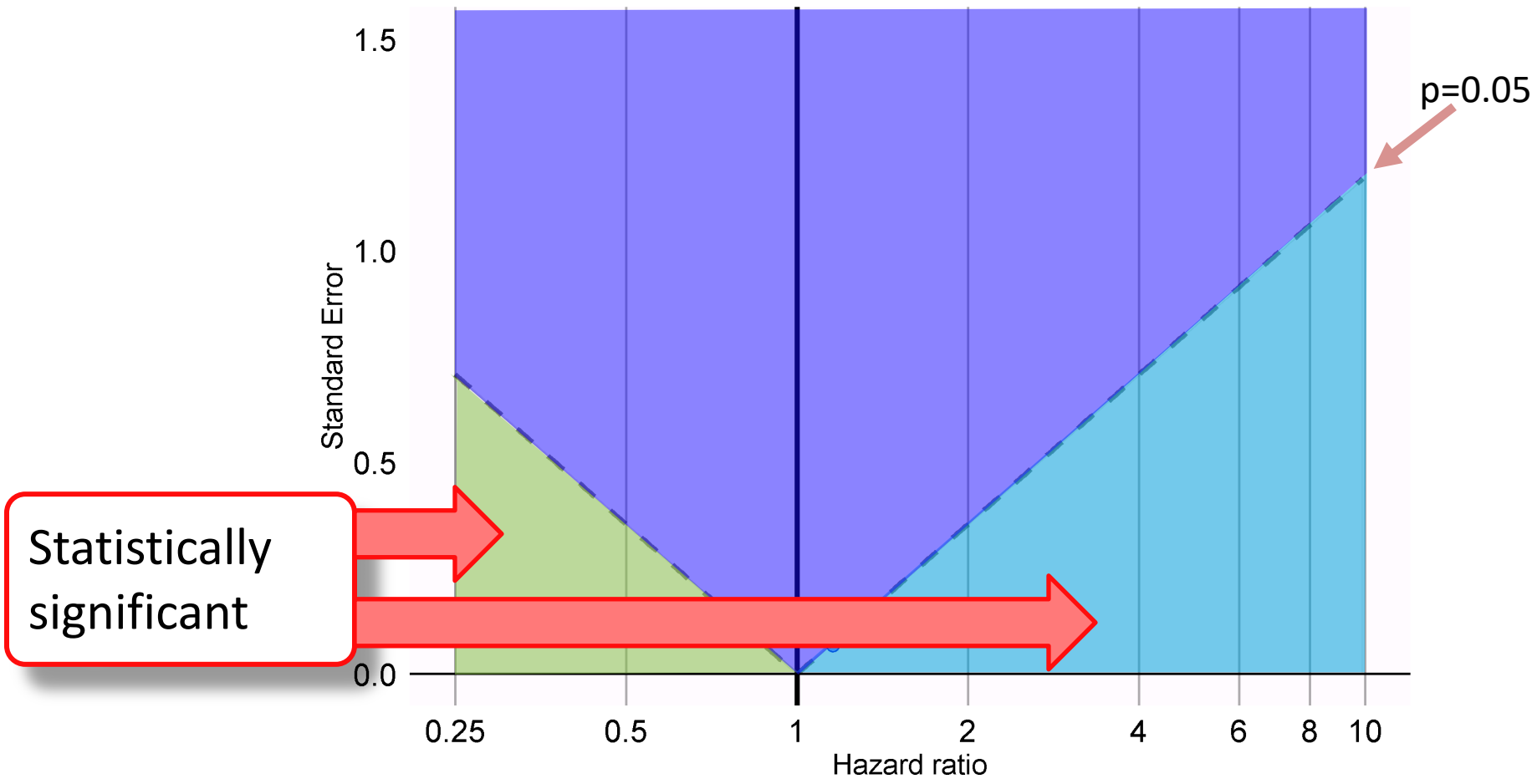
Conclusions: Network research

- It is feasible to encode the world population in a single data model
  - Over 10% now
- Generating evidence is feasible
- Stakeholders willing to share results
- Able to accommodate vast differences in privacy and research regulation

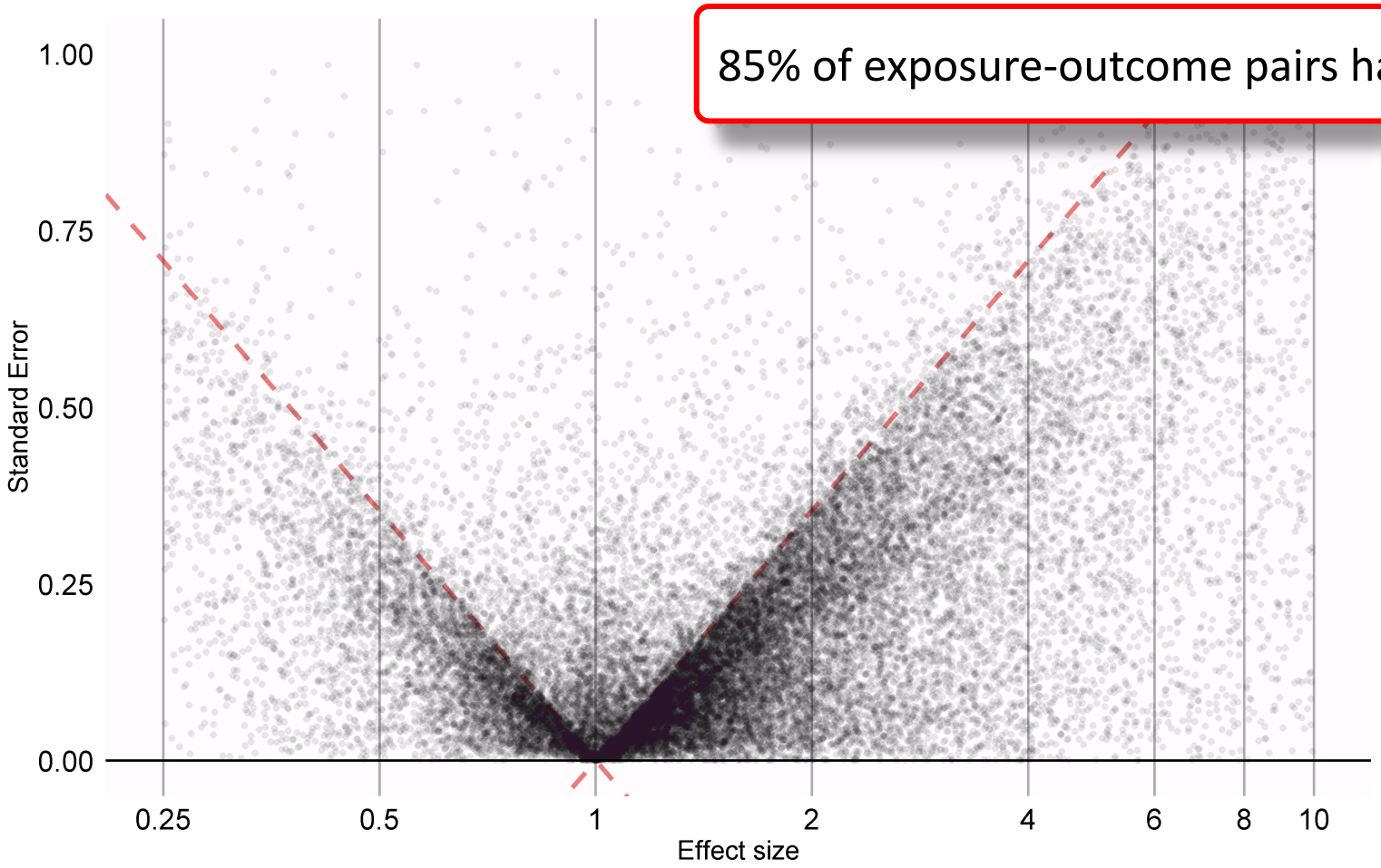
Population-level estimation

- Causal inference, hypothesis testing
- Creating reliable evidence
- OHDSI: study it scientifically
  - Distribution of study designs, parameters, databases, hypotheses

Standard error vs effect size

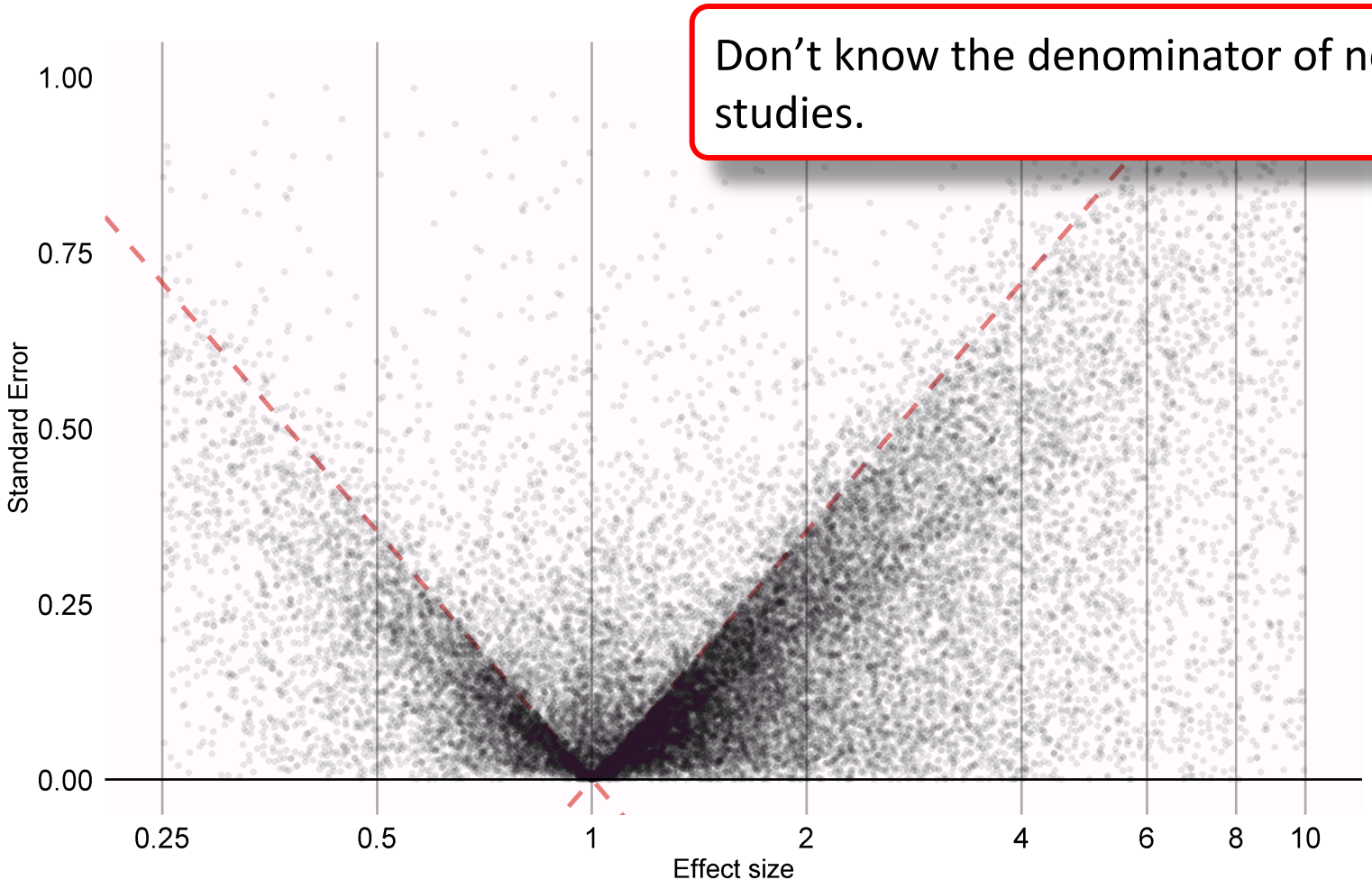


Observational research results in literature

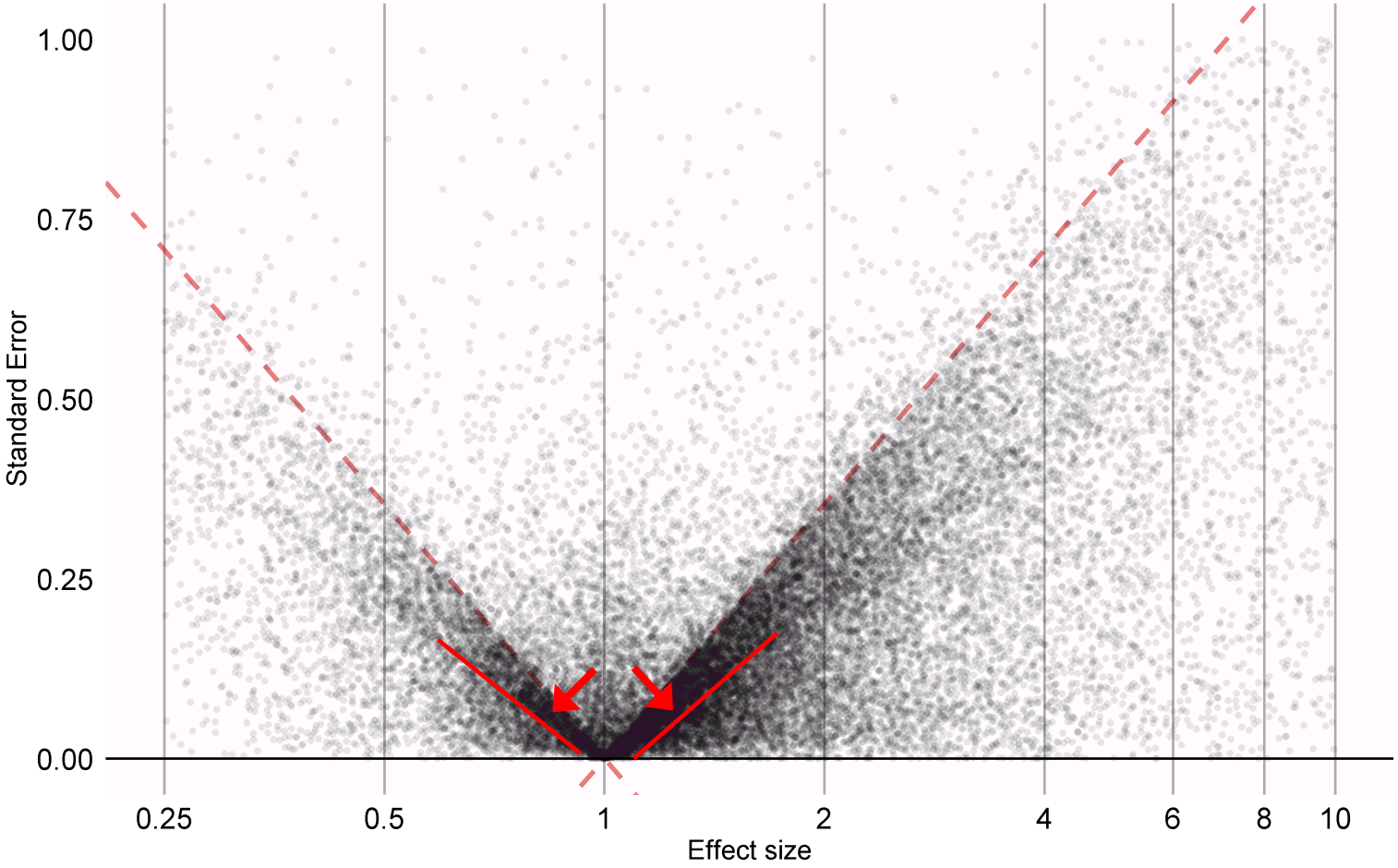


Schuemie Phil Trans A 2018

Observational research results in literature

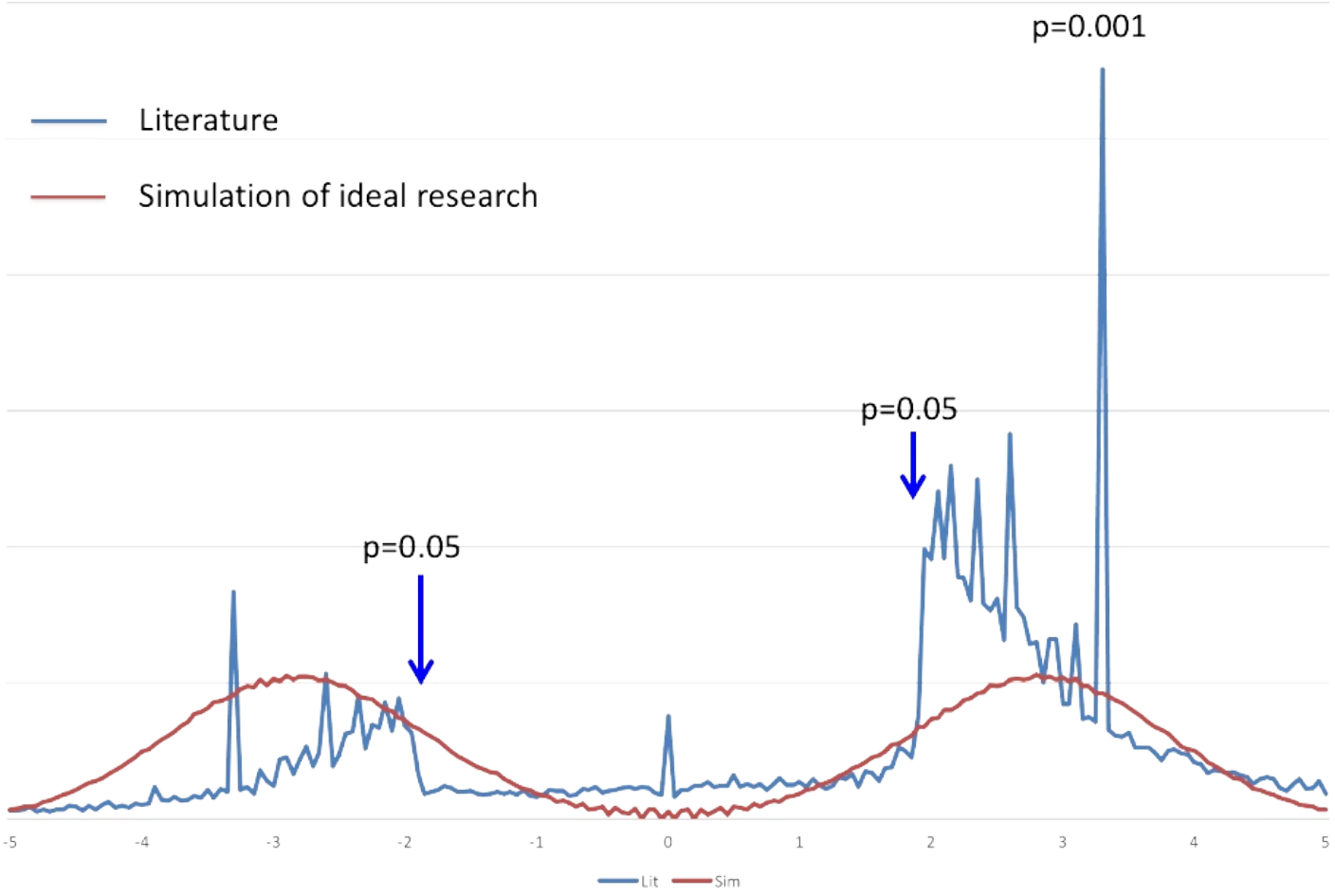


Observational research results in literature



29,982 estimates
11,758 papers

We’re not just guessing right

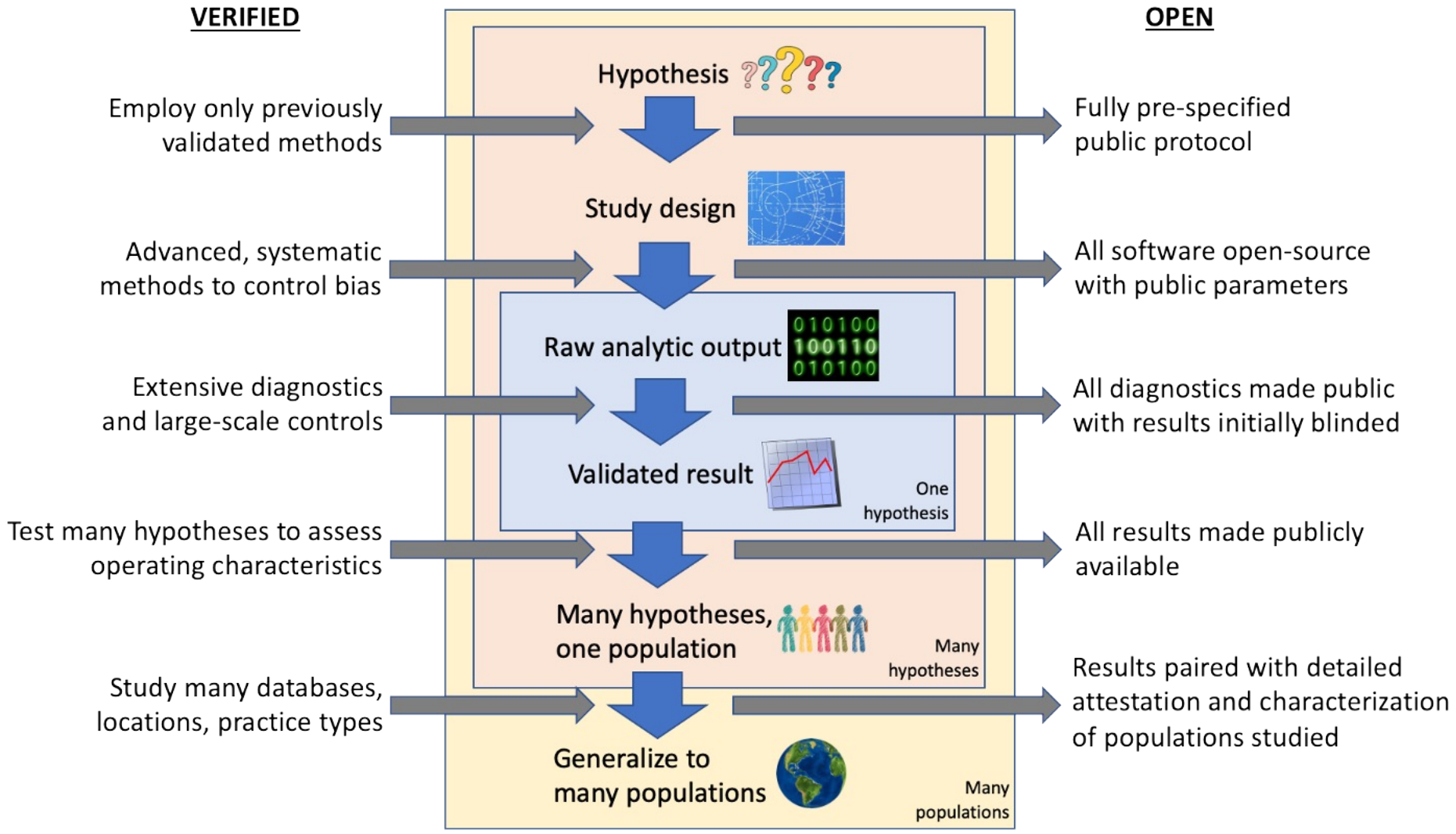


Hripcsak Yearb Med Inform 2021

Observational research results in literature

- Individuals may produce good research studies
- In aggregate, the medical observational research system is a data dredging machine

Verified and open



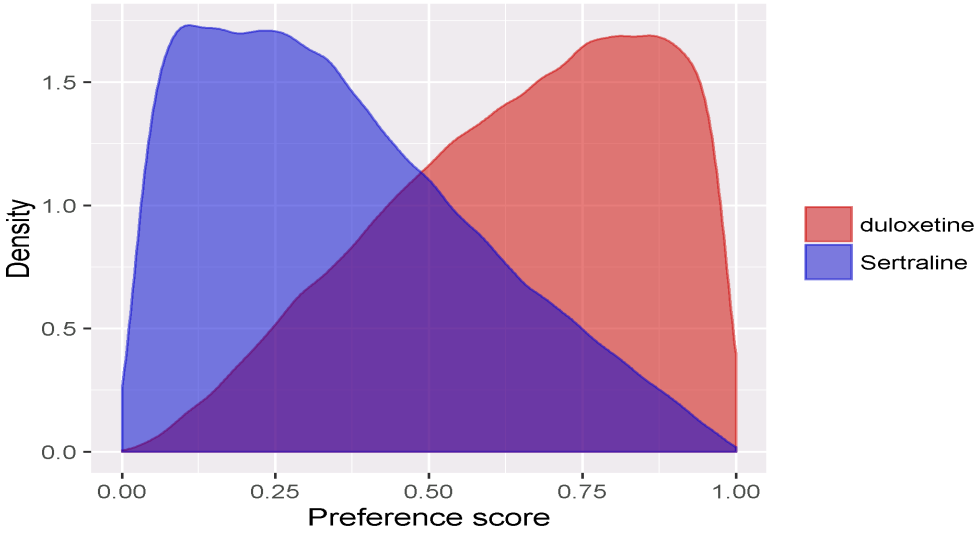
10 LEGEND Principles (Large-scale Evidence Generation and Evaluation across a Network of Databases)

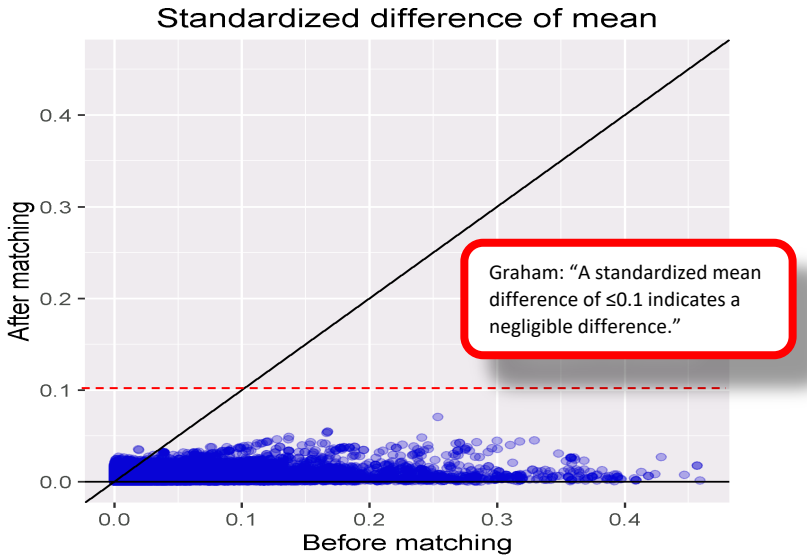
- LEGEND will generate evidence at a large scale
- Dissemination of the evidence will not depend on the estimated effects
- LEGEND will generate evidence using a prespecified analysis design
- LEGEND will generate evidence by consistently applying a systematic process across all research questions
  - No thumb on the scale
- LEGEND will generate evidence using best practices
- LEGEND will include empirical evaluation through the use of control questions
- LEGEND will generate evidence using open-source software that is freely available to all
- LEGEND will not be used to evaluate new methods
- LEGEND will generate evidence across a network of multiple databases
- LEGEND will maintain data confidentiality; patient-level data will not be shared between sites in the network

Schuemie JAMIA 2020

Addressing reproducibility #1

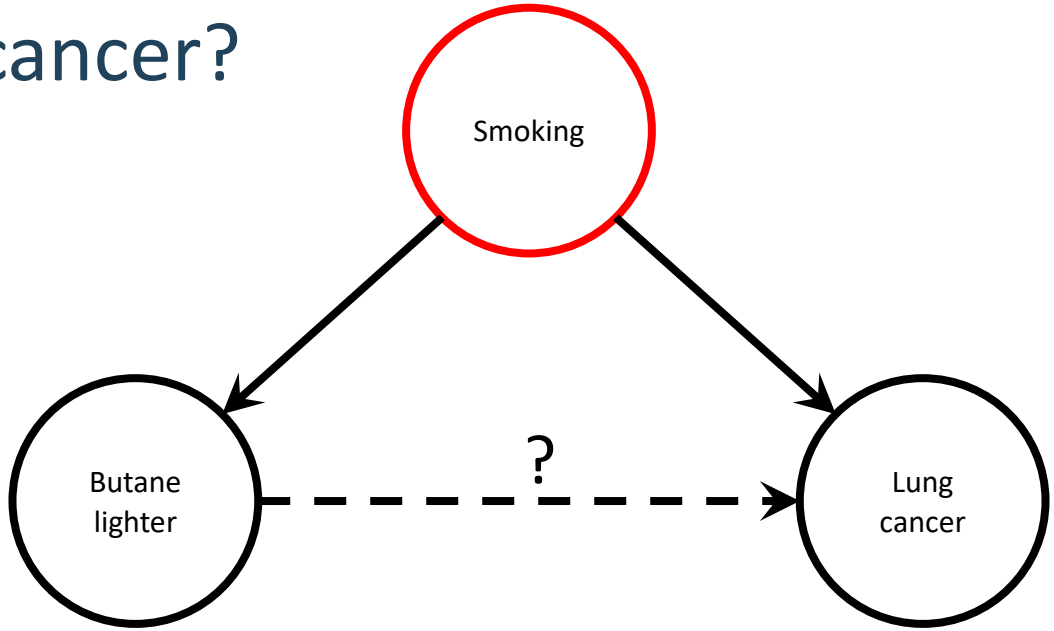
1. **Propensity score adjustment with large-scale covariate set: measured confounding (and some unmeasured?)**
   - Take advantage of the huge databases and balance on tens of thousands of covariates, pulling in other variables (BP)
   - Mimic balance of randomization (imperfect)
   - Don't rely on human expertise to select confounders: **systematic**
   - Diagnostics





Confounding

- Does butane gas cause lung cancer?



Propensity score to address confounding

- Propensity score = probability of belonging to the target cohort vs. the comparator cohort, given the baseline covariates
- Propensity score can be used as a 'balancing score': if the two cohorts have similar propensity score distribution, then the distribution of covariates should be the similar
- Balance the propensity -> balance the covariates
- Balance the covariates -> the comparisons are similar
  - Make a causal assertion: must be due to the treatment

Rubin Biometrika 1983

How to select the confounders

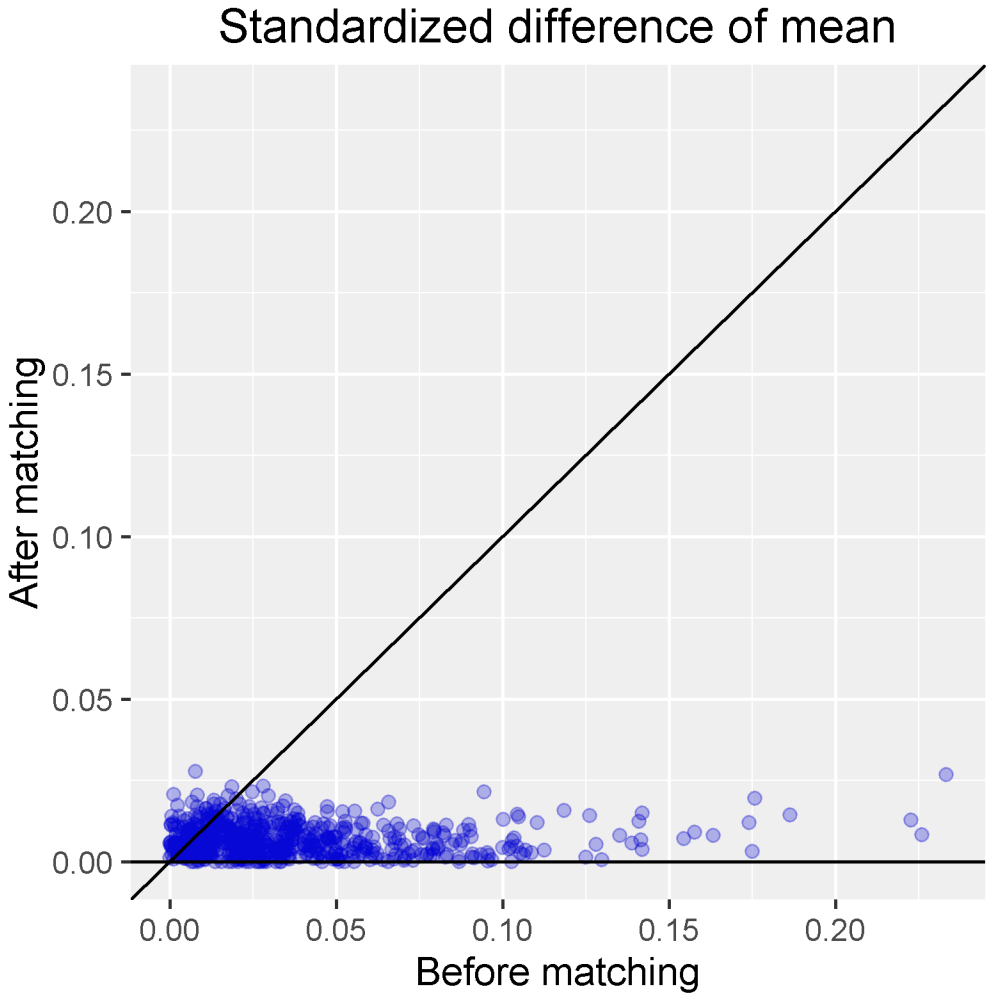
- Manual selection -> poor agreement
  - Chien 2015: age, month, gender, #visits, income urbanization, #drugs, specific drugs, Charlson, comorbidities (16), +HDPS variables
  - Hicks 2018: age, sex, year of cohort entry, body mass index, smoking status, alcohol related disorders (including alcoholism, alcoholic cirrhosis of the liver, alcoholic hepatitis, and hepatic failure), and history of lung diseases (including pneumonia, tuberculosis, and chronic obstructive pulmonary disease), duration of HTN Rx, statin use, #drugs
  - Ku 2018: age, sex, race, income status, baseline HF, baseline myocardial infarction, baseline peripheral artery disease, baseline stroke, baseline eGFR, baseline proteinuria, and time-dependent covariates including diabetes mellitus, obesity, systolic blood pressure, statin use, aspirin use, diuretic use, and concurrent use of other antihypertensive agents for the outcome of HF
  - Magid 2010: age, gender, days on thiazide prior to 2nd agent start, # of visits prior to thiazide, Mean Systolic BP, Mean Diastolic BP, Chronic Obstructive Pulmonary Disease, Hyperlipidemia, Cancer, Dementia, Chronic liver disease, Depression
  - Hasvold 2014: age, gender, elevated blood glucose, overweight and low socio-economic status are known risk factors for diabetes, High cholesterol and hypertension are additionally known risk factors for CVD
- Empirical selection

Large-scale propensity score (LSPS)

- A **systematic** approach to propensity adjustment
- Use a large set of covariates (10,000 < n < 100,000)
- But don't want to balance *everything*
  - Mediators – pre-treatment
  - Simple colliders – pre-treatment
  - Instruments – diagnostics, domain knowledge
  - M-bias – correlation with underlying causes
- Fit a propensity model
  - LASSO (regularized regression) because #variables > #cases
- Match or stratify on propensity score
- Diagnostic: check that covariate balance is achieved on all observed variables

Tian Int J Epi 2018

Diagnostic: Covariate balance



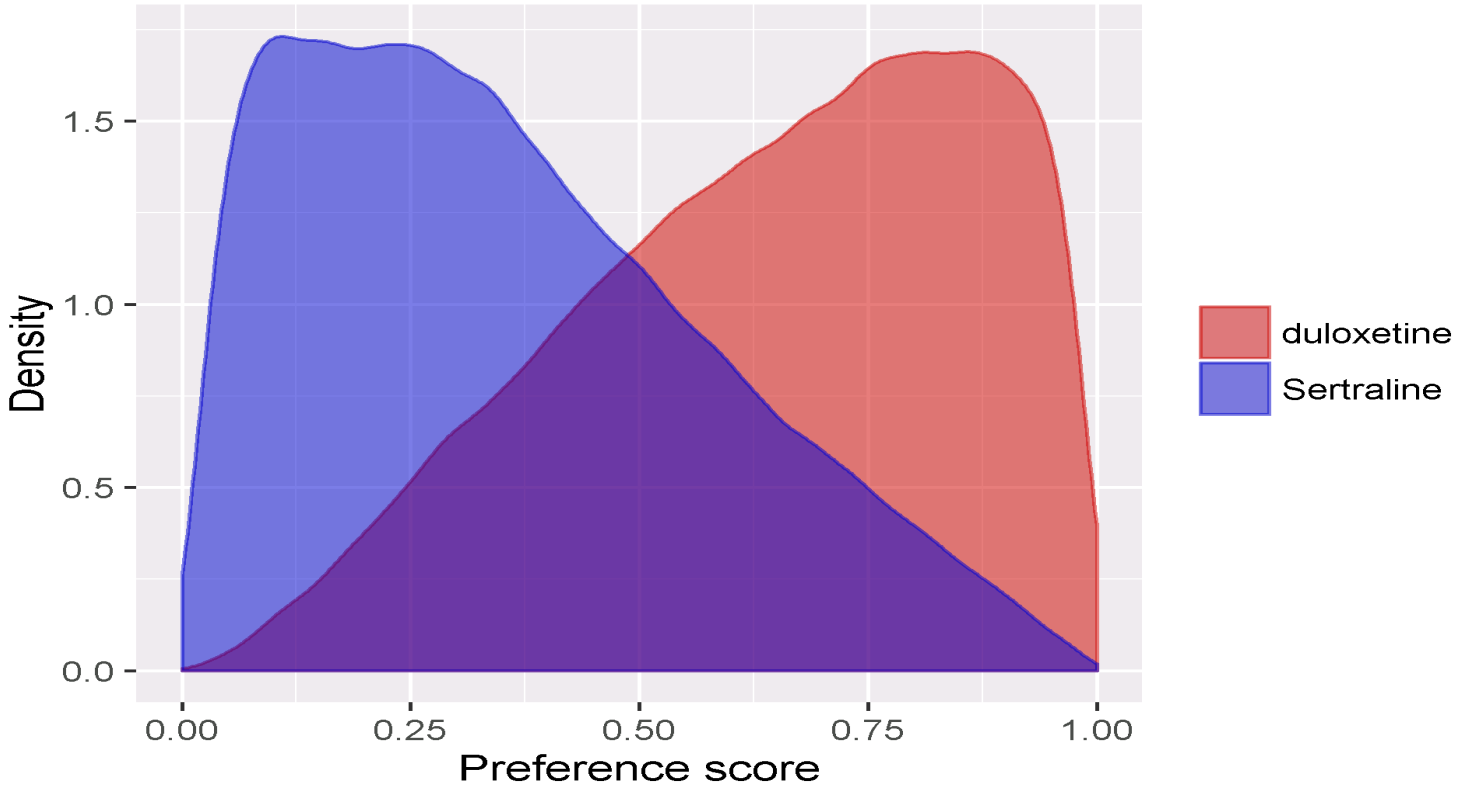
Plot 60,000 covariates;
most are binary:

$$\frac{\text{abs}(P_{\text{target group}} - P_{\text{comparator group}})}{\text{standard deviation}}$$

Graham: "A standardized mean difference of ≤0.1 indicates a negligible difference."

Diagnostic: equipoise

- What is the overlap between the groups
  - If too small, poor generalizability and stability



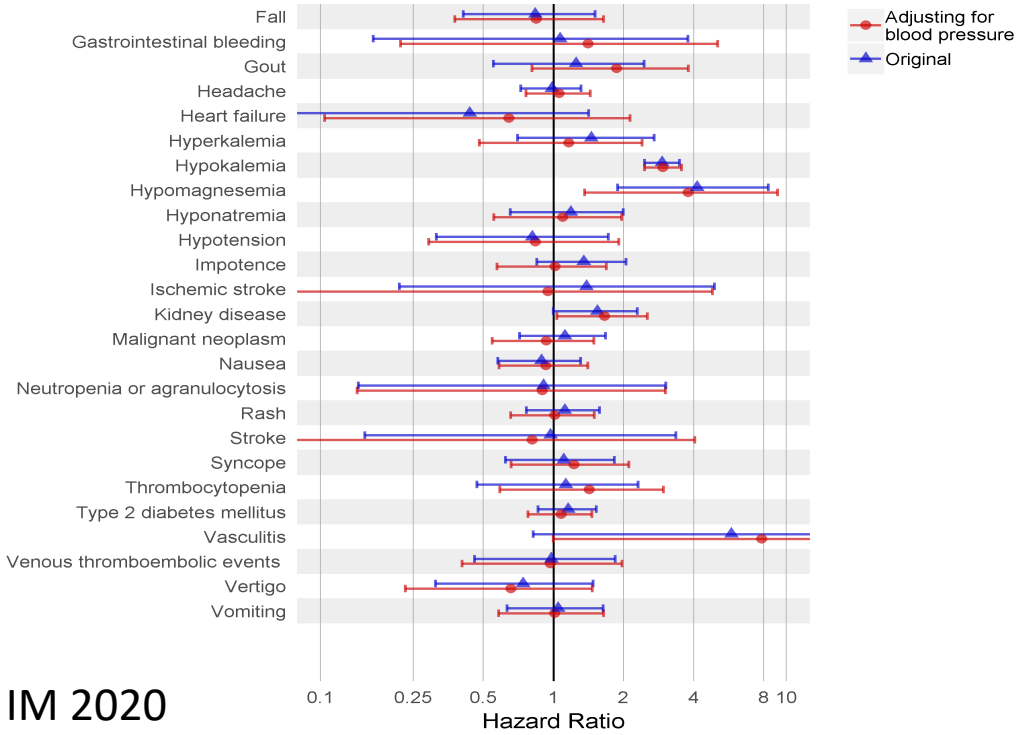
What about confounding that is not measured?

- Some confounders are not directly measured but may be correlated with the many variables used by LSPS
- Hypertension study
  - Baseline blood pressure is an important confounder
  - But not measured in most databases, except Optum EHR database

Same result with or without BP

Achieves (near) balance of BP despite its absence

	Systolic BP	Diastolic BP
Unadjusted	0.200	0.168
LSPS 60K	0.126	0.094
LSPS 60K+BP	0.046	0.001



Hripcsak JAMIA IM 2020

Adjusting for other variables that are not directly measured

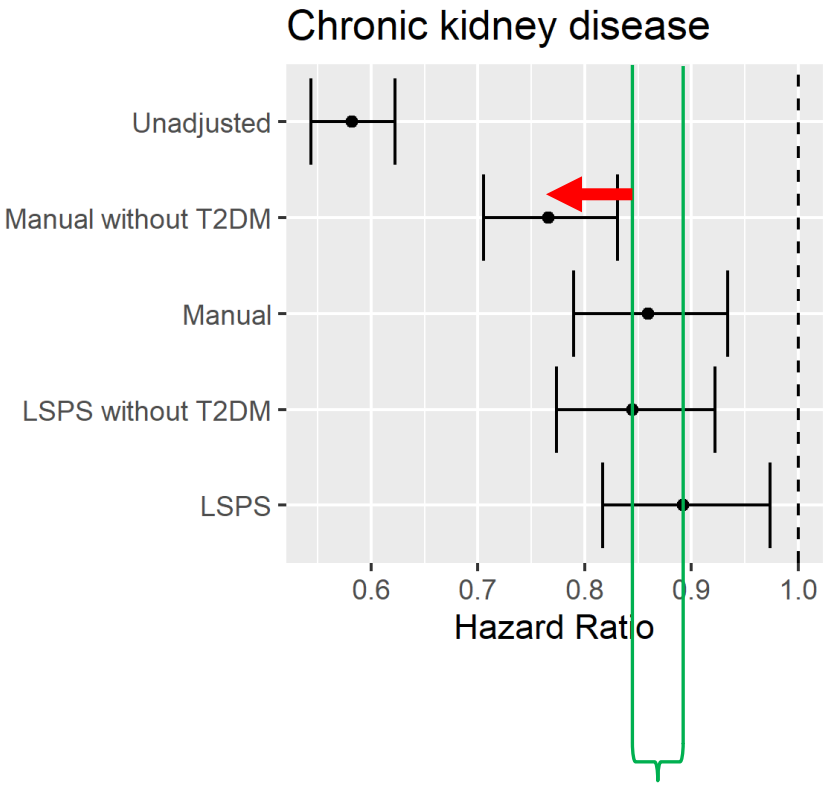
- We seem to balance lots of stuff beyond what was put into LSPS
  - Medications balance conditions (Dx)
  - Conditions balance medications (mostly)
  - Non-CV balances cardiovascular
  - Demographics do **not** work

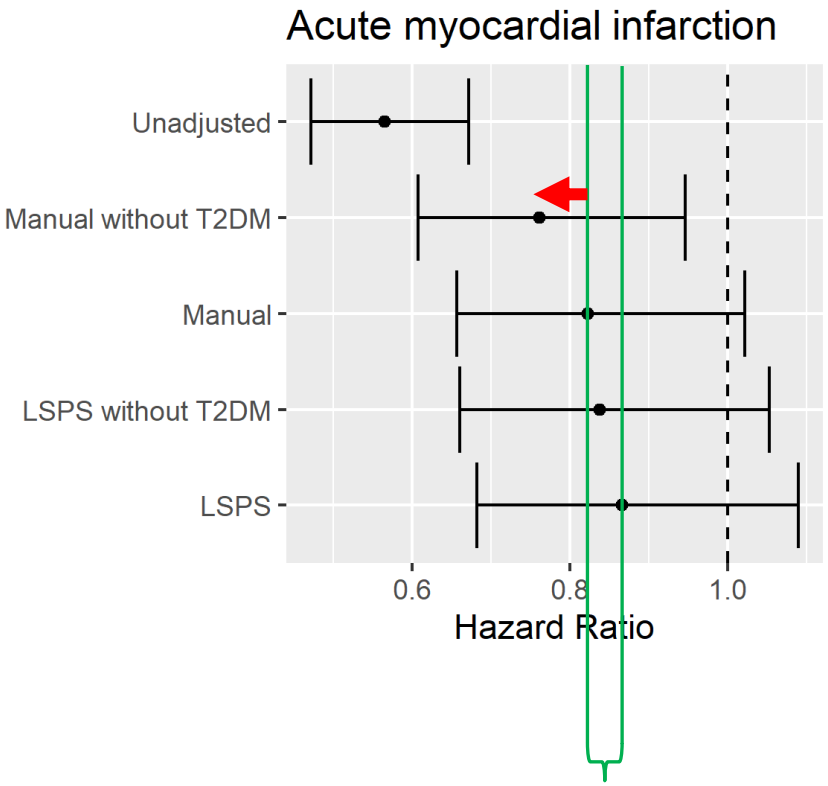
Test	Total Covariates	# unbal. before matching	# unbal. after matching	Max diff before	Max diff after
Full covariates	72,203	166	0	0.3975	0.035
Demographics only	72,203	166	158	0.3975	0.375
No conditions	72,203	166	0	0.3975	0.068
No drugs	72,203	166	8	0.3975	0.108
No procedures	72,203	166	0	0.3975	0.059
No cardiovascular-related concepts	72,203	166	0	0.3975	0.074

Chen AMIA 2020

LSPS vs. manual selection on the effect of a missing confounder

- Lisinopril vs hydrochlorothiazide
  - Confounder type 2 diabetes

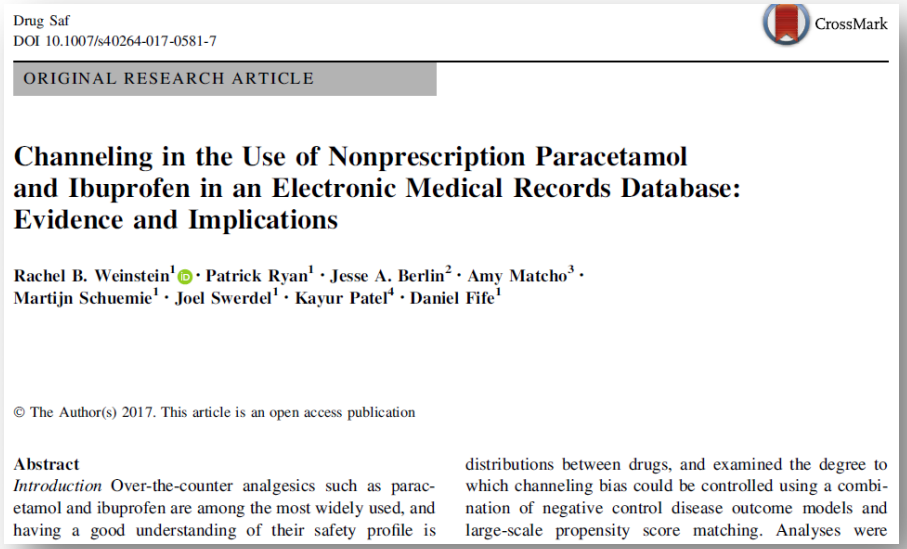




Zhang arXiv 2021

LSPS

- Reduce bias if balance on many covariates instead of a few human-selected covariates (bias measured via negative controls)



Drug Saf
DOI 10.1007/s40264-017-0581-7

CrossMark

ORIGINAL RESEARCH ARTICLE

Channeling in the Use of Nonprescription Paracetamol and Ibuprofen in an Electronic Medical Records Database: Evidence and Implications

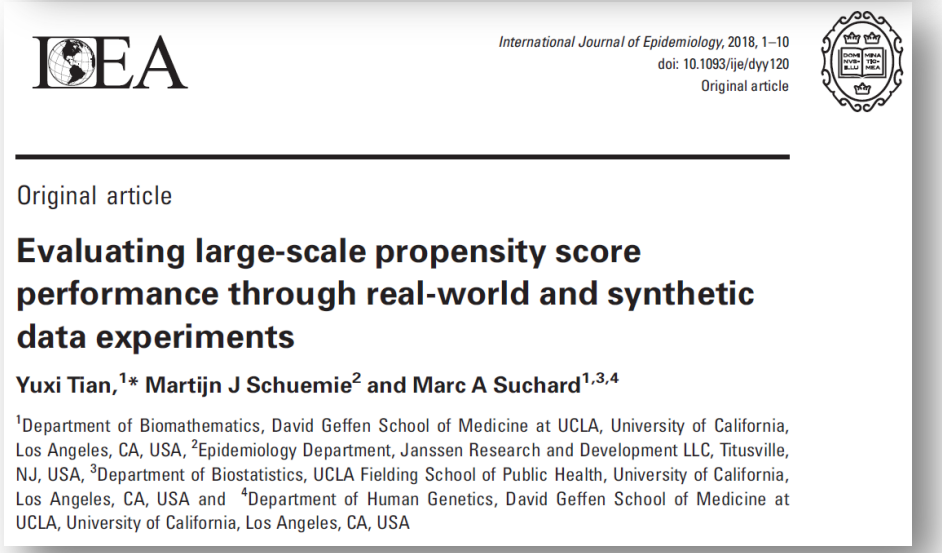
Rachel B. Weinstein[1] · Patrick Ryan[1] · Jesse A. Berlin[2] · Amy Matcho[3] · Martijn Schuemie[1] · Joel Swerdel[1] · Kayur Patel[4] · Daniel Fife[1]

Abstract

Introduction Over-the-counter analgesics such as paracetamol and ibuprofen are among the most widely used, and having a good understanding of their safety profile is

distributions between drugs, and examined the degree to which channeling bias could be controlled using a combination of negative control disease outcome models and large-scale propensity score matching. Analyses were

- LSPS performs better than competing methods like high-dimensional propensity score (HDPS)



IJEA

International Journal of Epidemiology, 2018, 1–10
doi: 10.1093/ije/dyy120
Original article

Original article

Evaluating large-scale propensity score performance through real-world and synthetic data experiments

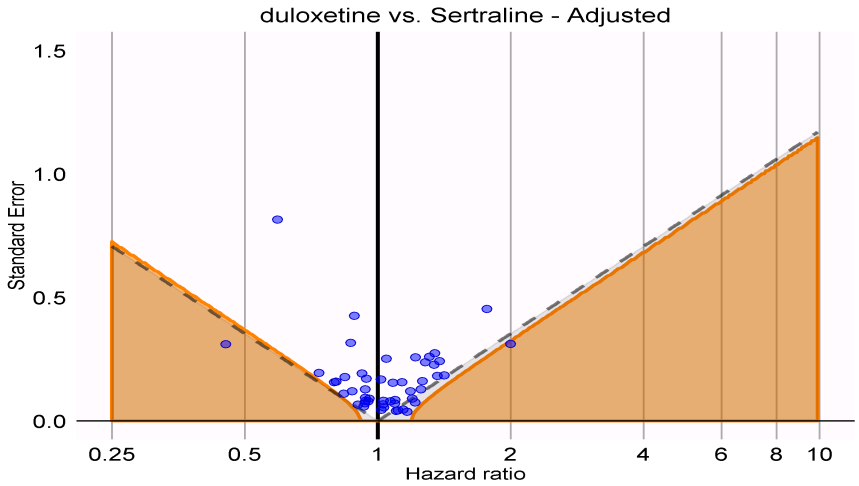
Yuxi Tian,[1]* Martijn J Schuemie[2] and Marc A Suchard[1,3,4]

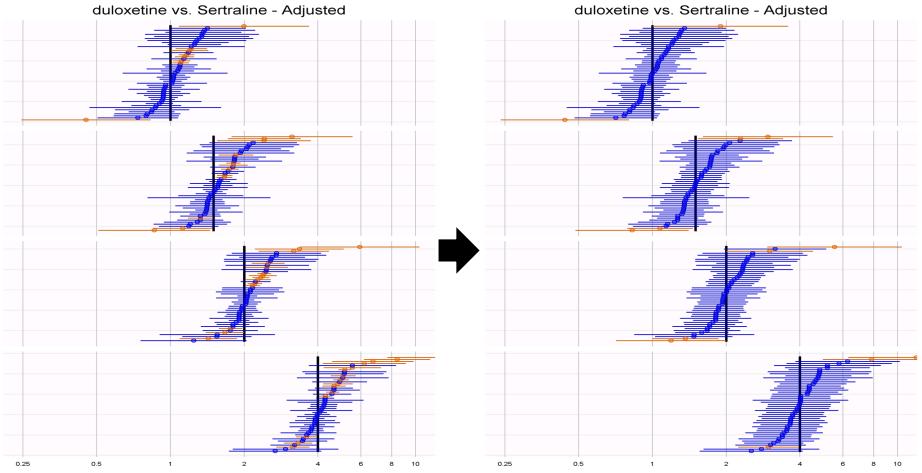
[1]Department of Biomathematics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA, [2]Epidemiology Department, Janssen Research and Development LLC, Titusville, NJ, USA, [3]Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, CA, USA and [4]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA

Addressing reproducibility #2

2. Confidence interval calibration using negative controls: unmeasured confounding

- Address residual confounding using hypotheses you know the answer for
- If too many are positive, then systematic error is operative
- Calibrate to keep the type 1 error at 0.05
- Diagnostics





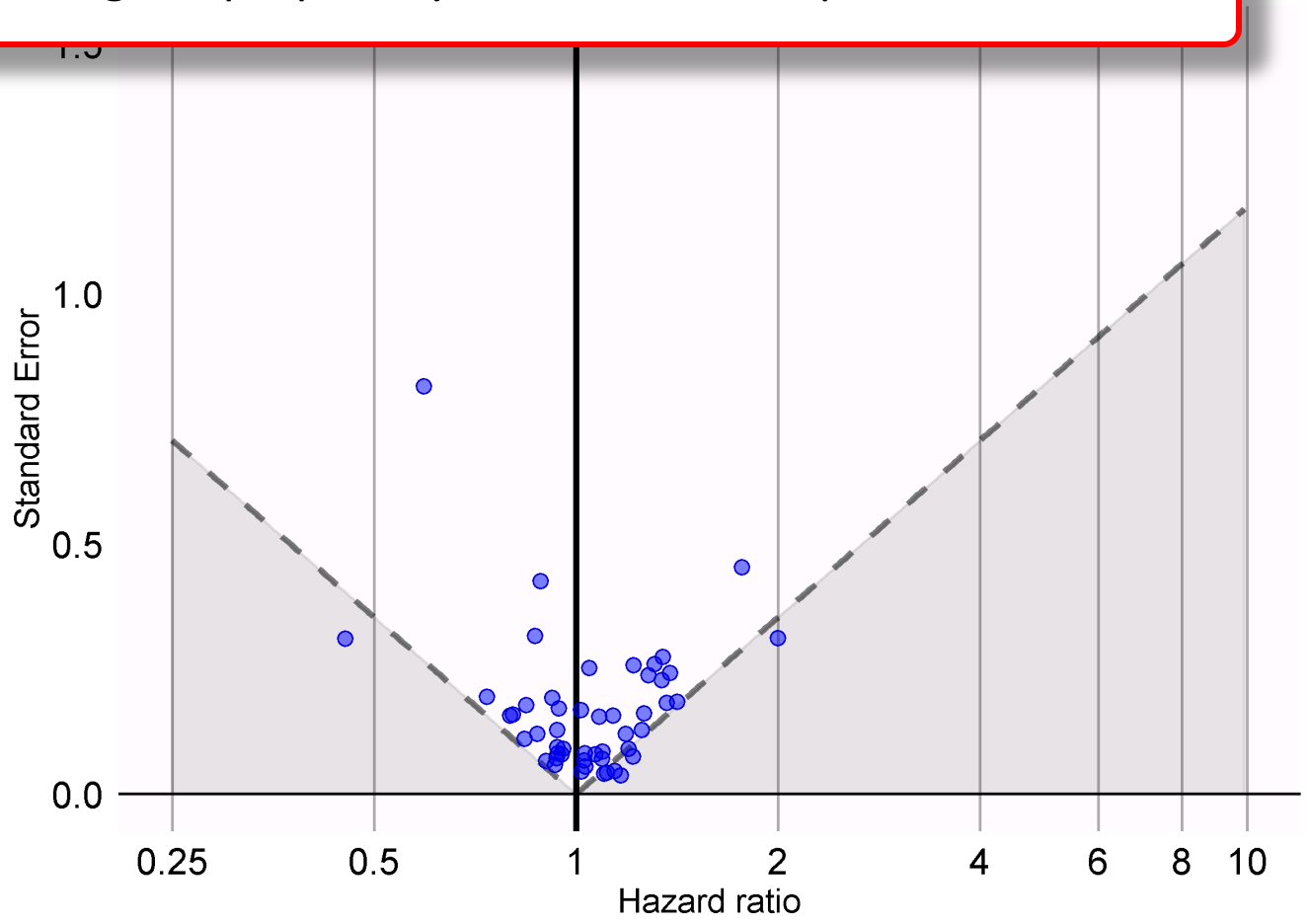
Negative controls

- Negative control
  - exposure-outcome where relative risk is believed to be 1
  - example: ingrowing nail

- OHDSI employs 50-100 negative controls
  - systematic methods allow large scale

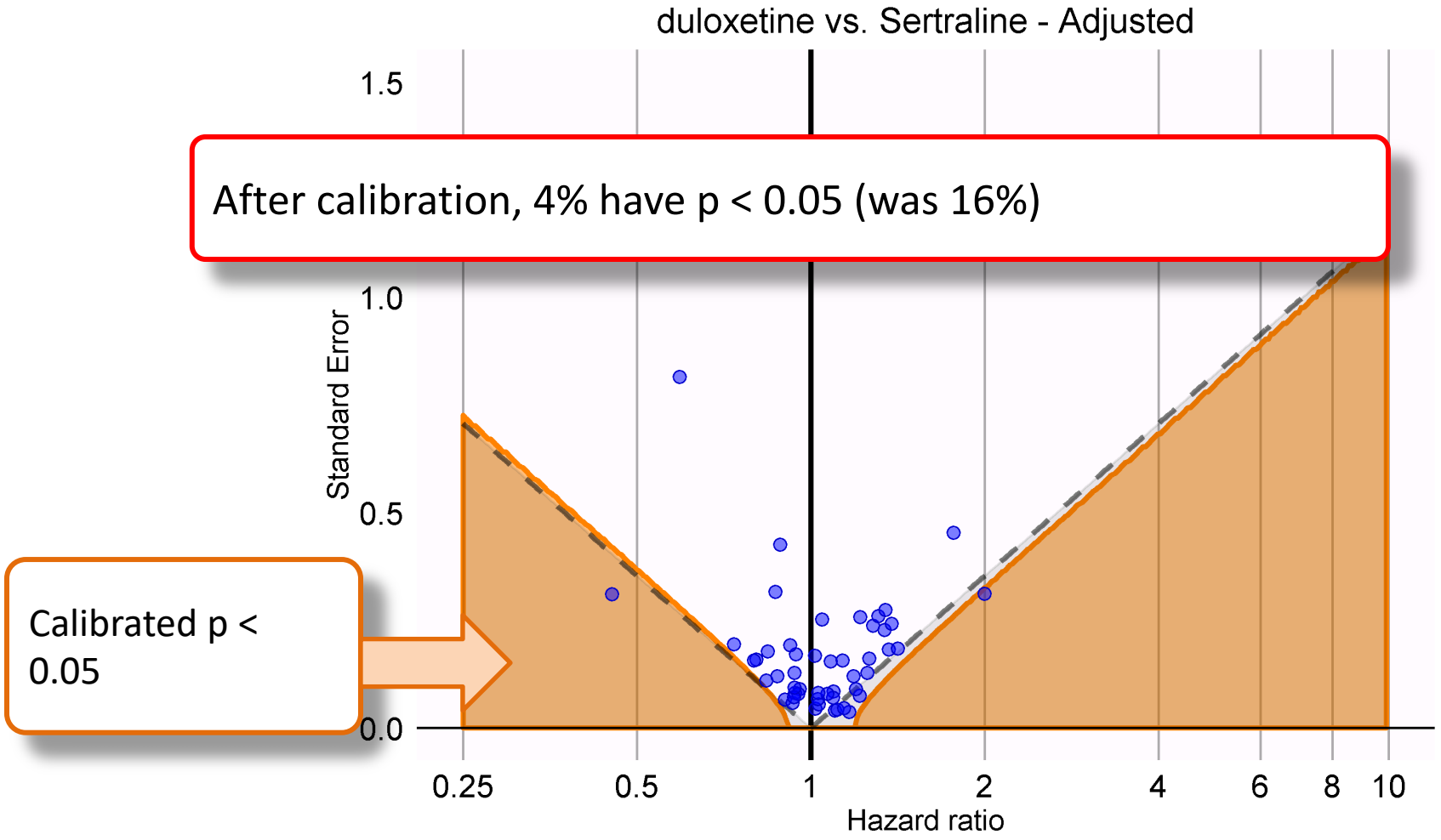
All negative controls - adjusted

When using the propensity score, 16% have p < 0.05



Schuemie OHDSI 2016

P-value calibration

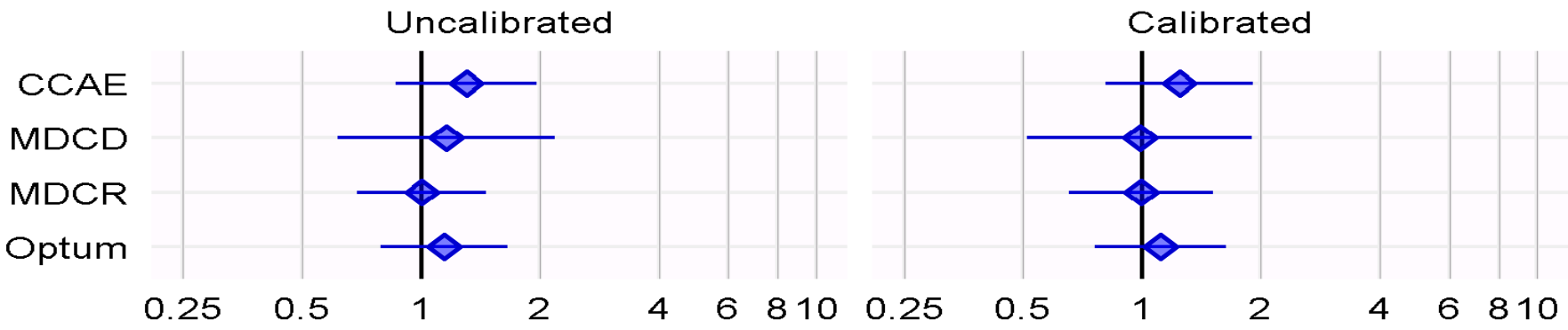


Schuemie OHDSI 2016

Addressing reproducibility #3

3. Multiple databases, locations, practice types

- Look for consistency among databases, practices
- Combine via meta-analysis
- Aids generalizability

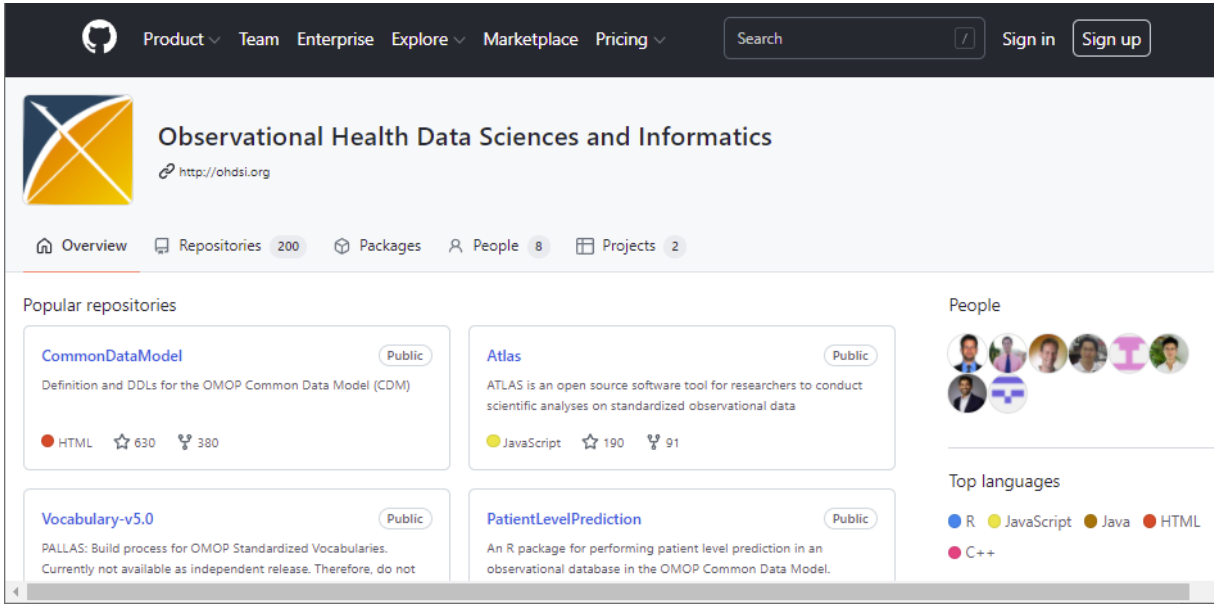


(Recent grant review)

Addressing reproducibility #4

4. Publish all hypotheses, code, parameters, runs

- Pre-specify protocol so cannot cheat
- Publish all code so that others can run it
- Publish masked results, check diagnostics, reveal results

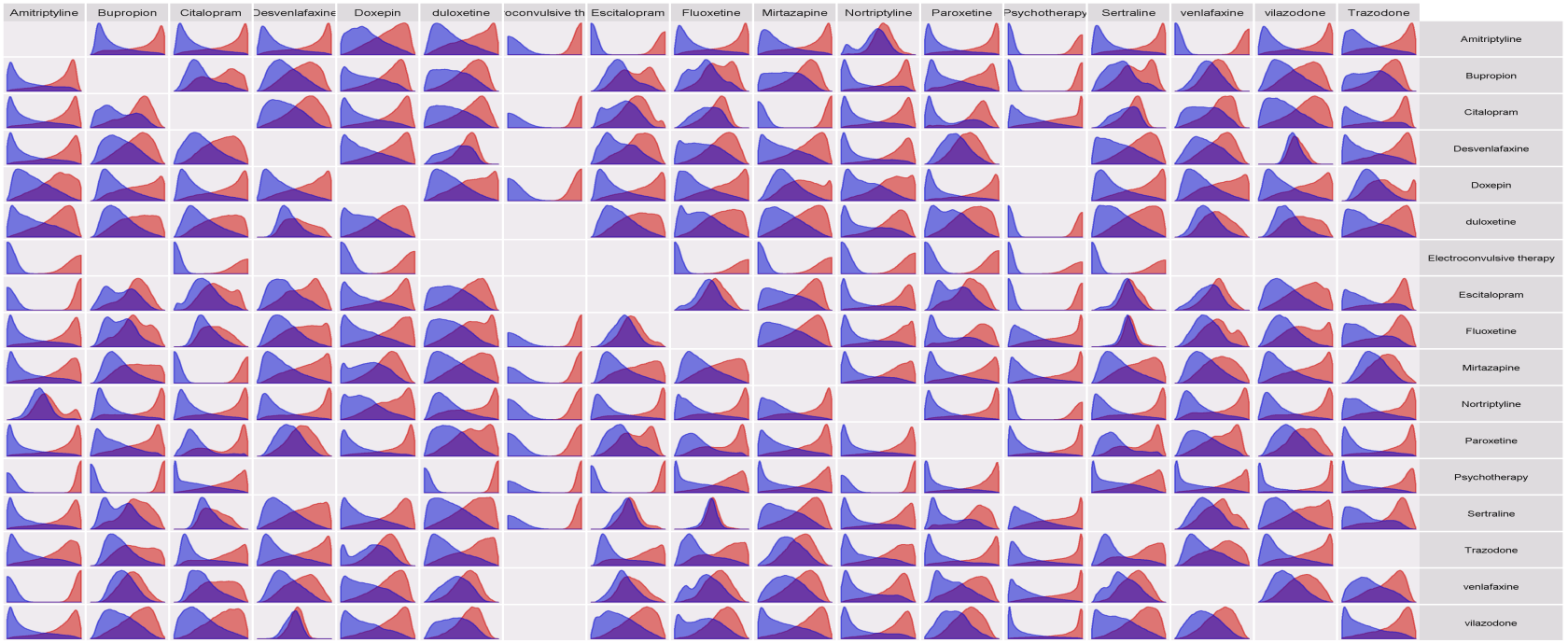


(Sharing source code)

Addressing reproducibility #5

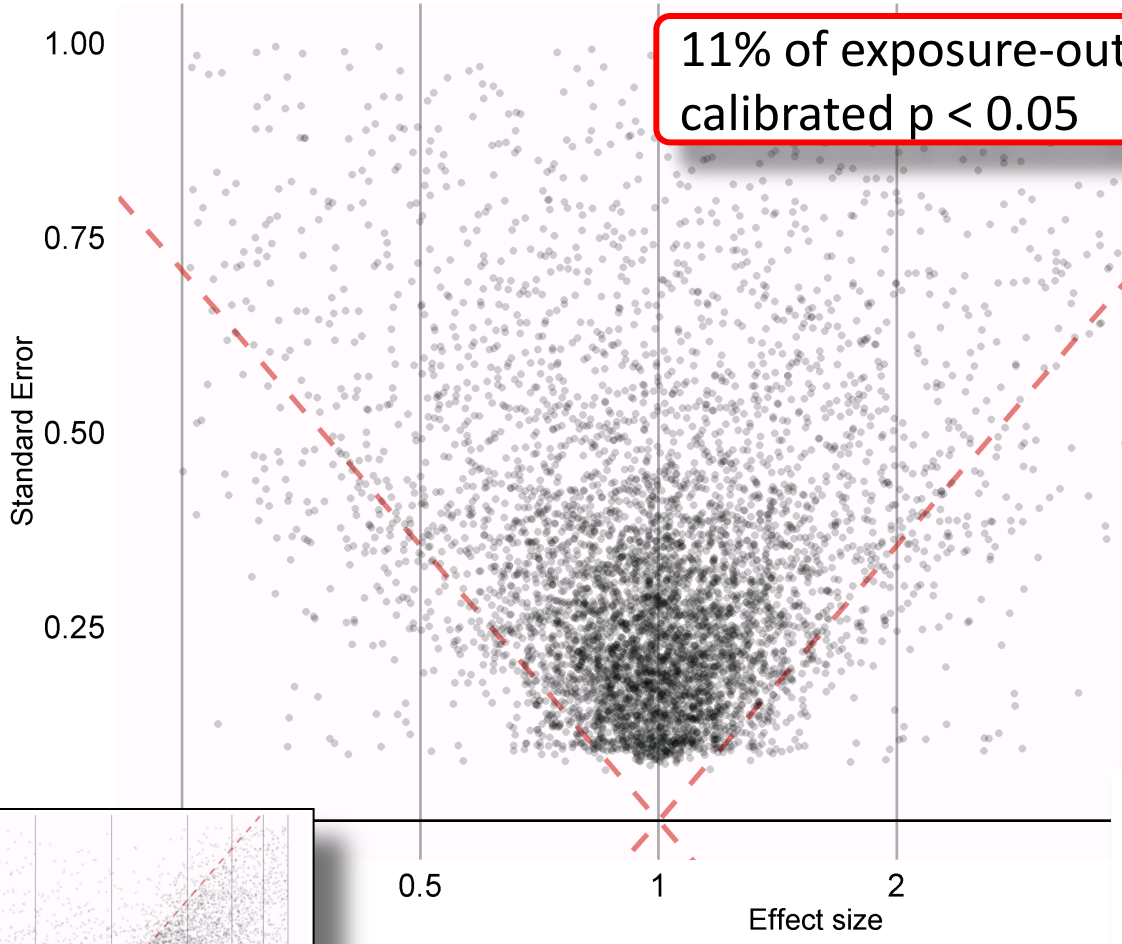
5. Carry out on aligned hypotheses at scale

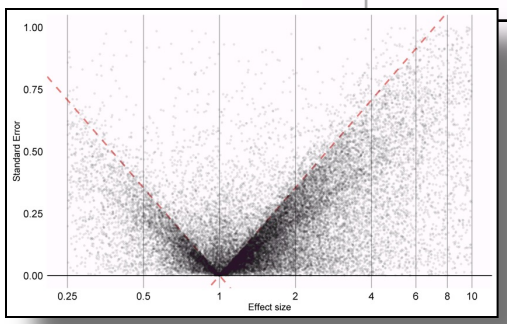
- Operating characteristics of the analysis
- Large-scale diagnostics



OHDSI results in line with expectations

11% of exposure-outcome pairs have calibrated p < 0.05





PHILOSOPHICAL TRANSACTIONS A

rsta.royalsocietypublishing.org

Research

Cite this article: Schuemie MJ, Ryan PB, Hripcsak G, Madigan D, Suchard MA. 2018 Improving reproducibility by using high-throughput observational studies with empirical calibration. *Phil. Trans. R. Soc. A* **376**: 20170356.
http://dx.doi.org/10.1098/rsta.2017.0356

Accepted: 8 May 2018

Improving reproducibility by using high-throughput observational studies with empirical calibration

Martijn J. Schuemie[1,2], Patrick B. Ryan[1,2,3], George Hripcsak[1,3,4], David Madigan[1,5] and Marc A. Suchard[1,6,7,8]

[1]Observational Health Data Sciences and Informatics (OHDSI), New York, NY 10032, USA
[2]Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08560, USA
[3]Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032, USA

Large-scale estimation

- How to use it
  - Decide what question you are asking, then correct for multiple hypotheses
- Not “data-dredging”!
  - Data-dredging is not about what you do but about what you *throw out*

OHDSI "LEGEND" Hypertension Study
Filling in the evidence gaps

Clinical Practice Guideline: Executive Summary

2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary

A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines

WRITING COMMITTEE MEMBERS

Paul K. Whelton, MB, MD, MSc, FAHA, Chair; Robert M. Carey, MD, FAHA, Vice Chair; Wilbert S. Aronow, MD, FACC, FAHA*; Donald E. Casey, Jr, MD, MPH, MBA, FAHA†; Karen J. Collins, MBA‡; Cheryl Dennison Himmelfarb, RN, ANP, PhD, FAHA§; Sondra M. DePalma, MHS, PA-C, CLS, AACCǁ; Samuel Gidding, MD, FAHA¶; Kenneth A. Jamerson, MD#; Daniel W. Jones, MD, FAHA†; Eric J. MacLaughlin, PharmD**; Paul Muntner, PhD, FAHA†; Bruce Ovbiagele, MD, MSc, MAS, MBA, FAHA†; Sidney C. Smith, Jr, MD, MACC, FAHA††; Crystal C. Spencer, JD‡; Randall S. Stafford, MD, PhD‡‡; Sandra J. Taler, MD, FAHA§§; Randal J. Thomas, MD, MS, FACC, FAHAǁǁ; Kim A. Williams, Sr, MD, MACC, FAHA†; Jeff D. Williamson, MD, MHS¶¶; Jackson T. Wright, Jr, MD, PhD, FAHA##

ACC/AHA TASK FORCE MEMBERS

Glenn N. Levine, MD, FACC, FAHA, Chair; Patrick T. O'Gara, MD, FAHA, MACC, Chair-Elect; Jonathan L. Halperin, MD, FACC, FAHA, Immediate Past Chair; Sana M. Al-Khatib, MD, MHS, FACC, FAHA; Joshua A. Beckman, MD, MS, FAHA; Kim K. Birtcher, MS, PharmD, AACC; Biykem Bozkurt, MD, PhD, FACC, FAHA***; Ralph G. Brindis, MD, MPH, MACC***; Joaquin E. Cigarroa, MD, FACC; Lesley H. Curtis, PhD, FAHA***; Anita Deswal, MD, MPH, FACC, FAHA; Lee A. Fleisher, MD, FACC, FAHA; Federico Gentile, MD, FACC; Samuel Gidding, MD, FAHA***; Zachary D. Goldberger, MD, MS, FACC, FAHA; Mark A. Hlatky, MD, FACC, FAHA; John Ikonomidis, MD, PhD, FAHA; José A. Joglar, MD, FACC, FAHA; Laura Mauri, MD, MSc, FAHA; Susan J. Pressler, PhD, RN, FAHA***; Barbara Riegel, PhD, RN, FAHA; Duminda N. Wijeysundera, MD, PhD

Table 18. Oral Antihypertensive Drugs

Class	Drug	Usual Dose, Range (mg/d)*	Daily Frequency	Comments
Primary agents				
Thiazide or thiazide-type diuretics	Chlorthalidone	12.5–25	1	• Chlorthalidone is preferred on the basis of prolonged half-life and proven trial reduction of CVD. • Monitor for hyponatremia and hypokalemia, uric acid and calcium levels. • Use with caution in patients with history of acute gout unless patient is on uric acid–lowering therapy.
	Hydrochlorothiazide	25–50	1	
	Indapamide	1.25–2.5	1	
	Metolazone	2.5–10	1	
ACE inhibitors	Benazepril	10–40	1 or 2	• Do not use in combination with ARBs or direct renin inhibitor. • There is an increased [illegible]alemia, especially in patients with [illegible] supplements • [illegible] • [illegible]
	Captopril	12.5–150	2 or 3	
	Enalapril	5–40	1 or 2	
	Fosinopril	10–40	1	
	Lisinopril	10–40	1	
	Moexipril	7.5–30	1 or 2	
	Perindopril	4–16	1	
	Quinapril	10–80	1 or 2	
	Ramipril	2.5–10	1 or 2	
	Trandolapril	1–4	1	
ARBs	Azilsartan	40–80	1	[illegible]
	Candesartan	8–32	1	[illegible]
	Eprosartan	600–800	1 or 2	• There is [illegible] or in those [illegible]
	Irbesartan	150–300	1	
	Losartan	50–100	1 or 2	• There is a [illegible] with severe bila[illegible]nal artery stenosis.
	Olmesartan	20–40	1	
	Telmisartan	20–80	1	• Do not use if patient has history of angioedema with ARBs. Patients with a history of angioedema with an ACE inhibitor can receive an ARB beginning 6 [illegible]
	Valsartan	80–320	1	
CCB—dihydropyridines	Amlodipine	2.5–10	1	
	Felodipine	5–10	1	
	Isradipine	5–10	2	
	Nicardipine SR	5–20	1	
	Nifedipine LA	60–120	1	
	Nisoldipine	30–90	1	
CCB—nondihydropyridines	Diltiazem SR	180–360	2	
	Diltiazem ER	120–480	1	
	Verapamil IR	40-80	3	
	Verapamil SR	120–480	1 or 2	
	Verapamil-delayed onset ER (various forms)	100–480	1 (in the evening)	

Only **29** different drugs in **5** different classes to choose from!

Distinguished from **28** drugs in **12** other classes that are classified as potential secondary agents **(including Beta Blockers)**

Whelton et al., Hypertension 2018

Evidence to support the guideline

- 40 randomized trials
- Most decisions are expert opinion

Comparisons of hypertension treatments

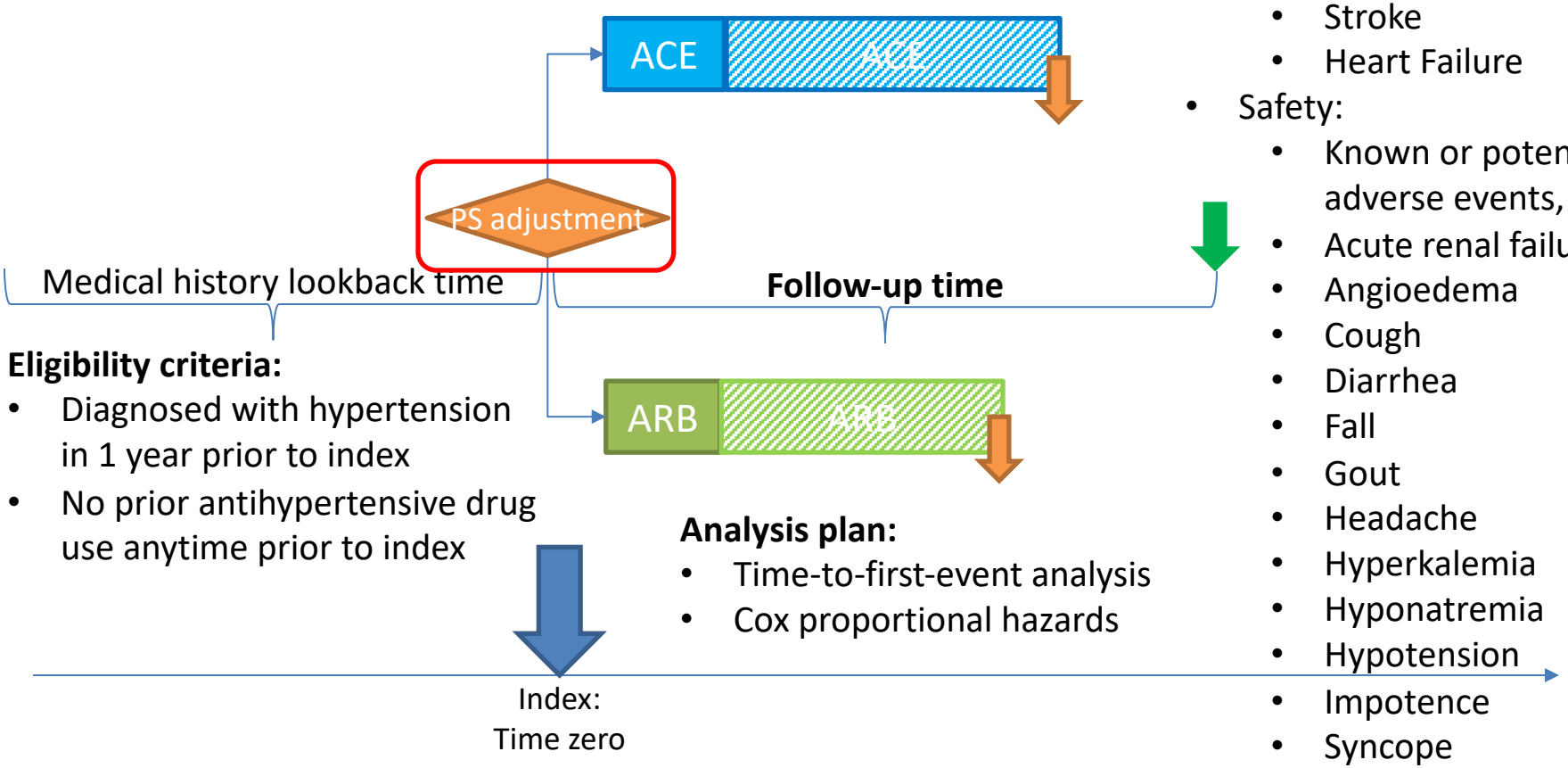
	Theoretical	Observed (n > 2,500)
Single ingredients	58	39
Single ingredient comparisons	58 * 57 = 3,306	1,296
Single drug classes	15	13
Single class comparisons	15 * 14 = 210	156
Dual ingredients	58 * 57 / 2 = 1,653	58
Single vs duo drug comparisons	58 * 1,653 = 95,874	3,810
Dual classes	15 * 14 / 2 = 105	32
Single vs duo class comparisons	15 * 105 = 1,575	832
Duo vs duo drug comparisons	1,653 * 1,652 = 2,730,756	2,784
Duo vs duo class comparisons	105 * 104 = 10,920	992
…	…	…
Total comparisons	2,843,250	10,278
Outcomes of interest	58	58
Target-comparator-outcomes	2,843,250 * 58 = 164,908,500	587,020

Observational study to compare two initial therapies

Treatment strategies:
- Monotherapy with ACE
- Monotherapy with ARB

Causal contrasts of interest:
- Intent-to-treat effect
- On-treatment effect

Outcomes:
- Efficacy:
  - Myocardial infarction
  - Stroke
  - Heart Failure
- Safety:
  - Known or potential adverse events, e.g.
  - Acute renal failure
  - Angioedema
  - Cough
  - Diarrhea
  - Fall
  - Gout
  - Headache
  - Hyperkalemia
  - Hyponatremia
  - Hypotension
  - Impotence
  - Syncope
  - Vertigo



Eligibility criteria:
- Diagnosed with hypertension in 1 year prior to index
- No prior antihypertensive drug use anytime prior to index

Analysis plan:
- Time-to-first-event analysis
- Cox proportional hazards

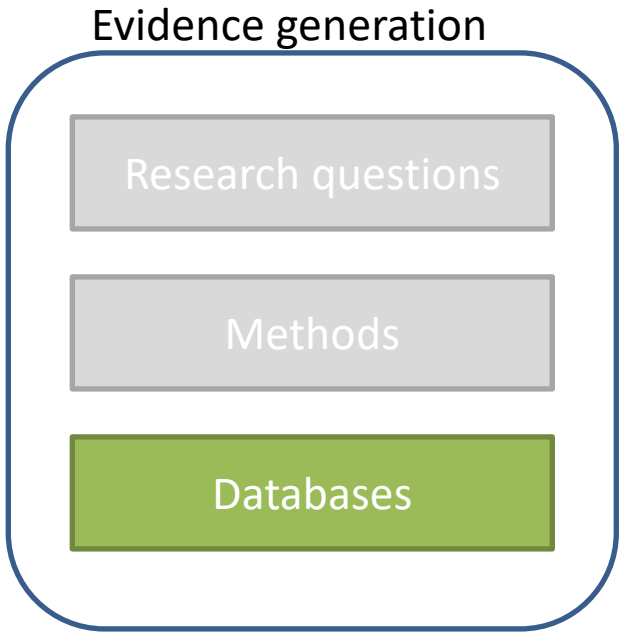
58 outcomes of interest

Abdominal pain
Abnormal weight gain
Abnormal weight loss
Acute myocardial infarction
Acute pancreatitis
Acute renal failure
All-cause mortality
Anaphylactoid reaction
Anemia
Angioedema
Anxiety
Bradycardia
Cardiac arrhythmia
Cardiovascular disease
Cardiovascular-related mortality
Chest pain or angina
Chronic kidney disease
Coronary heart disease
Cough
Decreased libido
Dementia
Depression
Diarrhea
Edema
End stage renal disease
Fall
Gastrointestinal bleeding
Gout
Headache
Heart failure
Hemorrhagic stroke
Hepatic failure
Hospitalization with heart failure
Hospitalization with preinfarction syndrome
Hyperkalemia
Hypokalemia
Hypomagnesemia
Hyponatremia
Hypotension
Impotence
Ischemic stroke
Kidney disease
Malignant neoplasm
Measured renal dysfunction
Nausea
Neutropenia or agranulocytosis
Rash
Rhabdomyolysis
Stroke
Sudden cardiac death
Syncope
Thrombocytopenia
Transient ischemic attack
Type 2 diabetes mellitus
Vasculitis
Venous thromboembolic events
Vertigo
Vomiting

76 negative controls

Abnormal cervical smear
Abnormal pupil
Abrasion and/or friction burn of trunk without infection
Absence of breast
Absent kidney
Acid reflux
Acquired hallux valgus
Acquired keratoderma
Acquired trigger finger
Acute conjunctivitis
Amputated foot
Anal and rectal polyp
Burn of forearm
Calcaneal spur
Cannabis abuse
Cervical somatic dysfunction
Changes in skin texture
Chondromalacia of patella
Cocaine abuse
Colostomy present
Complication due to Crohn's disease
Contact dermatitis
Contusion of knee
Crohn's disease
Derangement of knee
Difficulty sleeping
Disproportion of reconstructed breast
Effects of hunger
Endometriosis
Epidermoid cyst
Feces contents abnormal
Foreign body in orifice
Ganglion cyst
Genetic predisposition
Hammer toe
Hereditary thrombophilia
Herpes zoster without complication
High risk sexual behavior
Homocystinuria
Human papilloma virus infection
Ileostomy present
Impacted cerumen
Impingement syndrome of shoulder region
Ingrowing nail
Injury of knee
Irregular periods
Kwashiorkor
Late effect of contusion
Late effect of motor vehicle accident
Leukorrhea
Macular drusen
Melena
Nicotine dependence
Noise effects on inner ear
Nonspecific tuberculin test reaction
Non-toxic multinodular goiter
Onychomycosis due to dermatophyte
Opioid abuse
Passing flatus
Postviral fatigue syndrome
Presbyopia
Problem related to lifestyle
Psychalgia
Ptotic breast
Regular astigmatism
Senile hyperkeratosis
Somatic dysfunction of lumbar region
Splinter of face, without major open wound
Sprain of ankle
Strain of rotator cuff capsule
Tear film insufficiency
Tobacco dependence syndrome
Vaginitis and vulvovaginitis
Verruca vulgaris
Wrist joint pain
Wristdrop

Databases



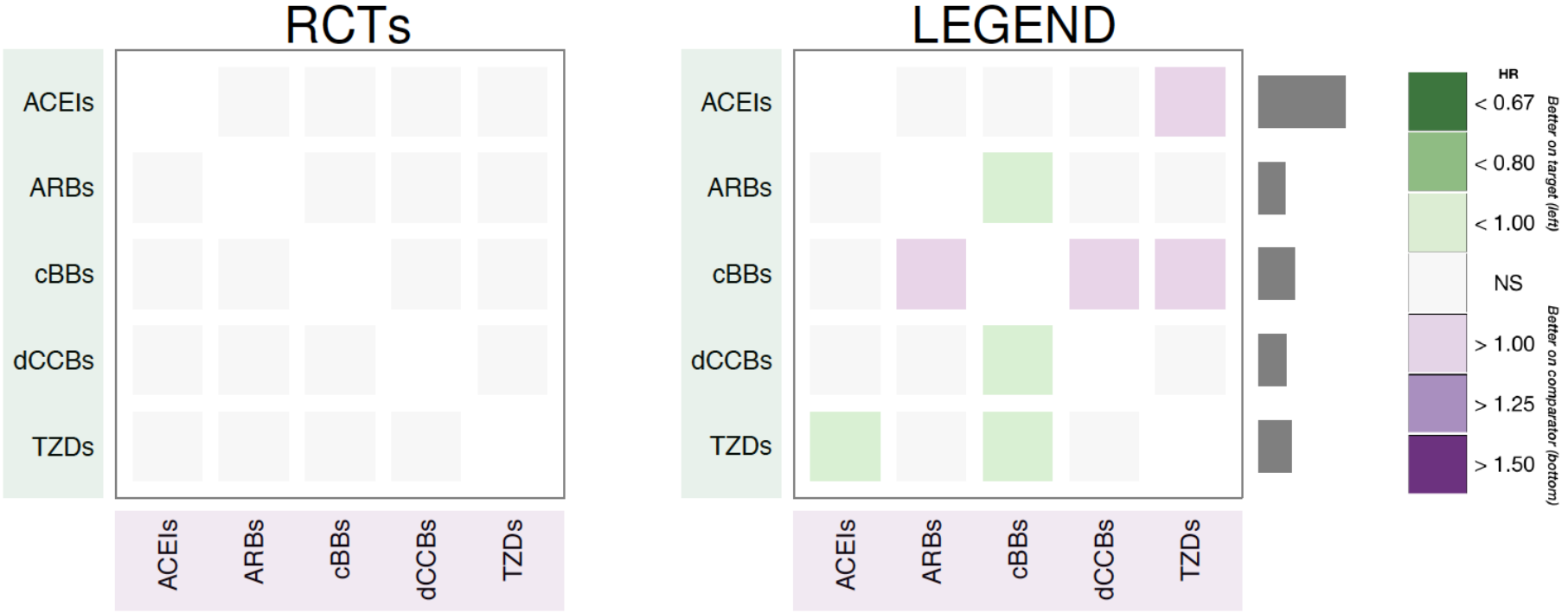
- US insurance databases
  - IBM® MarketScan® CCAE
  - IBM® MarketScan® MDCD
  - IBM® MarketScan® MDCR
  - Optum© Clinformatics®
- Japanese insurance database
  - Japan Medical Data Center
- Korean national insurance database
  - NHIS-NSC
- US EHR databases
  - Columbia University Irving Medical Center
  - Optum© PANTHER®
- German EHR database
  - QuintilesIMS Disease Analyzer (DA) Germany

Ajou University

Columbia University

First-line agents: comparisons from LEGEND

Efficacy outcome: **myocardial infarction**, heart failure, stroke

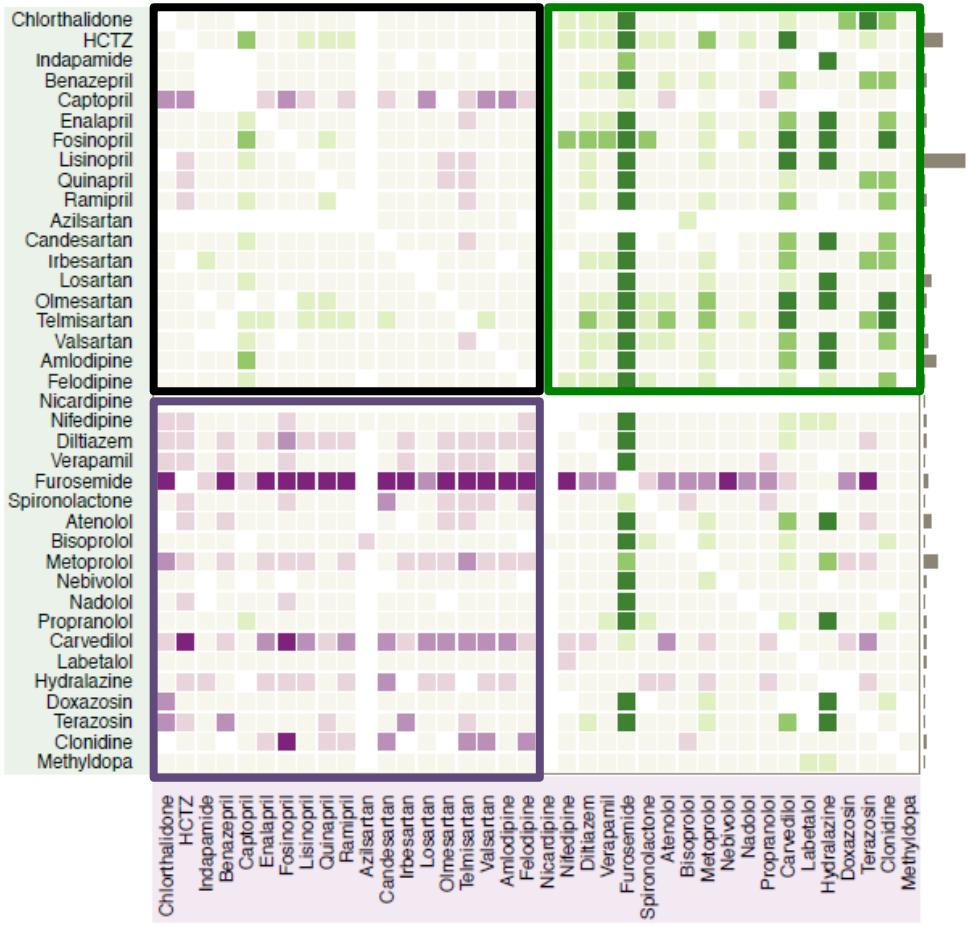


Data source: meta-analysis, $\sim 1 - 2$M total patients per study

- Beta blockers underperform alternatives
- Unexpected: TZDs $>$ ACEIs. Reliable?

Lancet 2019

Cardiovascular efficacy by drug



Composite (MI, HF, stroke) outcome in meta-analysis

Prescriptions are not written at the class-level; must choose an individual drug for the patient

- 1st-line > 2nd-line
- Some within-class differences failed diagnostics, e.g. captopril

Chlorthalidone vs hydrochlorthiazide:
worse safety without real world effectiveness

Risk estimates and meta-analysis across LEGEND databases:

Analysis	Data source	HR	LB	UB	P	Cal.HR	Cal.LB	Cal.UB	Cal.P
PS stratification, on-treatment	CCAE	0.65	0.33	1.14	0.17	0.66	0.37	1.19	0.18
PS stratification, on-treatment	Meta-analysis	0.79	0.54	1.16	0.24	0.81	0.56	1.17	0.30
PS stratification, on-treatment	Optum	0.90	0.52	1.44	0.67	0.93	0.57	1.53	0.82
PS stratification, on-treatment	Panther	0.98	0.05	5.06	0.99	0.91	0.26	3.42	0.96

Showing 1 to 4 of 4 entries Previous 1 Next

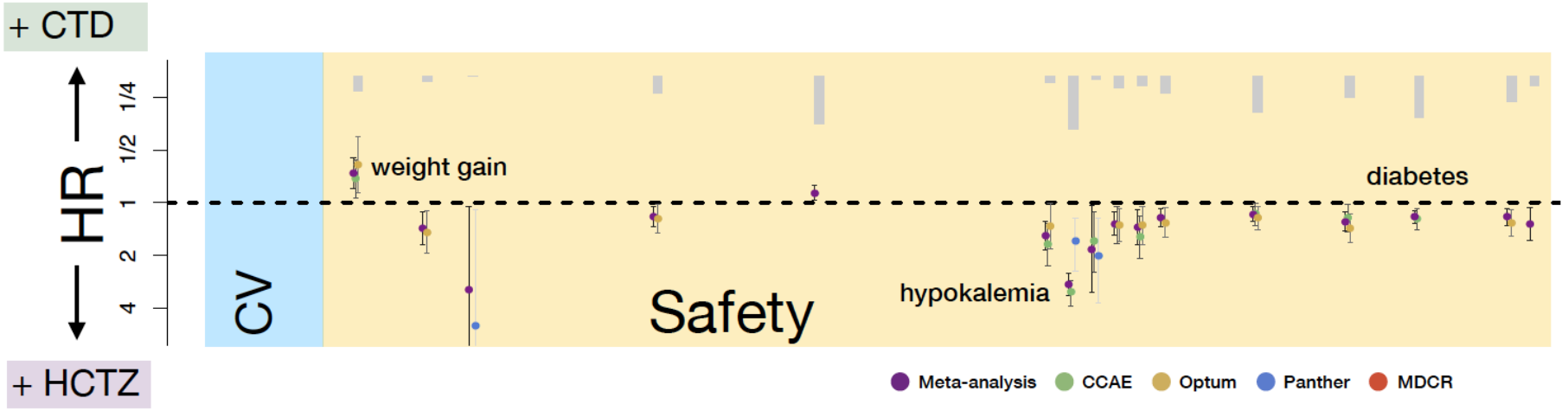
Power | Systematic error | Subgroups

Table 1a. Number of subjects, follow-up time (in years), number of outcome events, and event incidence rate (IR) per 1,000 patient years (PY) in the target (*Chlorthalidone*) and comparator (*Hydrochlorothiazide*) group after stratification, as well as the minimum detectable relative risk (MDRR). Note that the IR does not account for any stratification.

Target subjects	Comparator subjects	Target years	Comparator years	Target events	Comparator events	Target IR (per 1,000 PY)	Comparator IR (per 1,000 PY)	MDRR	i2
25,566	528,202	14,047	339,516	<32	819	<2.28	2.41	>1.58	0.00

Hripcsak JAMA IM 2020

CTD vs. HCTZ: safety profile



- Safety favors HCTZ – electrolyte imbalance
- CTD is more potent, longer half-life

HCTZ vs chlorthalidone

- Physiology
  - Chlorthalidone is more potent and longer lasting
- Indirect (network) meta-analysis favors chlorthalidone
  - Combine RCT results
  - Bias: heterogeneity of treatment effect + different populations
  - Also: differential RCT design and execution
- Recent observational research favors HCTZ
- VA Diuretic Comparison Project
  - RCT with completion 2023
  - Different question: of patients tolerating HCTZ, should they switch to chlorthalidone
- Response
  - 50% failure off diuretic (chlorthalidone is faster) in 2 months
  - Time at risk too short (but 25% are long term); do have sufficient power
  - Anecdotes
- Chlorthalidone is a potent drug

Choice of drug therapy in primary (essential) hypertension – **UpToDate**

By contrast, other observational studies suggest that chlorthalidone and hydrochlorothiazide lead to similar rates of cardiovascular events but that chlorthalidone increases the risk of adverse metabolic effects [35,36].

Based upon the above observations, we and other experts suggest that thiazide-like diuretics (such as chlorthalidone…

ACEi versus ARB

On the other hand, once the field is confident in a result…

Hypertension

ORIGINAL ARTICLE

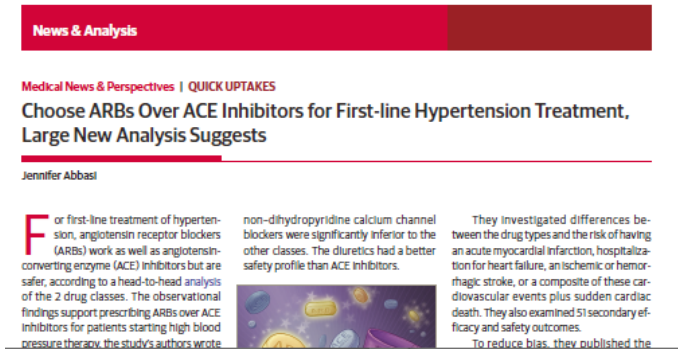
Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers

A Multinational Cohort Study

RuiJun Chen, Marc A. Suchard, Harlan M. Krumholz, Martijn J. Schuemie, Steven Shea, Jon Duke, Nicole Pratt, Christian G. Reich, David Madigan, Seng Chan You, Patrick B. Ryan, George Hripcsak

ABSTRACT: ACE (angiotensin-converting enzyme) inhibitors and angiotensin receptor blockers (ARBs) are equally guideline-recommended first-line treatments for hypertension, yet few head-to-head studies exist. We compared the real-world effectiveness and safety of ACE inhibitors versus ARBs in the first-line tre retrospective, new-user comparative cohort design to estimate hazard rat confounding and bias, specifically large-scale propensity score adjustment, en

Circulation
Volume 145, Issue 6, 8 February 2022; Pages 413-415
https://doi.org/10.1161/CIRCULATIONAHA.121.057835



News & Analysis

Medical News & Perspectives | QUICK UPTAKES

Choose ARBs Over ACE Inhibitors for First-line Hypertension Treatment, Large New Analysis Suggests

Jennifer Abbasi

For first-line treatment of hypertension, angiotensin receptor blockers (ARBs) work as well as angiotensin-converting enzyme (ACE) inhibitors but are safer, according to a head-to-head analysis of the 2 drug classes. The observational findings support prescribing ARBs over ACE inhibitors for patients starting high blood pressure therapy, the study's authors wrote

non-dihydropyridine calcium channel blockers were significantly inferior to the other classes. The diuretics had a better safety profile than ACE inhibitors.

They investigated differences between the drug types and the risk of having an acute myocardial infarction, hospitalization for heart failure, an ischemic or hemorrhagic stroke, or a composite of these cardiovascular events plus sudden cardiac death. They also examined 51 secondary efficacy and safety outcomes.

To reduce bias, they published the

American Heart Association.

ON MY MIND

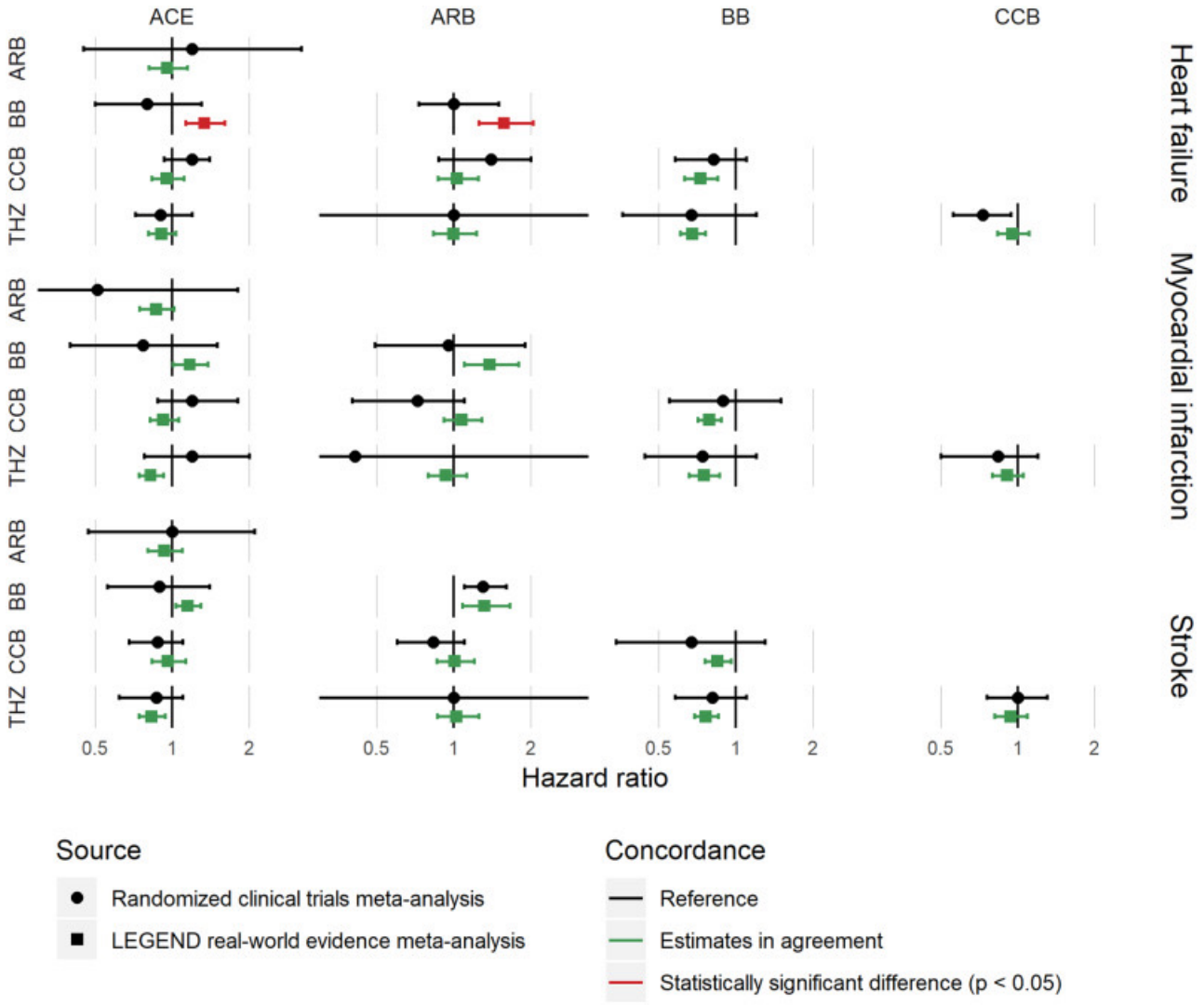
Why Are We Still Prescribing Angiotensin-Converting Enzyme Inhibitors?

Franz H. Messerli, MD , Chirag Bavishi, MD, MPH, and Sripal Bangalore, MD, MHA

The human understanding when it has once adopted an opinion (either as being the received opinion or as being agreeable to itself) draws all things else to support and agree with it.

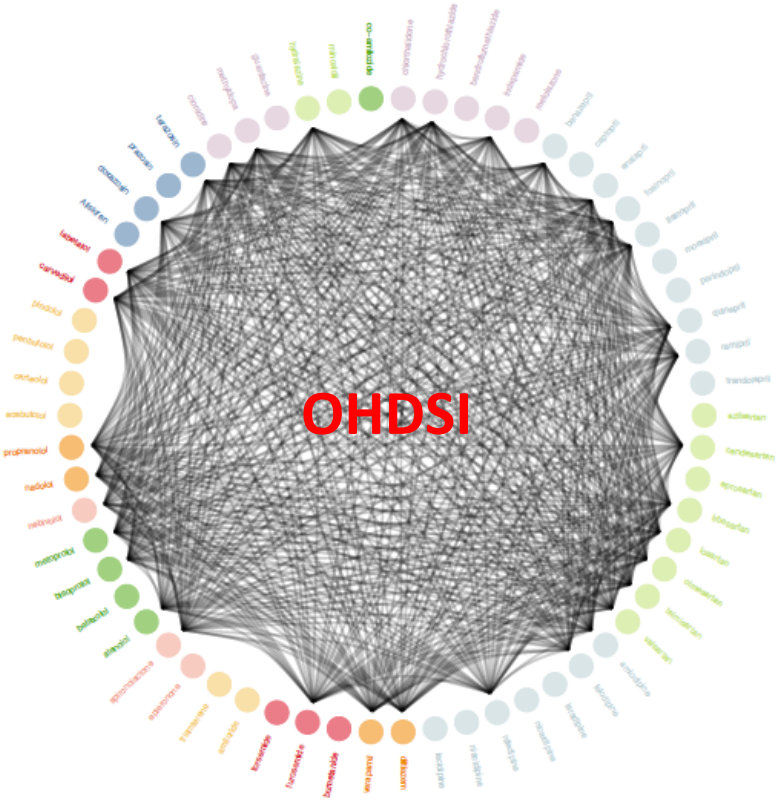
—Sir Francis Bacon, 1620

LEGEND vs RCTs: 28/30



OHDSI “LEGEND” Hypertension Study
Filling in the evidence gaps

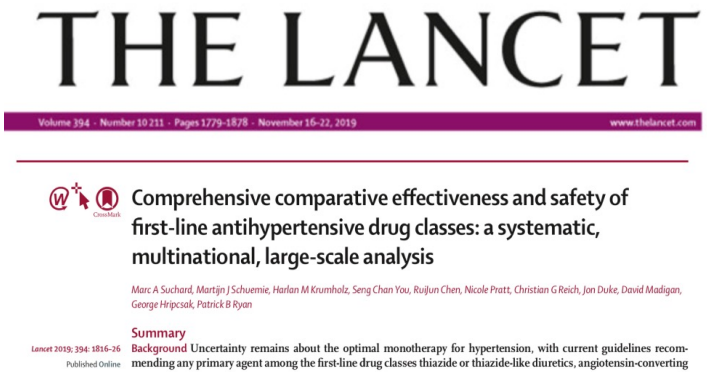
Head-to-head HTN drug comparisons



- Trials: 40
- $N = 102 - [1148] - 33K$

- Comparisons: $10,278$
- $N = 3502 - [212K] - 1.9M$

LEGEND Hypertension



THE LANCET

Volume 394 · Number 10211 · Pages 1779-1878 · November 16-22, 2019 www.thelancet.com

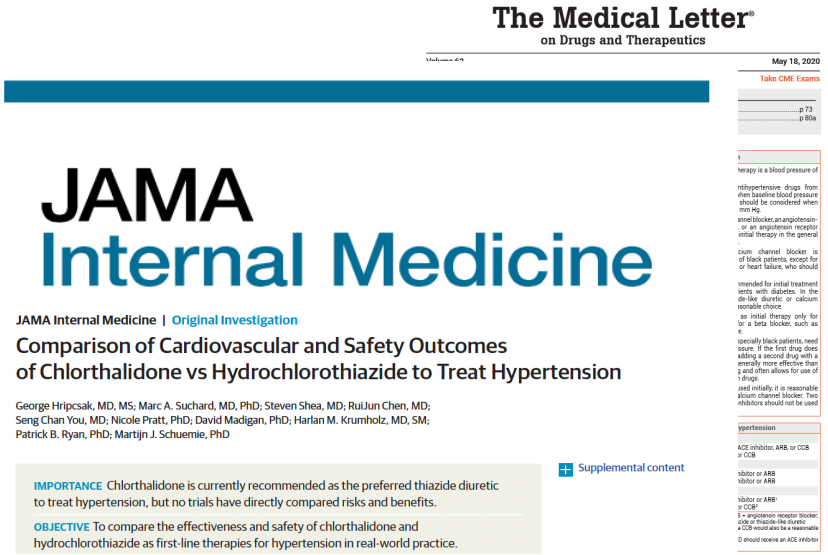
Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis

Marc A Suchard, Martijn J Schuemie, Harlan M Krumholz, Seng Chan You, RuiJun Chen, Nicole Pratt, Christian G Reich, Jon Duke, David Madigan, George Hripcsak, Patrick B Ryan

Summary

Lancet 2019; 394: 1816–26 Published Online

Background Uncertainty remains about the optimal monotherapy for hypertension, with current guidelines recommending any primary agent among the first-line drug classes thiazide or thiazide-like diuretics, angiotensin-converting



The Medical Letter on Drugs and Therapeutics

May 18, 2020

Take CME Exams

JAMA Internal Medicine

JAMA Internal Medicine | Original Investigation

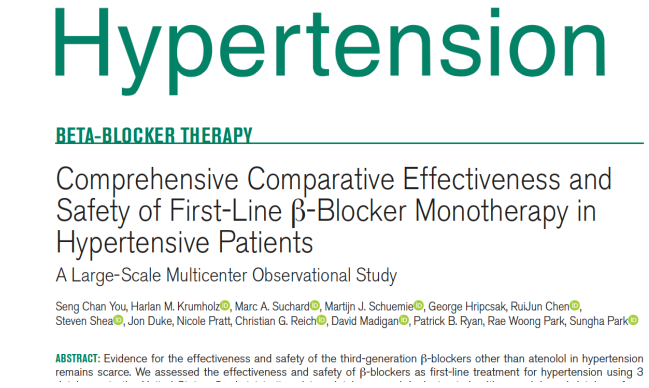
Comparison of Cardiovascular and Safety Outcomes of Chlorthalidone vs Hydrochlorothiazide to Treat Hypertension

George Hripcsak, MD, MS; Marc A. Suchard, MD, PhD; Steven Shea, MD; RuiJun Chen, MD; Seng Chan You, MD; Nicole Pratt, PhD; David Madigan, PhD; Harlan M. Krumholz, MD, SM; Patrick B. Ryan, PhD; Martijn J. Schuemie, PhD

Supplemental content

IMPORTANCE Chlorthalidone is currently recommended as the preferred thiazide diuretic to treat hypertension, but no trials have directly compared risks and benefits.

OBJECTIVE To compare the effectiveness and safety of chlorthalidone and hydrochlorothiazide as first-line therapies for hypertension in real-world practice.



Hypertension

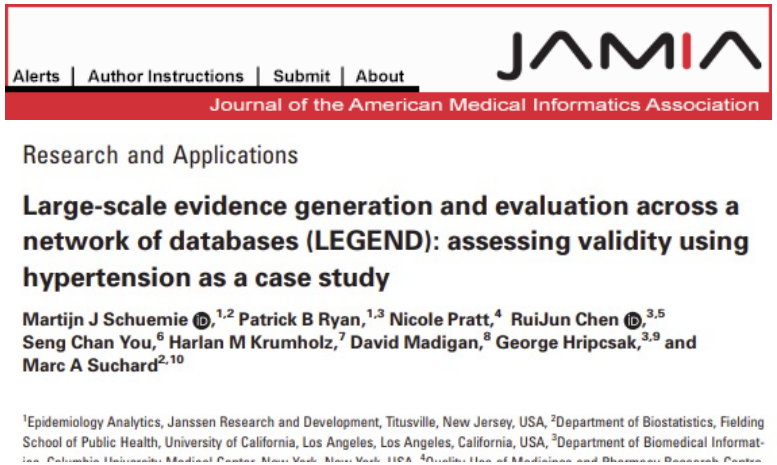
BETA-BLOCKER THERAPY

Comprehensive Comparative Effectiveness and Safety of First-Line β-Blocker Monotherapy in Hypertensive Patients

A Large-Scale Multicenter Observational Study

Seng Chan You, Harlan M. Krumholz, Marc A. Suchard, Martijn J. Schuemie, George Hripcsak, RuiJun Chen, Steven Shea, Jon Duke, Nicole Pratt, Christian G. Reich, David Madigan, Patrick B. Ryan, Rae Woong Park, Sungha Park

ABSTRACT: Evidence for the effectiveness and safety of the third-generation β-blockers other than atenolol in hypertension remains scarce. We assessed the effectiveness and safety of β-blockers as first-line treatment for hypertension using 3



Alerts | Author Instructions | Submit | About

JAMIA

Journal of the American Medical Informatics Association

Research and Applications

Large-scale evidence generation and evaluation across a network of databases (LEGEND): assessing validity using hypertension as a case study

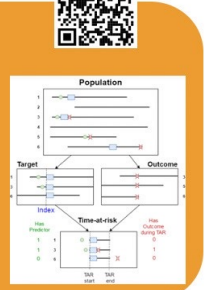
Martijn J Schuemie [1,2] Patrick B Ryan,[1,3] Nicole Pratt,[4] RuiJun Chen [3,5] Seng Chan You,[6] Harlan M Krumholz,[7] David Madigan,[8] George Hripcsak,[3,9] and Marc A Suchard[2,10]

[1]Epidemiology Analytics, Janssen Research and Development, Titusville, New Jersey, USA, [2]Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California, USA, [3]Department of Biomedical Informat-

OHDSI predictive algorithm validation framework

Clear specification of the prediction task:

- **Target Population: patients at risk**
- **Outcome: medical event to predict**
- **Time-at-risk (TAR): interval to predict if outcome will occur**

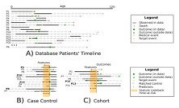


The patient-level prediction journey is more than just classification…

Join the PLP journey
PLP GitHub:
github.com/OHDSI/PatientLevelPrediction

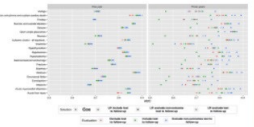
Design and Extraction

Study design



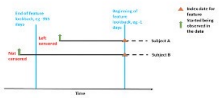
Case-control prone to misclassification and should be avoided; use cohort design

Loss to follow-up



Excluding non-outcomes lost to follow-up can bias the data

Feature extraction



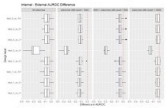
Feature lookback can make an impact on model performance if it is too short (<180 days)

Model Development

Learning across datasets

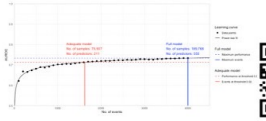
Models can be learned across datasets while maintaining privacy

Test/Train split



The design used to pick hyper-parameters and estimate internal validation matters, even with big p and big n data.

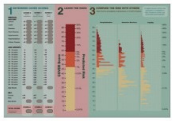
Sample size



Learning curves provide a way for model developers to determine whether they have sufficiently sized data

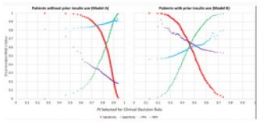
Model Evaluation

Model usability



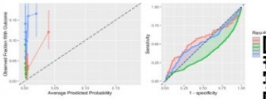
Simple score-based models are easier to apply and can be benchmarked against large-scale models

Visualizing performance



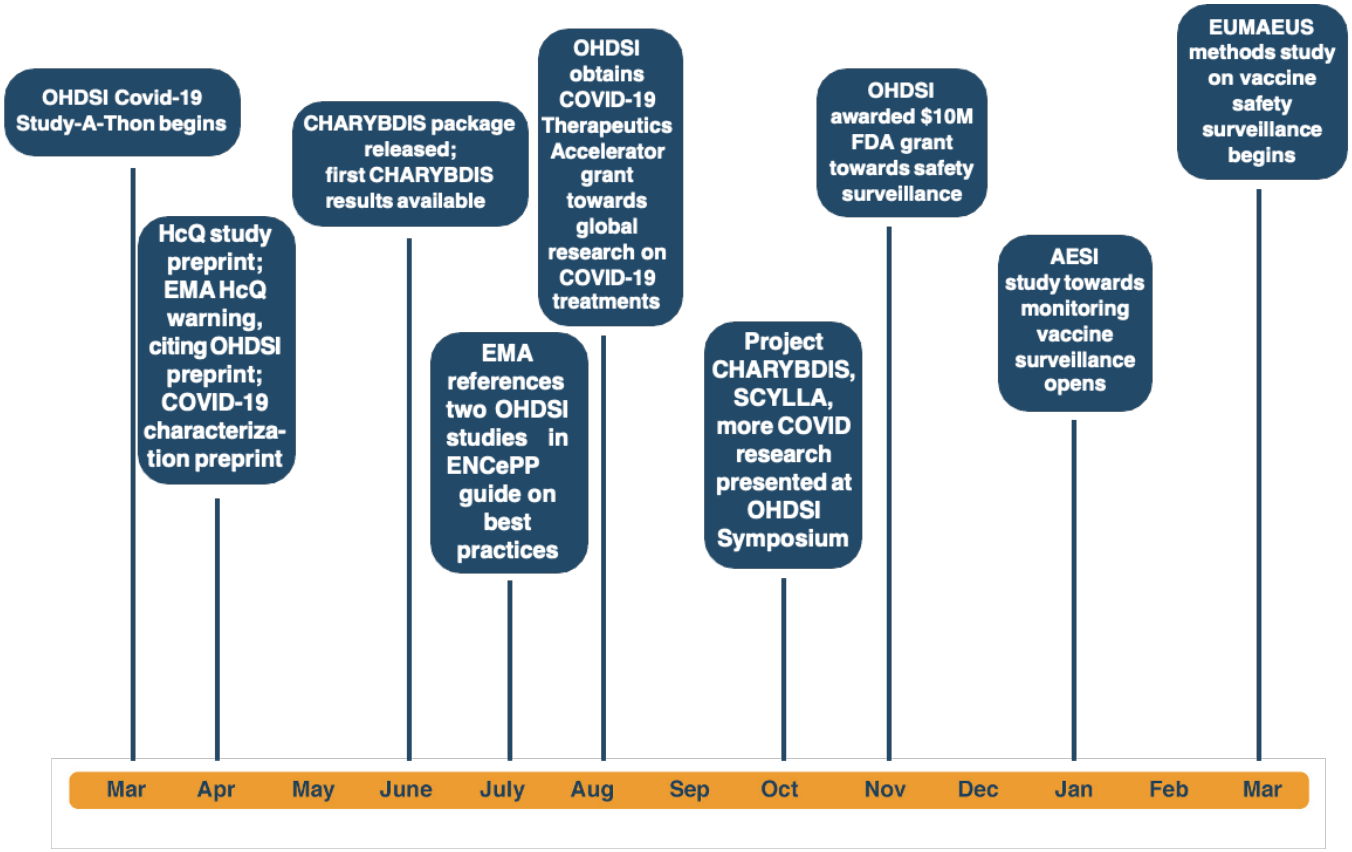
A simple plot with the operating characteristics for all cut-offs informs model usefulness

Network validation



The OHDSI network enables large-scale external validation and improves our understanding of models

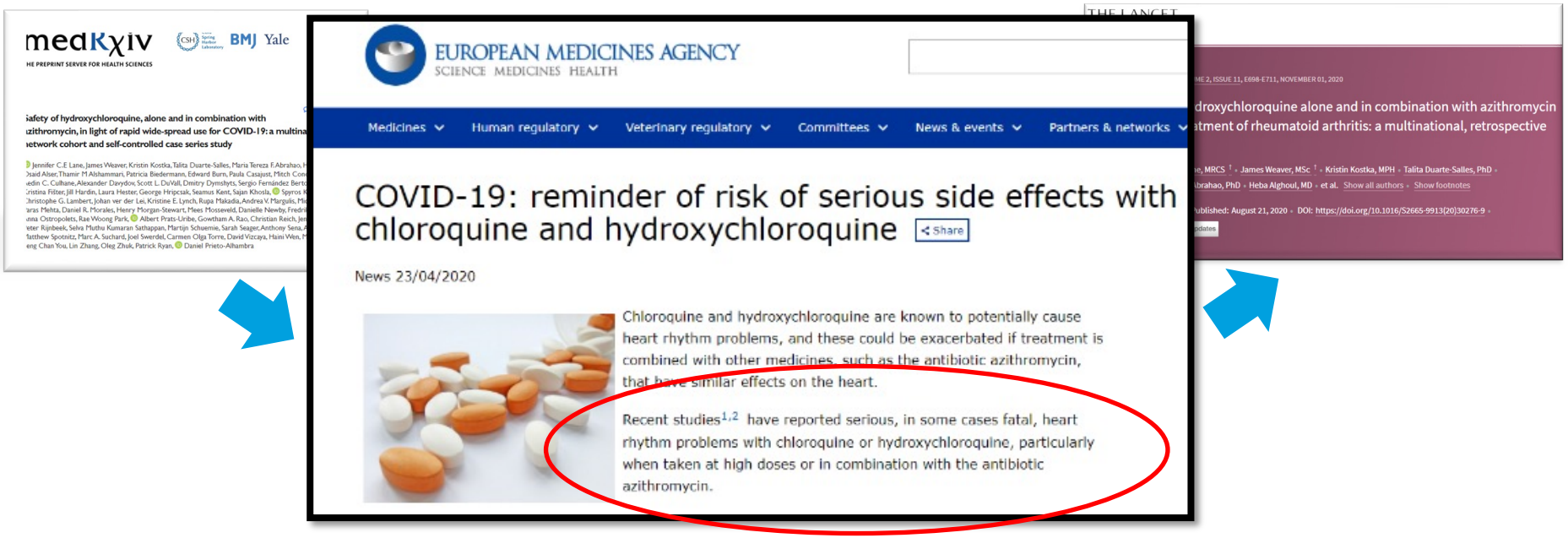
COVID-19


OHDSI Covid-19 Study-A-Thon begins
HcQ study preprint; EMA HcQ warning, citing OHDSI preprint; COVID-19 characteriza-tion preprint
CHARYBDIS package released; first CHARYBDIS results available
EMA references two OHDSI studies in ENCePP guide on best practices
OHDSI obtains COVID-19 Therapeutics Accelerator grant towards global research on COVID-19 treatments
Project CHARYBDIS, SCYLLA, more COVID research presented at OHDSI Symposium
OHDSI awarded $10M FDA grant towards safety surveillance
AESI study towards monitoring vaccine surveillance opens
EUMAEUS methods study on vaccine safety surveillance begins
Mar
Apr
May
June
July
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar

Safety of hydroxychloroquine

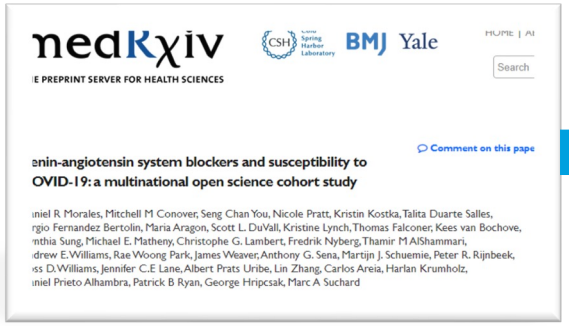
- Evidence was needed around the use of hydroxychloroquine (HCQ) alone and in combination with azithromycin (AZ). We examined the use of these drugs in rheumatoid arthritis (RA) patients.

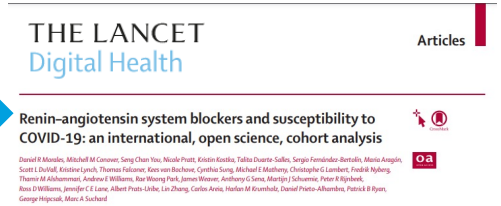
- Findings:
  - In history use in RA population, HCQ alone is generally safe but in combination with AZ it shows a doubling of risk of 30-day cardiovascular mortality.



ACE Inhibitors and susceptibility to COVID-19

- Patients with cardiovascular diseases and hypertension treated with angiotensin converting enzyme inhibitors (ACEs) angiotensin-II receptor blockers (ARBs) may influence susceptibility to COVID-19 and worsen its severity.

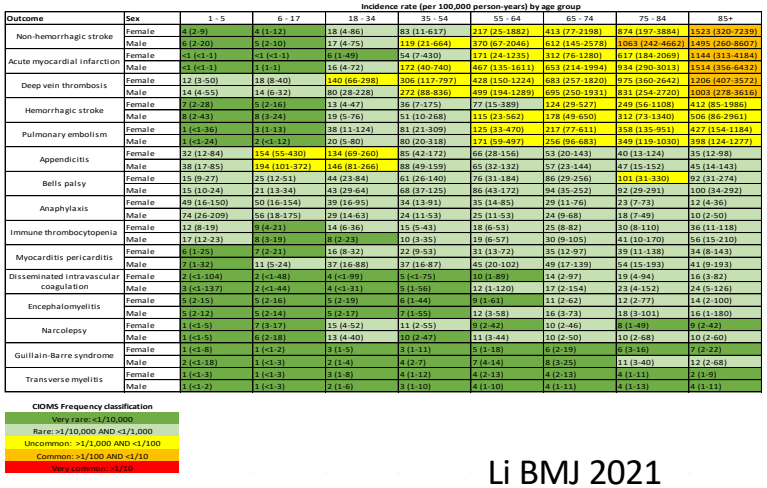




As stated by Watson et al.in relation to one of the published studies, lack of transparency and uncertainties about research standards applied raise doubts about published results. Morales et al. supported the reproducibility of their study by publishing the study protocol in the EU PAS Register ahead of time, providing a start-to-finish executable code, facilitating the sharing and exploration of the complete result set with an interactive web application and asking clinicians and epidemiologists to perform a blinded evaluation of propensity score diagnostics for the treatment comparisons.

COVID-19 Vaccine Safety Methods Research

- AstraZeneca vaccine
  - March 11-15, 2021 – 13 European countries suspend use for fears of blood clots
    - Denmark, Norway, Iceland, Bulgaria, Ireland, Netherlands, Spain, Germany, Italy, France, …
  - March 18, 2021 – EMA determines benefits outweigh the risks
    - Thromboembolic events "lower than that **expected in the general population**"
    - DIC and CVST above baseline but very rare
    - "The number of reported events **exceeds those expected**, and causality although not confirmed, cannot therefore be excluded. However, given the rarity of the events, and the **difficulty of establishing baseline incidence** since COVID-19 itself is resulting in hospitalisations with thromboembolic complications, the strength of any association is uncertain."
- Partnered with FDA Center for Biologics Evaluation and Research (CBER)
  - Vaccine safety methods research, network and local studies



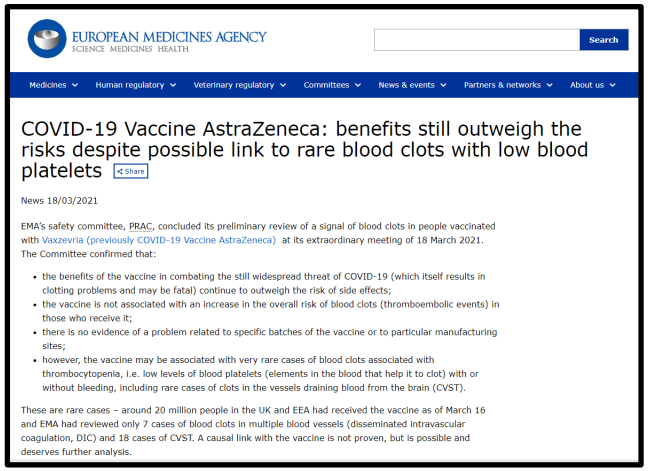
Incidence rate (per 100,000 person-years) by age group

Outcome	Sex	1 - 5	6 - 17	18 - 34	35 - 54	55 - 64	65 - 74	75 - 84	85+
Non-hemorrhagic stroke	Female	4 (2-9)	4 (1-12)	18 (4-86)	83 (11-617)	217 (25-1882)	413 (77-2198)	874 (197-3884)	1523 (320-7239)
	Male	6 (2-20)	5 (2-10)	17 (4-75)	119 (21-664)	370 (67-2046)	612 (145-2578)	1063 (242-4662)	1495 (260-8607)
Acute myocardial infarction	Female	<1 (<1-1)	<1 (<1-1)	6 (1-49)	54 (7-430)	171 (24-1235)	312 (76-1280)	617 (184-2069)	1144 (313-4184)
	Male	<1 (<1-1)	1 (1-1)	16 (4-72)	172 (40-740)	467 (135-1611)	653 (214-1994)	934 (290-3013)	1514 (356-6432)
Deep vein thrombosis	Female	12 (3-50)	18 (8-40)	140 (66-298)	306 (117-797)	428 (150-1224)	683 (257-1820)	975 (360-2642)	1206 (407-3572)
	Male	14 (4-55)	14 (6-32)	80 (28-228)	272 (88-836)	499 (194-1289)	695 (250-1931)	831 (254-2720)	1003 (278-3616)
Hemorrhagic stroke	Female	7 (2-28)	5 (2-16)	13 (4-47)	36 (7-175)	77 (15-389)	124 (29-527)	249 (56-1108)	412 (85-1986)
	Male	8 (2-43)	8 (3-24)	19 (5-76)	51 (10-268)	115 (23-562)	178 (49-650)	312 (73-1340)	506 (86-2961)
Pulmonary embolism	Female	1 (<1-36)	3 (1-13)	38 (11-124)	81 (21-309)	125 (33-470)	217 (77-611)	358 (135-951)	427 (154-1184)
	Male	1 (<1-24)	2 (<1-12)	25 (5-80)	80 (20-318)	171 (59-497)	256 (96-683)	349 (119-1030)	398 (124-1277)
Appendicitis	Female	32 (12-84)	154 (55-430)	134 (69-260)	85 (42-172)	66 (28-156)	53 (20-143)	40 (13-124)	35 (12-98)
	Male	38 (17-85)	194 (101-372)	146 (81-266)	88 (49-159)	65 (32-132)	57 (23-144)	47 (15-152)	45 (14-143)
Bells palsy	Female	15 (9-27)	25 (12-51)	44 (23-84)	61 (26-140)	76 (31-184)	86 (29-256)	101 (31-330)	92 (31-274)
	Male	15 (10-24)	21 (13-34)	43 (29-64)	68 (37-125)	86 (43-172)	94 (35-252)	92 (29-291)	100 (34-292)
Anaphylaxis	Female	49 (16-150)	50 (16-154)	39 (16-95)	34 (13-91)	35 (14-85)	29 (11-76)	23 (7-73)	12 (4-36)
	Male	74 (26-209)	56 (18-175)	29 (14-63)	24 (11-53)	25 (11-53)	24 (9-68)	18 (7-49)	10 (2-50)
Immune thrombocytopenia	Female	12 (8-19)	9 (4-21)	14 (6-36)	15 (5-43)	18 (6-53)	25 (8-82)	30 (8-110)	36 (11-118)
	Male	17 (12-23)	8 (3-19)	8 (2-23)	10 (3-35)	19 (6-57)	30 (9-105)	41 (10-170)	56 (15-210)
Myocarditis pericarditis	Female	6 (1-25)	7 (2-21)	16 (8-32)	22 (9-53)	31 (13-72)	35 (12-97)	39 (11-138)	34 (8-143)
	Male	7 (1-32)	11 (5-24)	37 (16-88)	37 (16-87)	45 (20-102)	49 (17-139)	54 (15-193)	41 (9-193)
Disseminated intravascular coagulation	Female	2 (<1-104)	2 (<1-48)	4 (<1-99)	5 (<1-75)	10 (1-89)	14 (2-97)	19 (4-94)	16 (3-82)
	Male	3 (<1-137)	2 (<1-44)	4 (<1-31)	5 (1-56)	12 (1-120)	17 (2-154)	23 (4-152)	24 (5-126)
Encephalomyelitis	Female	5 (2-15)	5 (2-16)	5 (2-19)	6 (1-44)	9 (1-61)	11 (2-62)	12 (2-77)	14 (2-100)
	Male	5 (2-12)	5 (2-14)	5 (2-17)	7 (1-55)	12 (3-58)	16 (3-73)	18 (3-101)	16 (1-180)
Narcolepsy	Female	1 (<1-5)	7 (3-17)	15 (4-52)	11 (2-55)	9 (2-42)	10 (2-46)	8 (1-49)	9 (2-42)
	Male	1 (<1-5)	6 (2-18)	13 (4-40)	10 (2-47)	11 (3-44)	10 (2-50)	10 (2-68)	10 (2-60)
Guillain-Barre syndrome	Female	1 (<1-8)	1 (<1-2)	3 (1-5)	3 (1-11)	5 (1-18)	6 (2-19)	6 (3-16)	7 (2-22)
	Male	2 (<1-18)	1 (<1-3)	2 (1-4)	4 (2-7)	7 (4-14)	8 (3-25)	11 (3-40)	12 (2-68)
Transverse myelitis	Female	1 (<1-3)	1 (<1-3)	3 (1-8)	4 (1-12)	4 (2-13)	4 (2-13)	4 (1-11)	2 (1-9)
	Male	1 (<1-2)	1 (<1-3)	2 (1-6)	3 (1-10)	4 (1-10)	4 (1-11)	4 (1-13)	4 (1-11)

CIOMS Frequency classification

Very rare: <1/10,000
Rare: >1/10,000 AND <1/1,000
Uncommon: >1/1,000 AND <1/100
Common: >1/100 AND <1/10
Very common: >1/10

Li BMJ 2021



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Search

Medicines | Human regulatory | Veterinary regulatory | Committees | News & events | Partners & networks | About us

COVID-19 Vaccine AstraZeneca: benefits still outweigh the risks despite possible link to rare blood clots with low blood platelets

Share

News 18/03/2021

EMA's safety committee, PRAC, concluded its preliminary review of a signal of blood clots in people vaccinated with Vaxzevria (previously COVID-19 Vaccine AstraZeneca) at its extraordinary meeting of 18 March 2021. The Committee confirmed that:

- the benefits of the vaccine in combating the still widespread threat of COVID-19 (which itself results in clotting problems and may be fatal) continue to outweigh the risk of side effects;
- the vaccine is not associated with an increase in the overall risk of blood clots (thromboembolic events) in those who receive it;
- there is no evidence of a problem related to specific batches of the vaccine or to particular manufacturing sites;
- however, the vaccine may be associated with very rare cases of blood clots associated with thrombocytopenia, i.e. low levels of blood platelets (elements in the blood that help it to clot) with or without bleeding, including rare cases of clots in the vessels draining blood from the brain (CVST).

These are rare cases – around 20 million people in the UK and EEA had received the vaccine as of March 16 and EMA had reviewed only 7 cases of blood clots in multiple blood vessels (disseminated intravascular coagulation, DIC) and 18 cases of CVST. A causal link with the vaccine is not proven, but is possible and deserves further analysis.

Standards enabling evidence for policy: COVID-19 treatment utilization patterns

thebmj

RESEARCH

Use of repurposed and adjuvant drugs in hospital patients with covid-19: multinational network cohort study

Albert Prats-Uribe,[1] Anthony G Sena,[2,3] Lana Yin Hui Lai, Heba Alghoul,[7] Osaid Alser,[8] Thamir M Alshammari,[9] Ca Paula Casajust,[12] Dalia Dawoud,[13,14] Asieh Golozar,[15,16] Paras P Mehta,[18] Mengchun Gong,[19] Daniel R Morales,[2] Martina Recalde,[24,25] Elena Roel,[24,25] Karishma Shah,[5] N Vignesh Subbian,[26] David Vizcaya,[27] Lin Zhang,[28,29] Ying Jaehyeong Cho,[31] Kristine E Lynch,[32] Michael E Matheny, Peter R Rijnbeek,[3] George Hripcsak,[36] Jennifer CE Lane,[5] Marc A Suchard,[38] Talita Duarte-Salles,[24] Kristin Kostka,[3] Daniel Prieto-Alhambra[1]

For numbered affiliations see end of the article.

Correspondence to: P B Ryan ryan@ohdsi.org (ORCID 0000-0002-9727-2138)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2021;373:n1038 http://dx.doi.org/10.1136/bmj.n1038

Accepted: 16 April 2021

ABSTRACT

OBJECTIVE

To investigate the use of repurposed and adjuvant drugs in patients admitted to hospital with covid-19 across three continents.

DESIGN

Multinational network cohort study.

SETTING

Hospital electronic health records from the United States, Spain, and China, and nationwide claims data from South Korea.

PARTICIPANTS

303 264 patients admitted to hospital with covid-19 from January 2020 to December 2020.

MAIN OUTCOME MEASURES

Prescriptions or dispensations of any drug on or 30 days after the date of hospital admission for covid-19.

RESULTS

Of the 303 264 patients included, 290 131 were from the US, 7599 from South Korea, 5230 from Spain, and 304 from China. 3455 drugs were identified. Common repurposed drugs were hydroxychloroquine (used in from <5 (<2%) patients in China to 2165 (85.1%)

in Spain 1473 (5 ritonavi (34.9%) and um and 238 varied g being e vitamin use incr decline the rest corticos

CONCLU Multiple the covi and ter azithror China o drugs. A antagor adjunctive treatments. Research is needed on the comparative risk and benefit of these treatments in the management of covid-19.

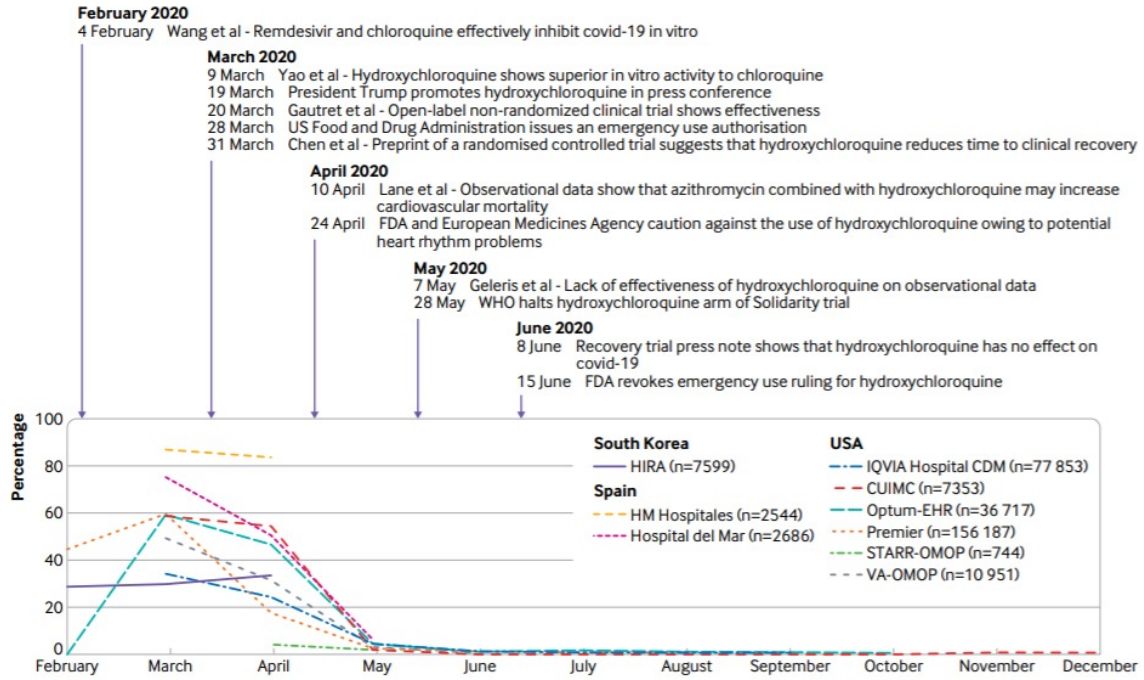


Fig 4 | Time trends in hydroxychloroquine use on days 0 to 30 after hospital admission in patients with a positive test result for or diagnosis of covid-19 by month. CUIMC=Columbia University Irving Medical Center; HIRA=Health Insurance Review and Assessment; OMOP=Observational Medical Outcomes Partnership; Optum-EHR=Optum deidentified electronic health record dataset; STARR=STAnford medicine Research data Repository; VA=Veterans Affairs

COVID-19

ARTICLE

https://doi.org/10.1038/s41467-020-18849-z OPEN

Check for updates

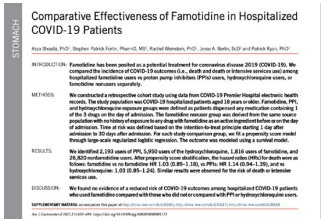
Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an inte[illegible] network study


International Journal of Obesity

ARTICLE OPEN

Epidemiology and Population Health

Characteristics and outcomes of 627 044 COVID-19 patients living with and without obesity in the United States, Spain, and the United Kingdom


STOMACH

Comparative Effectiveness of Famotidine in Hospitalized COVID-19 Patients


Prepublication Release

PEDIATRICS

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

30-Day Outcomes of Children and Adolescents With COVID-19: An International Experience

Talita Duarte-Salles, David Vizcaya, Andrea Pistillo, Paula Casajust, Anthony G. Sena, Lana

RESEARCH

Use of repurposed and adjuvant drugs in hospital patients with covid-19: multinational network cohort study

Albert Prats-Uribe,[1] Anthony G Sena,[2,3] Lana Yin Hui Lai,[4] Waheed-Ul-Rahman Ahmed,[5,6] Heba Alghoul,[7] Osaid Alser,[8] Thamir M Alshammari,[9] Carlos Areia,[10] William Carter,[11] Paula Casajust,[12] Dalia Dawoud,[13,14] Asieh Golozar,[15,16] Jitendra Jonnagaddala,[17] Paras P Mehta,[18] Mengchun Gong,[19] Daniel R Morales,[20,21] Fredrik Nyberg,[22] Jose D Posada,[23] Martina Recalde,[24,25] Elena Roel,[24,25] Karishma Shah,[5] Nigam H Shah,[2,3] Lisa M Schilling,[11] Vignesh Subbian,[26] David Vizcaya,[27] Lin Zhang,[28,29] Ying Zhang,[19] Hong Zhu,[30] Li Liu,[30] Jaehyeong Cho,[31] Kristine E Lynch,[32] Michael E Matheny,[33,34] Seng Chan You,[35] Peter R Rijnbeek,[5] George Hripcsak,[36] Jennifer CE Lane,[5] Edward Burn,[1,26] Christian Reich,[37] Marc A Suchard,[38] Talita Duarte-Salles,[24] Kristin Kostka,[37,39] Patrick B Ryan,[2,40] Daniel Prieto-Alhambra[1]

Renin–angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis

Daniel R Morales, Mitchell M Conover, Seng Chan You, Nicole Pratt, Kristin Kostka, Talita Duarte-Salles, Sergio Fernández-Bertolín, Maria Aragón, Scott L DuVall, Kristine Lynch, Thomas Falconer, Kees van Bochove, Cynthia Sung, Michael E Matheny, Christophe G Lambert, Fredrik Nyberg, Thamir M Alshammari, Andrew E Williams, Rae Woong Park, James Weaver, Anthony G Sena, Martijn J Schuemie, Peter R Rijnbeek, Ross D Williams, Jennifer C E Lane, Albert Prats-Uribe, Lin Zhang, Carlos Areia, Harlan M Krumholz, Daniel Prieto-Alhambra, Patrick B Ryan, George Hripcsak, Marc A Suchard

Summary

Background Angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin receptor blockers (ARBs) have been postulated to affect susceptibility to COVID-19. Observational studies so far have lacked rigorous ascertainment adjustment and internation[illegible] ciated with an increased susceptibi

Lancet Digit Health 2021; 3: e98–114

Published Online December 17, 2020 https://doi.org/10.1016/

Rheumatology 2021;0:1–14 doi:10.1093/rheumatology/keab250

Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study

Jennifer C E Lane, James Weaver*, Kristin Kostka, Talita Duarte-Salles, Maria Tereza F Abrahao, Heba Alghoul, Osaid Alser, Thamir M Alshammari, Patricia Biedermann, Juan M Banda, Edward Burn, Paula Casajust, Mitchell M Conover, Aedin C Culhane, Alexander Davydov, Scott L DuVall, Dmitry Dymshyts, Sergio Fernandez-Bertolin, Kristina Fitter, Jill Hardin, Laura Hester, George Hripcsak, Benjamin Skov Kaas-Hansen, Seamus Kent, Sajan Khosla, Spyros Kolovos, Christophe G Lambert, Johan van der Lei, Kristine E Lynch, Rupa Makadia, Andrea V Margulis, Michael E Matheny, Paras Mehta, Daniel R Morales, Henry Morgan-Stewart, Mees Mosseveld, Danielle Newby, Fredrik Nyberg, Anna Ostropolets, Rae Woong Park, Albert Prats-Uribe, Gowtham A Rao, Christian Reich, Jenna Reps, Peter Rijnbeek, Selva Muthu Kumaran Sathappan, Martijn Schuemie, Sarah Seager, Anthony G Sena, Azza Shoaibi, Matthew Spotnitz, Marc A Suchard, Carmen O Torre, David Vizcaya, Haini Wen, Marcel de Wilde, Junqing Xie, Seng Chan You, Lin Zhang, Oleg Zhuk, Patrick Ryan, Daniel Prieto-Alhambra, on behalf of the OHDSI-COVID-19 consortium*

Summary

Background Hydroxychloroquine, a drug commonly used in the treatment of rheumatoid arthritis, has received much negative publicity for adverse events associated with its authorisation for emergency use to treat patients with COVID-19 pneumonia. We studied the safety of hydroxychloroquine, alone and in combination with azithromycin, to determine the risk associated with its use in routine care in the patients with rheumatoid arthritis.

Lancet Rheumatol 2020

Published Online August 21, 2020 https://doi.org/10.1016/S2665-9913(20)30276-9

RHEUMATOLOGY

Original article

COVID-19 in patients with autoimmune diseases: characteristics and outcomes in a multinational network of cohorts across three countries

Eng Hooi Tan[1], Anthony G. Sena[2,3], Albert Prats-Uribe[1], Seng Chan You[4], Waheed-Ul-Rahman Ahmed[5,6], Kristin Kostka[7], Christian Reich[7], Scott L. Duvall[8,9], Kristine E. Lynch[8,9], Michael E. Matheny[10,11], Talita Duarte-Salles[12], Sergio Fernandez Bertolin[12], George Hripcsak[13,14], Karthik Natarajan[13,14], Thomas Falconer[13], Matthew Spotnitz[13], Anna Ostropolets[13], Clair Blacketer[2,3], Thamir M Alshammari[15], Heba Alghoul[16], Osaid Alser[17], Jennifer C. E. Lane[1], Dalia M. Dawoud[18], Karishma Shah[5], Yue Yang[19], Lin Zhang[20,21], Carlos Areia[22], Asieh Golozar[23,24], Martina Recalde[12,25], Paula Casajust[26], Jitendra Jonnagaddala[27], Vignesh Subbian[28], David Vizcaya[29], Lana Y. H. Lai[30], Fredrik Nyberg[31], Daniel R. Morales[32], Jose D. Posada[33], Nigam H. Shah[33], Mengchun Gong[34], Arani Vivekanantham[5], Aaron Abend[35], Evan P Minty[36], Marc Suchard[37], Peter Rijnbeek[3], Patrick B. Ryan[2,13] and Daniel Prieto-Alhambra[1]

Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational network cohort study

Xintong Li,[1] Anna Ostropolets,[2] Rupa Makadia,[3] Azza Shoaibi,[3] Gowtham Rao,[3] Anthony G Sena,[3,6] Eugenia Martinez-Hernandez,[4] Antonella Delmestri,[1] Katia Verhamme,[6,7] Peter R Rijnbeek,[6] Talita Duarte-Salles,[5] Marc A Suchard,[8,9] Patrick B Ryan,[2,3] George Hripcsak,[2] Daniel Prieto-Alhambra[1,6]


CANCER EPIDEMIOLOGY, BIOMARKERS & PREVENTION

Home About Articles For Authors Alerts News COVID-19 Webinars Search Q

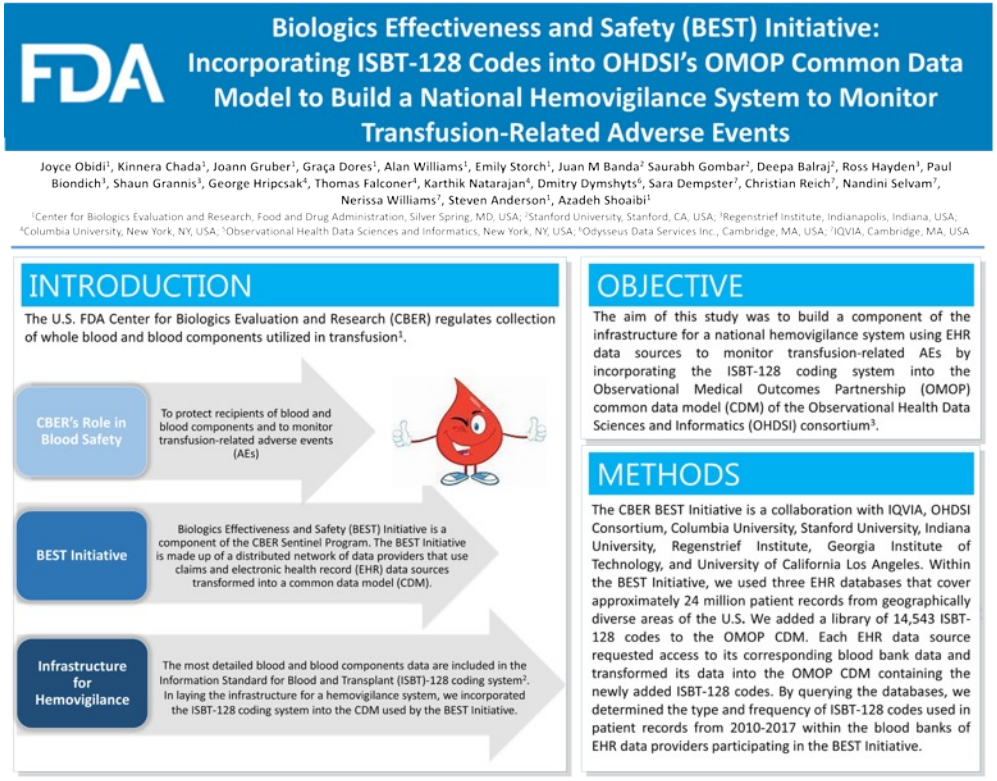
Research Article

Characteristics and outcomes of over 300,000 COVID-19 individuals with history of cancer in the United States and Spain

41 studies in the past year

OHDSI collaborations

US Food and Drug Administration CBER Biologics Effectiveness and Safety (BEST) OHDSI serves as the Convener



FDA

Biologics Effectiveness and Safety (BEST) Initiative: Incorporating ISBT-128 Codes into OHDSI's OMOP Common Data Model to Build a National Hemovigilance System to Monitor Transfusion-Related Adverse Events

Joyce Obidi[1], Kinnera Chada[1], Joann Gruber[1], Graça Dores[1], Alan Williams[1], Emily Storch[1], Juan M Banda[2] Saurabh Gombar[2], Deepa Balraj[2], Ross Hayden[3], Paul Biondich[3], Shaun Grannis[3], George Hripcsak[4], Thomas Falconer[4], Karthik Natarajan[4], Dmitry Dymshyts[6], Sara Dempster[7], Christian Reich[7], Nandini Selvam[7], Nerissa Williams[7], Steven Anderson[1], Azadeh Shoaibi[1]

[1]Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA; [2]Stanford University, Stanford, CA, USA; [3]Regenstrief Institute, Indianapolis, Indiana, USA; [4]Columbia University, New York, NY, USA; [5]Observational Health Data Sciences and Informatics, New York, NY, USA; [6]Odysseus Data Services Inc., Cambridge, MA, USA; [7]IQVIA, Cambridge, MA, USA

INTRODUCTION

The U.S. FDA Center for Biologics Evaluation and Research (CBER) regulates collection of whole blood and blood components utilized in transfusion[3].

CBER's Role in Blood Safety — To protect recipients of blood and blood components and to monitor transfusion-related adverse events (AEs)

BEST Initiative — Biologics Effectiveness and Safety (BEST) Initiative is a component of the CBER Sentinel Program. The BEST Initiative is made up of a distributed network of data providers that use claims and electronic health record (EHR) data sources transformed into a common data model (CDM).

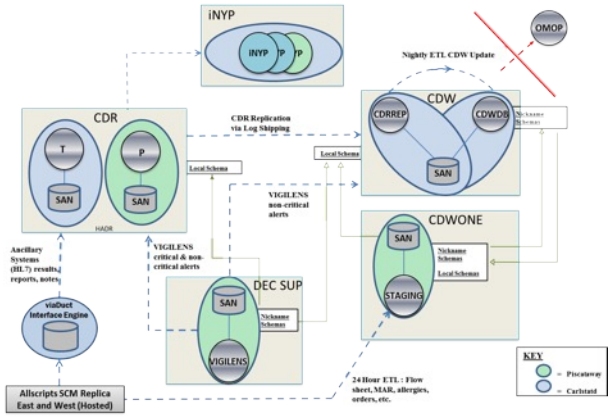
Infrastructure for Hemovigilance — The most detailed blood and blood components data are included in the Information Standard for Blood and Transplant (ISBT)-128 coding system[2]. In laying the infrastructure for a hemovigilance system, we incorporated the ISBT-128 coding system into the CDM used by the BEST Initiative.

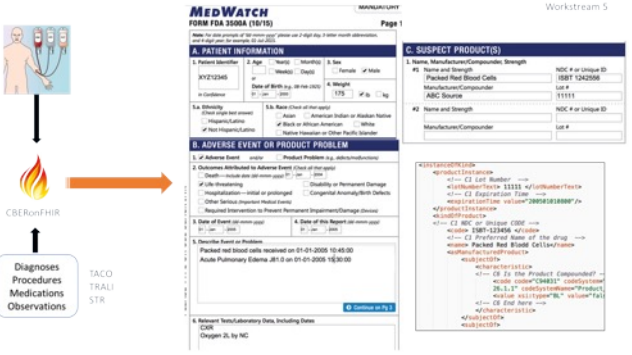
OBJECTIVE

The aim of this study was to build a component of the infrastructure for a national hemovigilance system using EHR data sources to monitor transfusion-related AEs by incorporating the ISBT-128 coding system into the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) of the Observational Health Data Sciences and Informatics (OHDSI) consortium[3].

METHODS

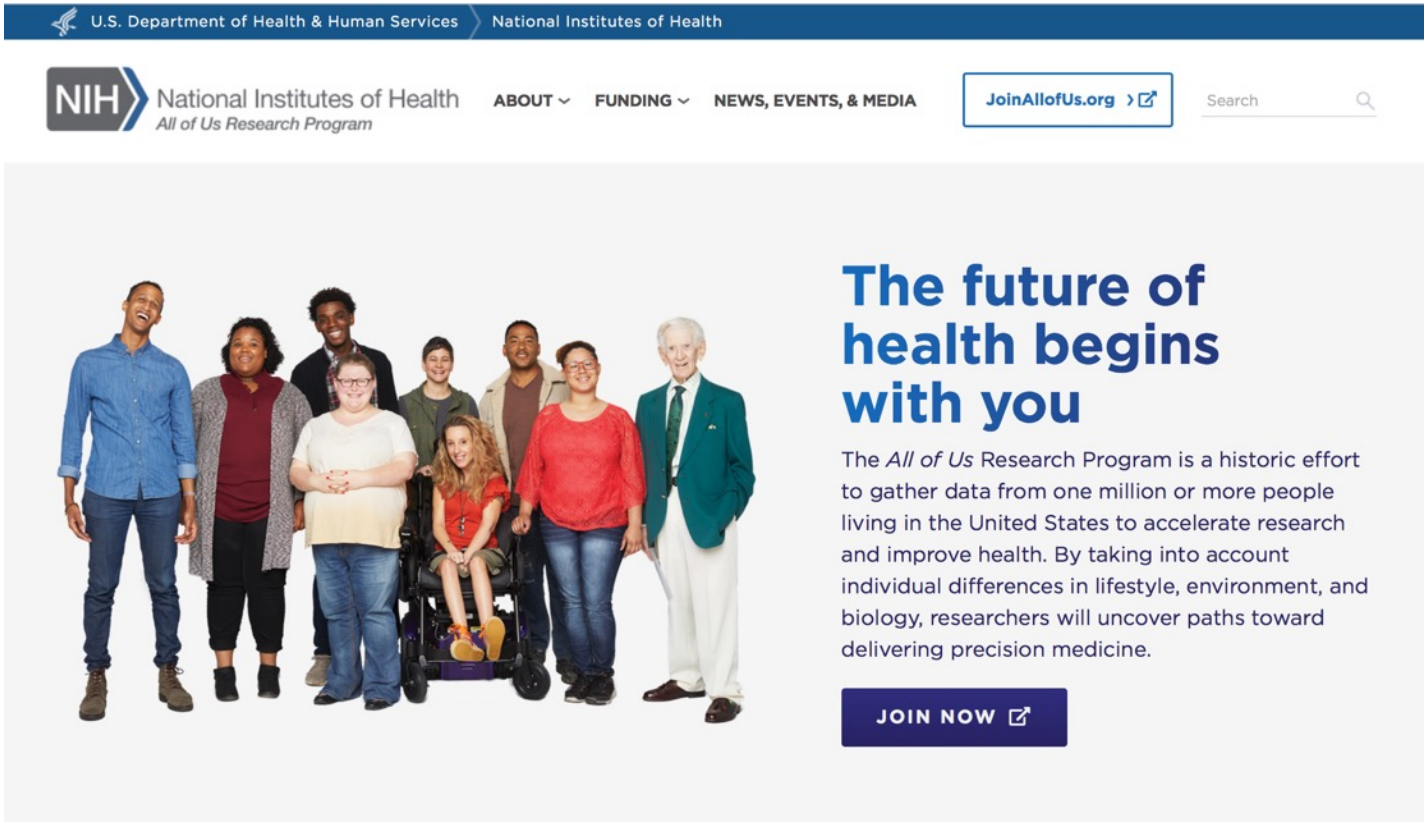
The CBER BEST Initiative is a collaboration with IQVIA, OHDSI Consortium, Columbia University, Stanford University, Indiana University, Regenstrief Institute, Georgia Institute of Technology, and University of California Los Angeles. Within the BEST Initiative, we used three EHR databases that cover approximately 24 million patient records from geographically diverse areas of the U.S. We added a library of 14,543 ISBT-128 codes to the OMOP CDM. Each EHR data source requested access to its corresponding blood bank data and transformed its data into the OMOP CDM containing the newly added ISBT-128 codes. By querying the databases, we determined the type and frequency of ISBT-128 codes used in patient records from 2010-2017 within the blood banks of EHR data providers participating in the BEST Initiative.



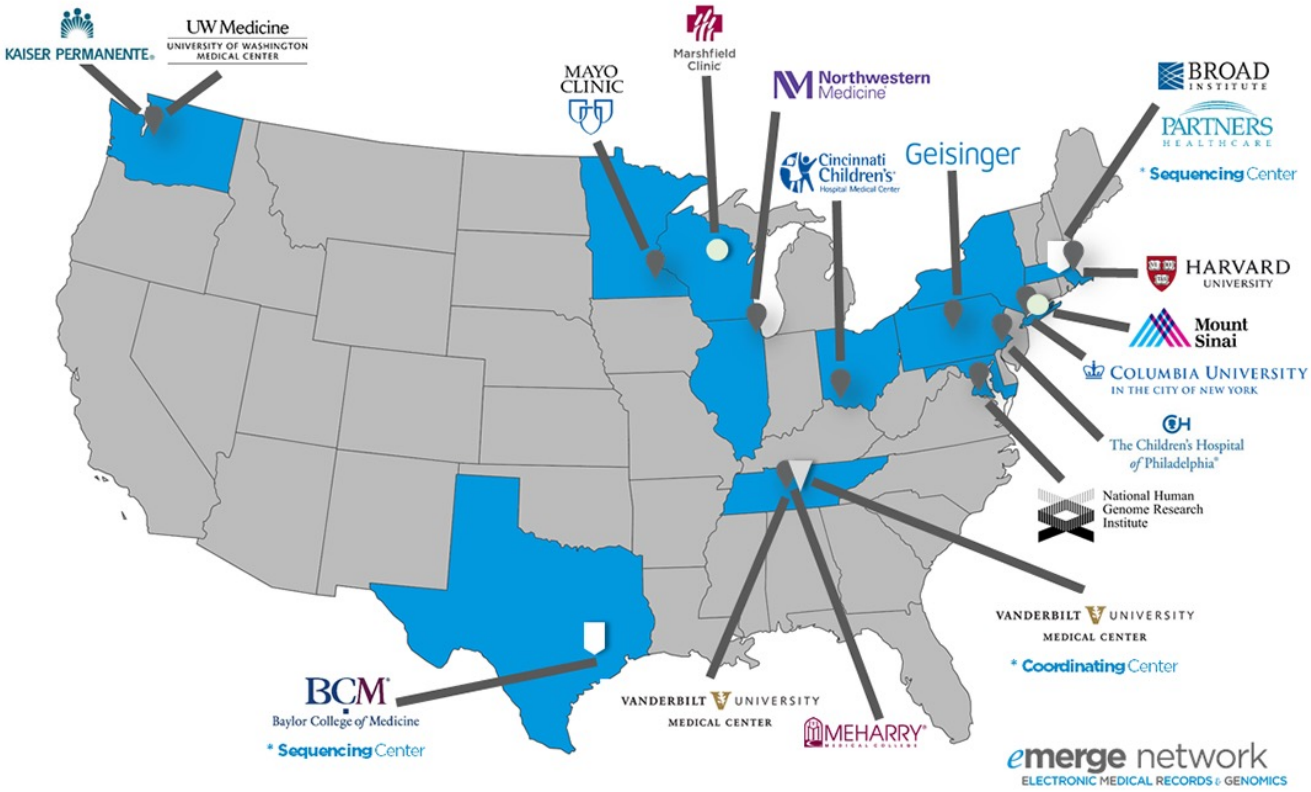


US National Institutes of Health
All of Us Research Program

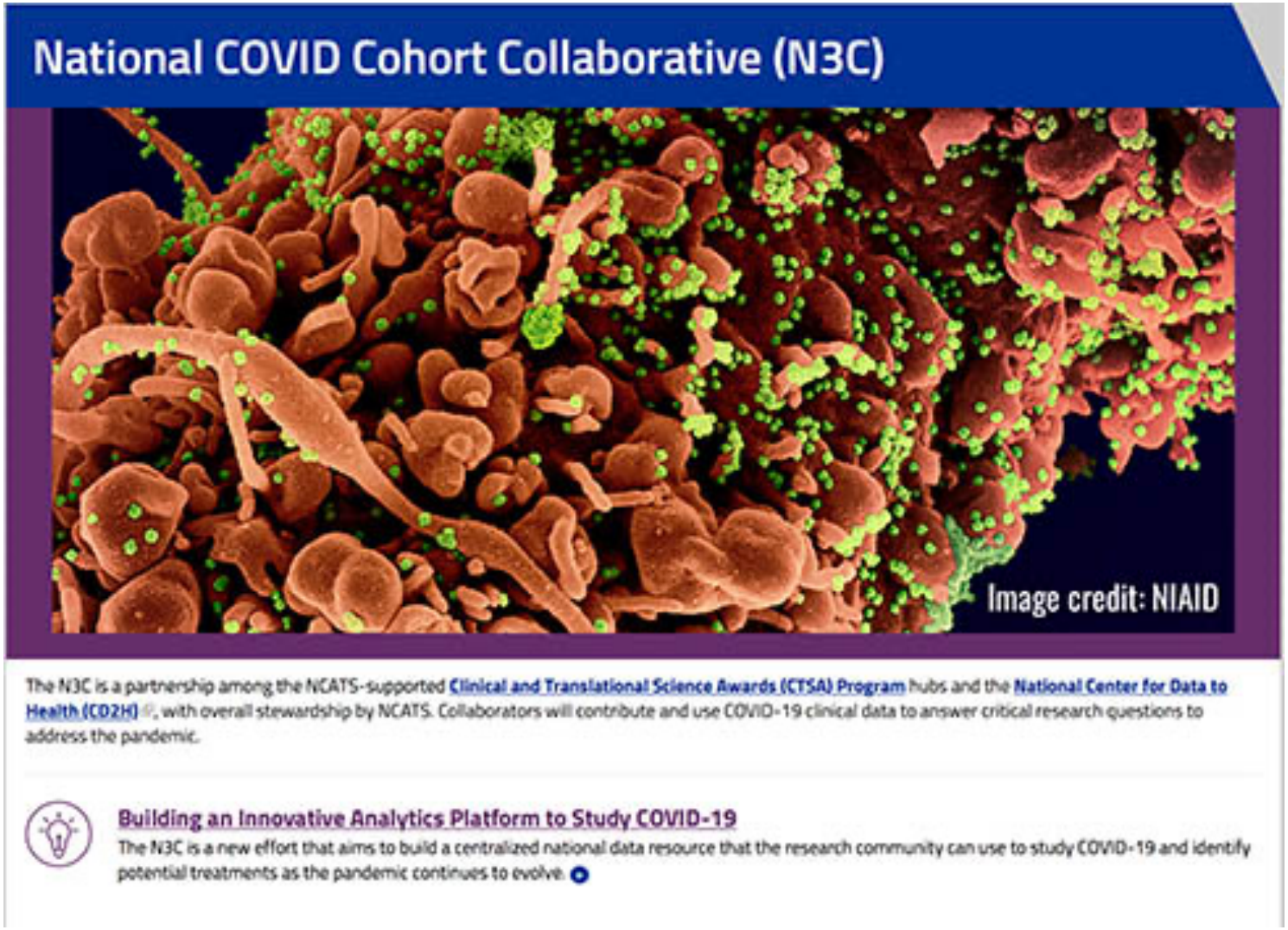
- 1,000,000 diverse participants
- Clinical data in OMOP CDM
- $100Ms



US National Institutes of Health Electronic Medical Records and Genomics (eMERGE)



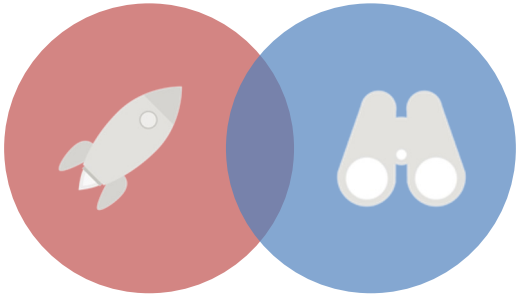
National COVID Cohort Collaborative (N3C)



The European Health Data and Evidence Network (EHDEN) 30M € Innovative Medicines Initiative

Mission

Our mission is to provide a new paradigm for the discovery and analysis of health data in Europe, by building a large-scale, federated network of data sources standardised to a common data model

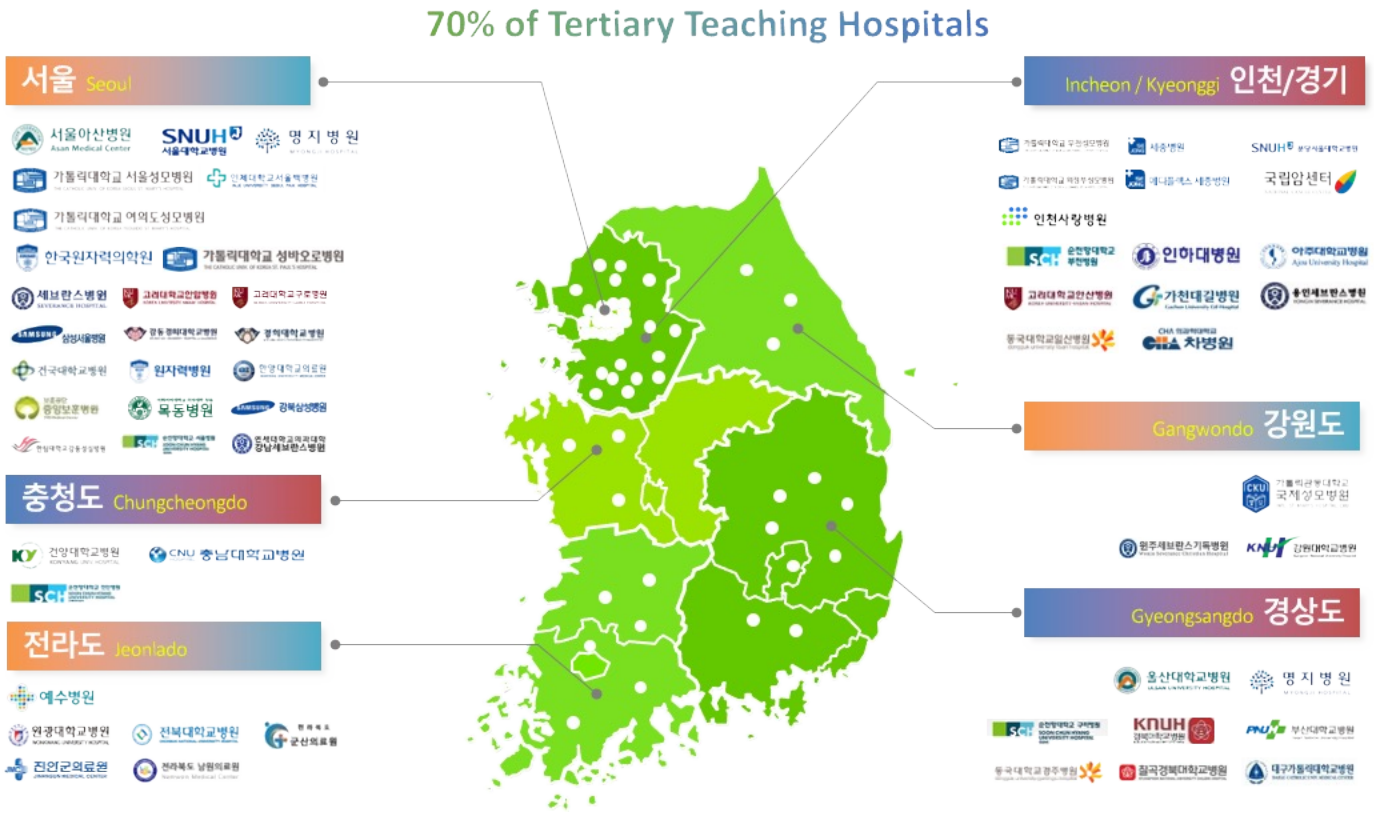


Vision

The European Health Data & Evidence Network (EHDEN) aspires to be the trusted observational research ecosystem to enable better health decisions, outcomes and care

DARWIN EU – European Medicines drug surveillance initiative
Erasmus MC using OHDSI

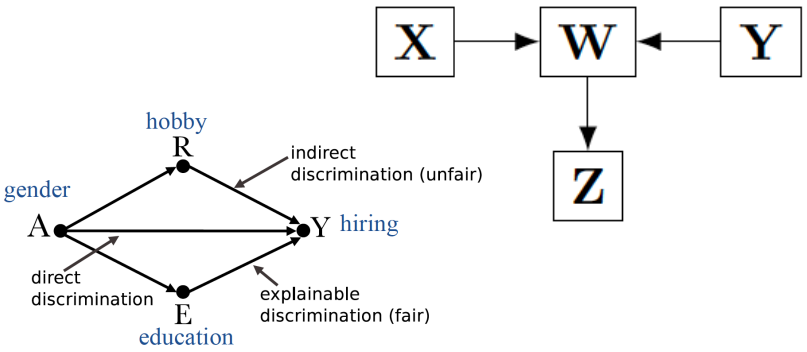
National CDM Projects in Korea
OHDSI Data Network



Lab's other causal inference work

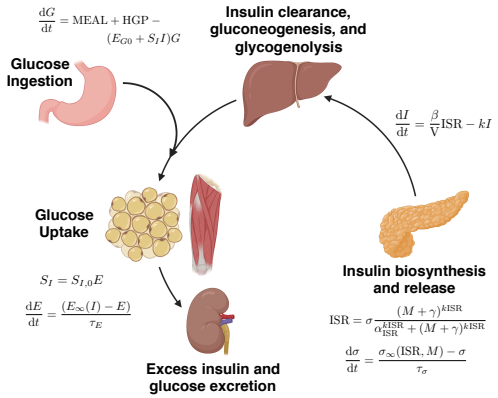
Structural causal models

Bareinboim, Blei, Zhang, Anand



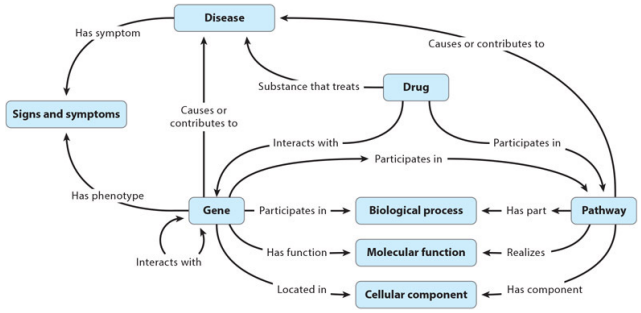
Mechanistic models

Albers, Richter, Albert



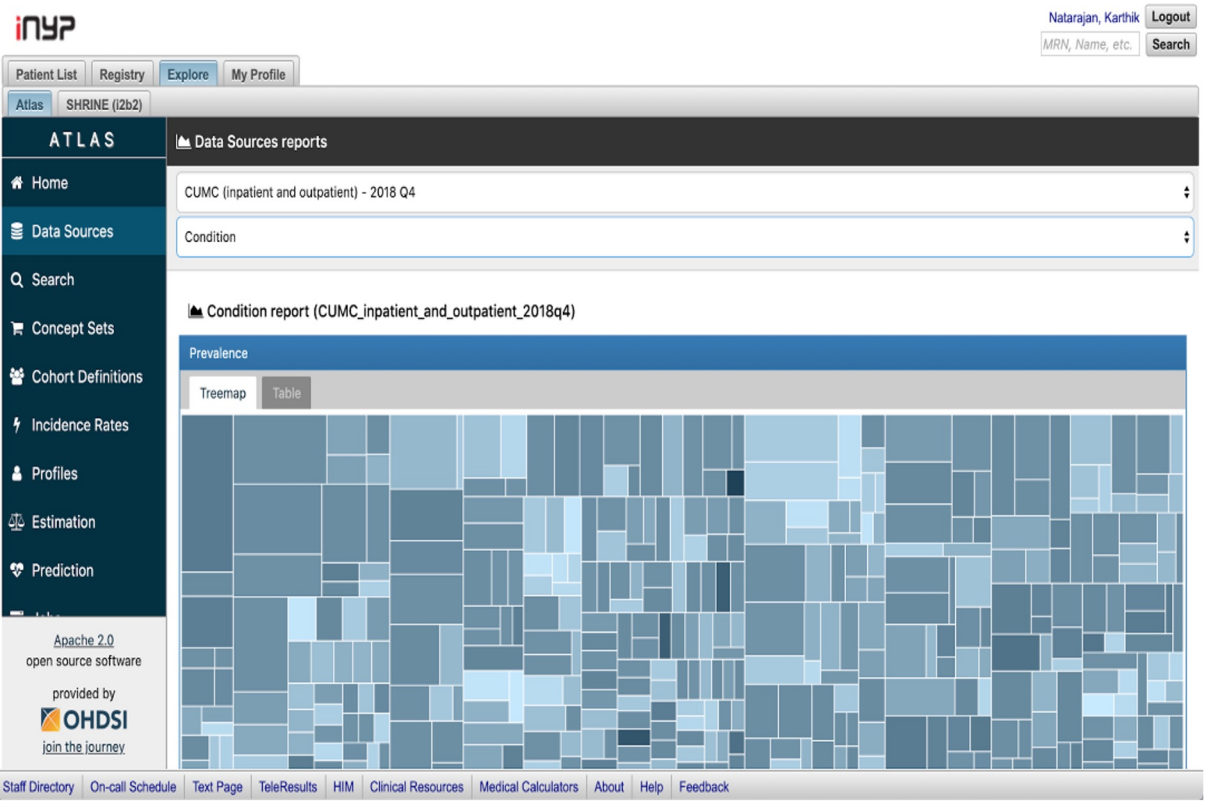
Ontologies

Callahan



OHDSI at CUIMC

iNYP -> Explore (Epic link)



- "De-identified" OMOP
  - refreshed 1-2 times a year
- Access to 15K users via clinical system
  - "Can I have this for my patient list"
- Use it to collaborate with clinical researchers
  - "Data consultation" ~10/yr
  - Improves our data

Columbia Data Consult Service

- Research project to study the effect of real-time evidence generation
- Put observational research into action
- 29 questions, 22 clinicians, 24 answers
  - Largely medicine, but due to recruitment
  - Mostly were about recurring issues
  - A fifth about a specific patient

Ostropolets JAMIA 2021

Observational Research Task Force

- Dawn Hershman, George Hripcsak, co-chairs
- Membership being confirmed
  - Clinical researchers, epidemiology, equity, informatics, privacy, health records

- Components
  - Research themes for emphasis
  - Faculty development and growth in these areas of research
  - Multi-PI collaborations
  - Education/training
  - Linkage to other initiatives and their task forces' reports: e.g. core facilities (completed); clinical trials (completed); biostatistics (Under Dr. Kiros Berhane, ongoing).

Summary

- Current observational research is suspect
- Large-scale observational research appears to be possible and more reliable than the current approach
- 6-million-patient database is available for research today (and Marketscan with funding)

Funding
National Library of Medicine
R01 LM006910 (2000-2024)
T15 LM007079 (1994-2027)
NIDDK R01 DK090372
NCI contracts
NHGRI All of Us
FDA CBER BEST