



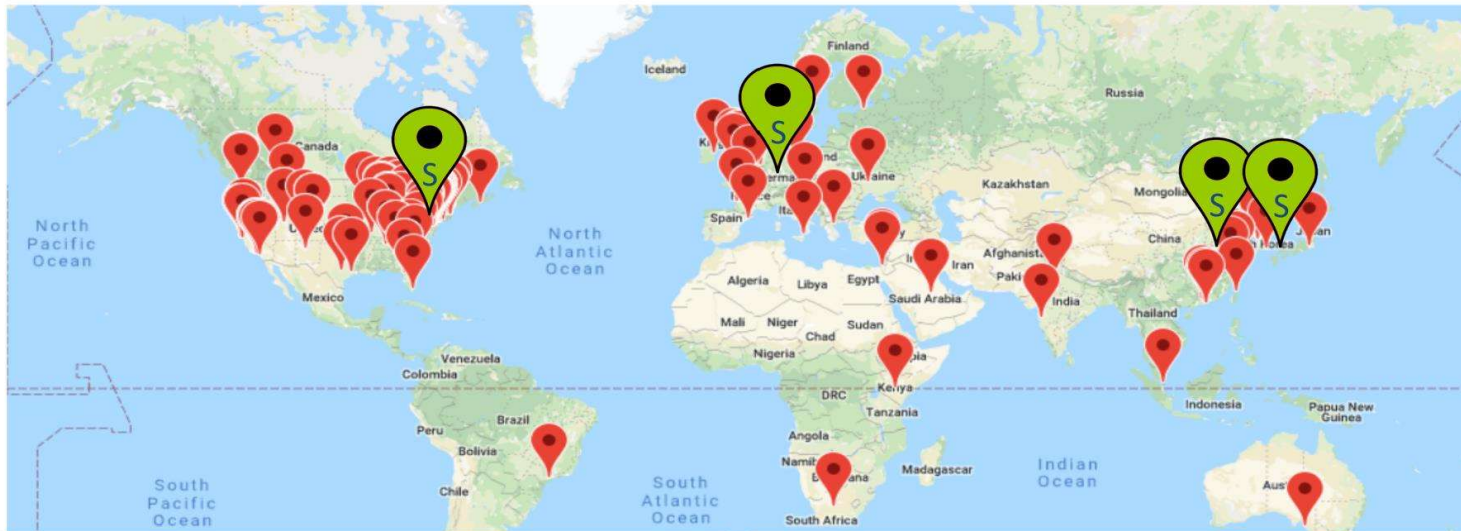
# Knowledge sharing to advance health: perspectives from OHDSI Data Collaborative

Dr. Juan M. Banda  
12/04/2020



# What is OHDSI?

- *Observational Health Data Sciences and Informatics (OHDSI)* program is a multi-stakeholder, interdisciplinary collaborative to create open-source solutions that bring out the value of observational health data through large-scale analytics
- Vision: OHDSI collaborators access a network of 1 billion patients to generate evidence about all aspects of healthcare. Patients and clinicians and other decision-makers around the world use OHDSI tools and evidence every day



### OHDSI Collaborators:

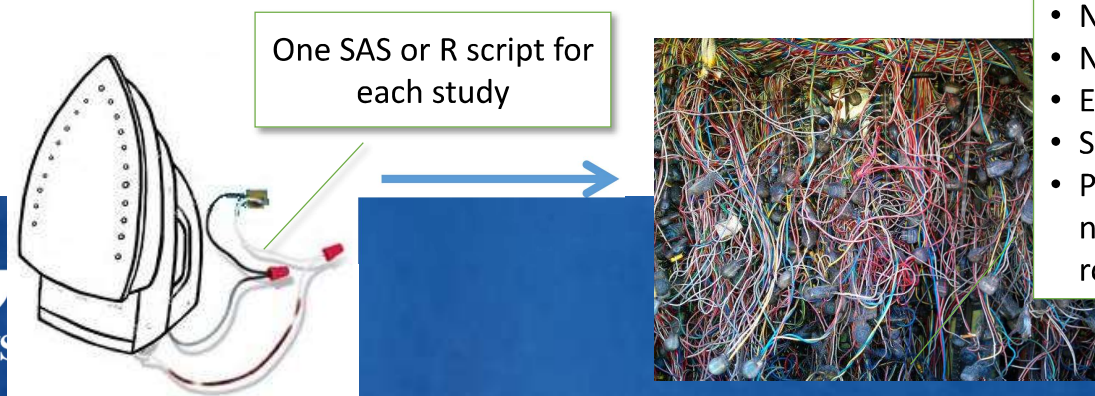
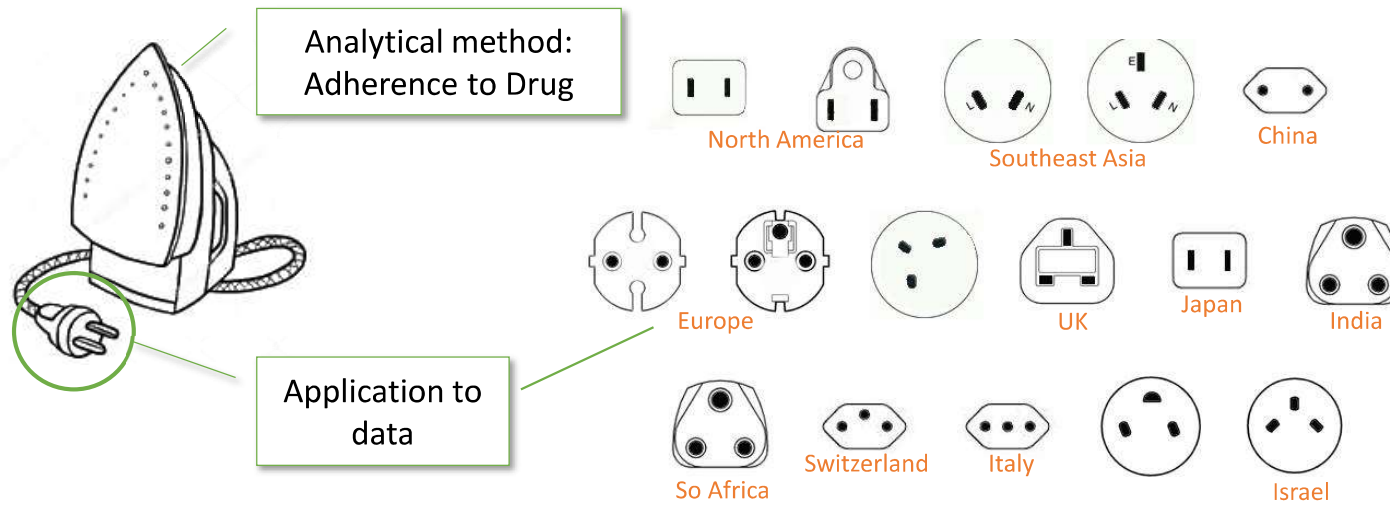
- 2,770 users
- 25 workgroups
- 18,700 posts on 3,250 topics

### OHDSI Network:

- 152 databases
- 18 countries
- approx. 600M patient records

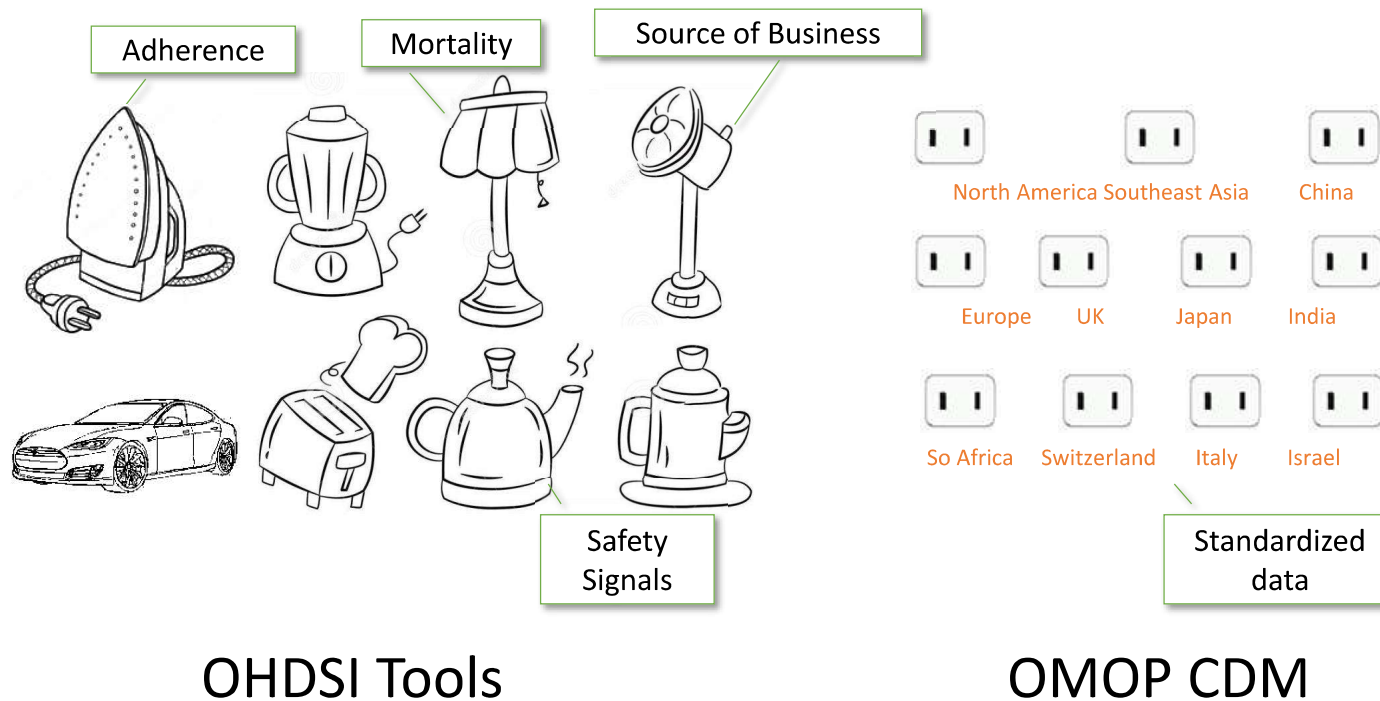
# Current Approach: "One Study – One Script"

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

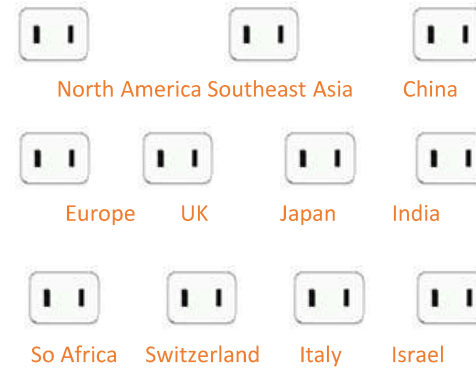


- Not scalable
- Not transparent
- Expensive
- Slow
- Prohibitive to non-expert routine use

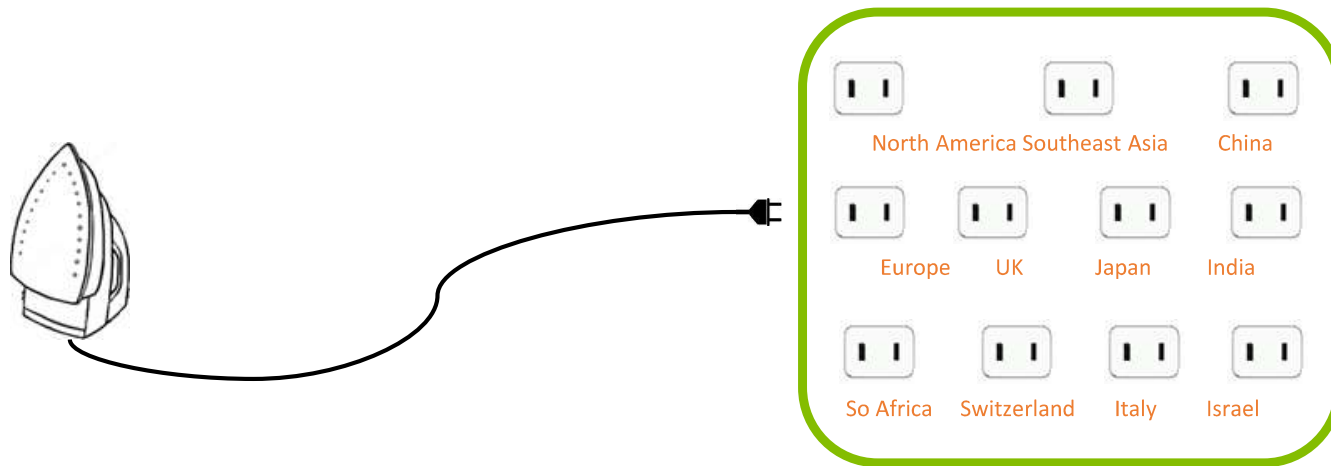
# Solution: Data Standardization Enables Systematic Research



# Analytics can be remote

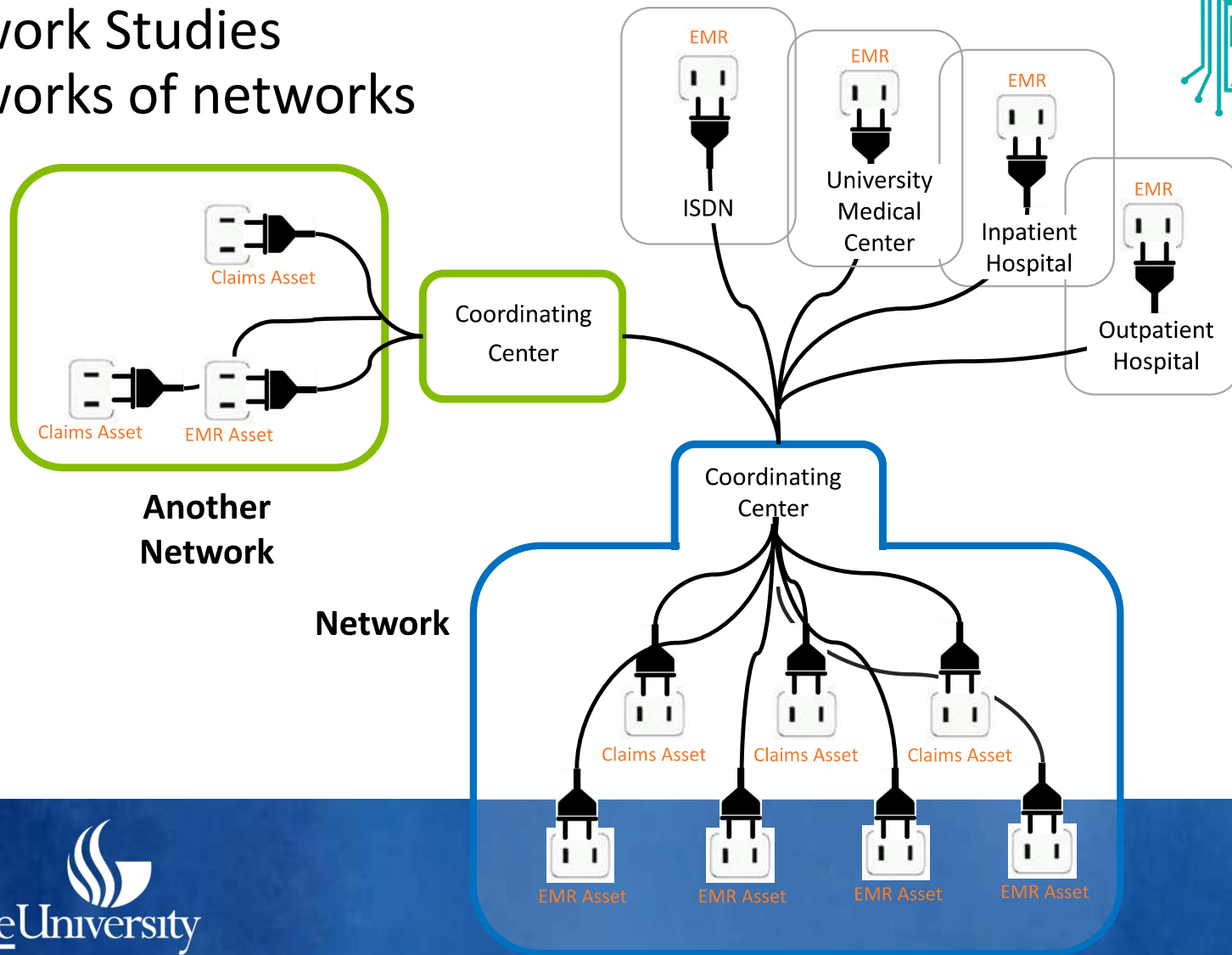


# Analytics can be behind firewall



# Network Studies

## Networks of networks







# OHDSI strategy

- Standardize the data
- Standardize the analytics

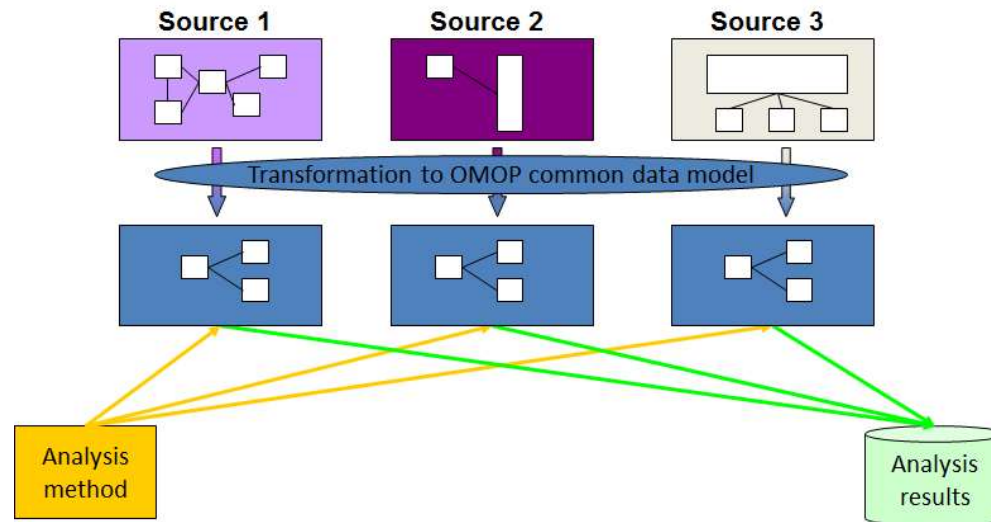
... simple right?



So, what is the cost of admission?

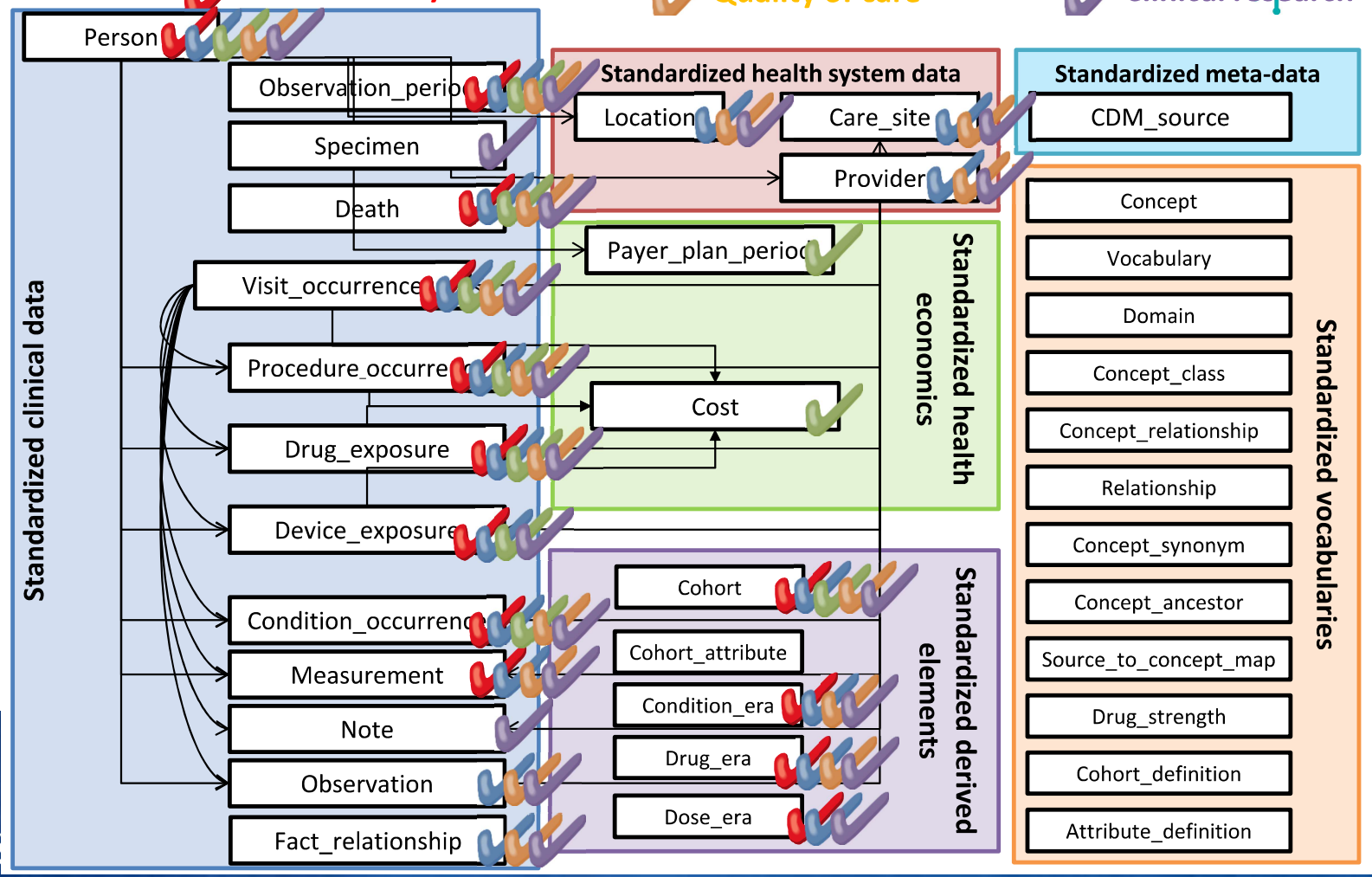
# OHDSI provides all the tools... Open Source!

- Standard Common Data Model is the cornerstone



- ✓ Drug safety surveillance
- ✓ Device safety surveillance
- ✓ Vaccine safety surveillance

- ✓ Comparative effectiveness
- ✓ Health economics
- ✓ Quality of care





## What OHDSI is not

- A centralized data network
- A federated query mechanism
- An NLP pipeline / data conversion tool
- A big dataset of publicly available data



OHDSI provides all the tools... Open Source!

- Standardized Common Vocabulary – Athena
  - <https://athena.ohdsi.org/search-terms/start>
- Data ETL Design – White Rabbit
  - <https://www.ohdsi.org/analytic-tools/whiterabbit-for-etl-design/>
- Data Quality - With Achilles and Data Quality Dashboard
  - <https://data.ohdsi.org/DataQualityDashboard/>

# Standardized Phenotyping in OHDSI

- You can use the traditional rule-based way of building phenotypes (like PheKB) and build them in ATLAS
  - <https://www.ohdsi.org/atlas-a-unified-interface-for-the-ohdsi-tools/>
- Brand new phenotype library contains a hefty list of several dozen phenotypes freely available
  - <https://data.ohdsi.org/PhenotypeLibrary/>
- Looking for machine learning, probabilistic phenotypes?
  - APHRODITE is the tool - <https://github.com/OHDSI/Aphrodite>





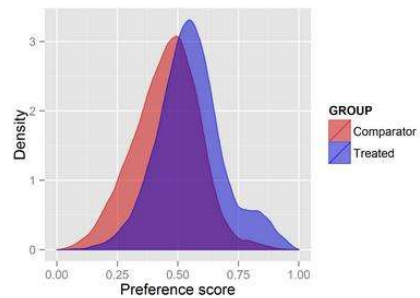
And you can SHARE all these objects!

- ATLAS cohorts are JSON objects that can be shared
- Phenotype Library comes as an R package
- APHRODITE phenotypes can be shared and are FAIR

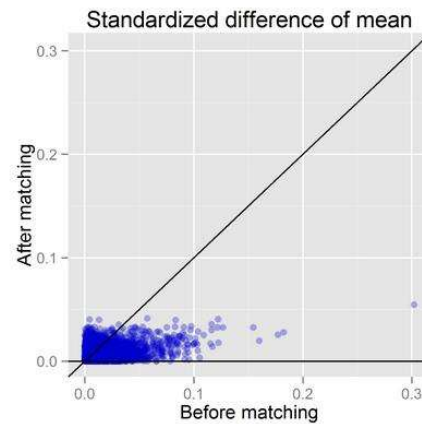


And plenty more tools/software available

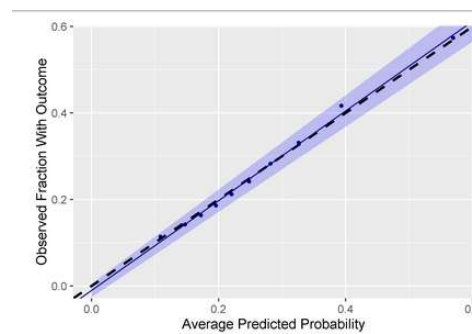
- R Packages – CohortMethod, PatientLevelPrediction,



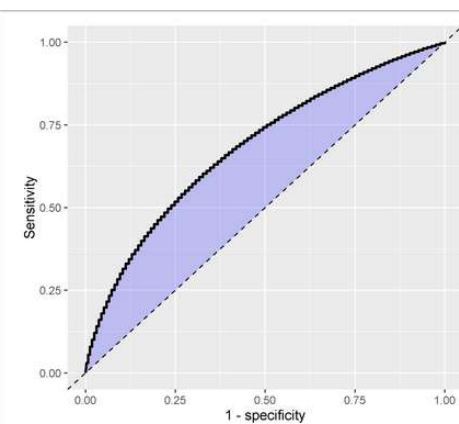
Propensity (preference score) distribution



Covariate balance plot



Calibration plot



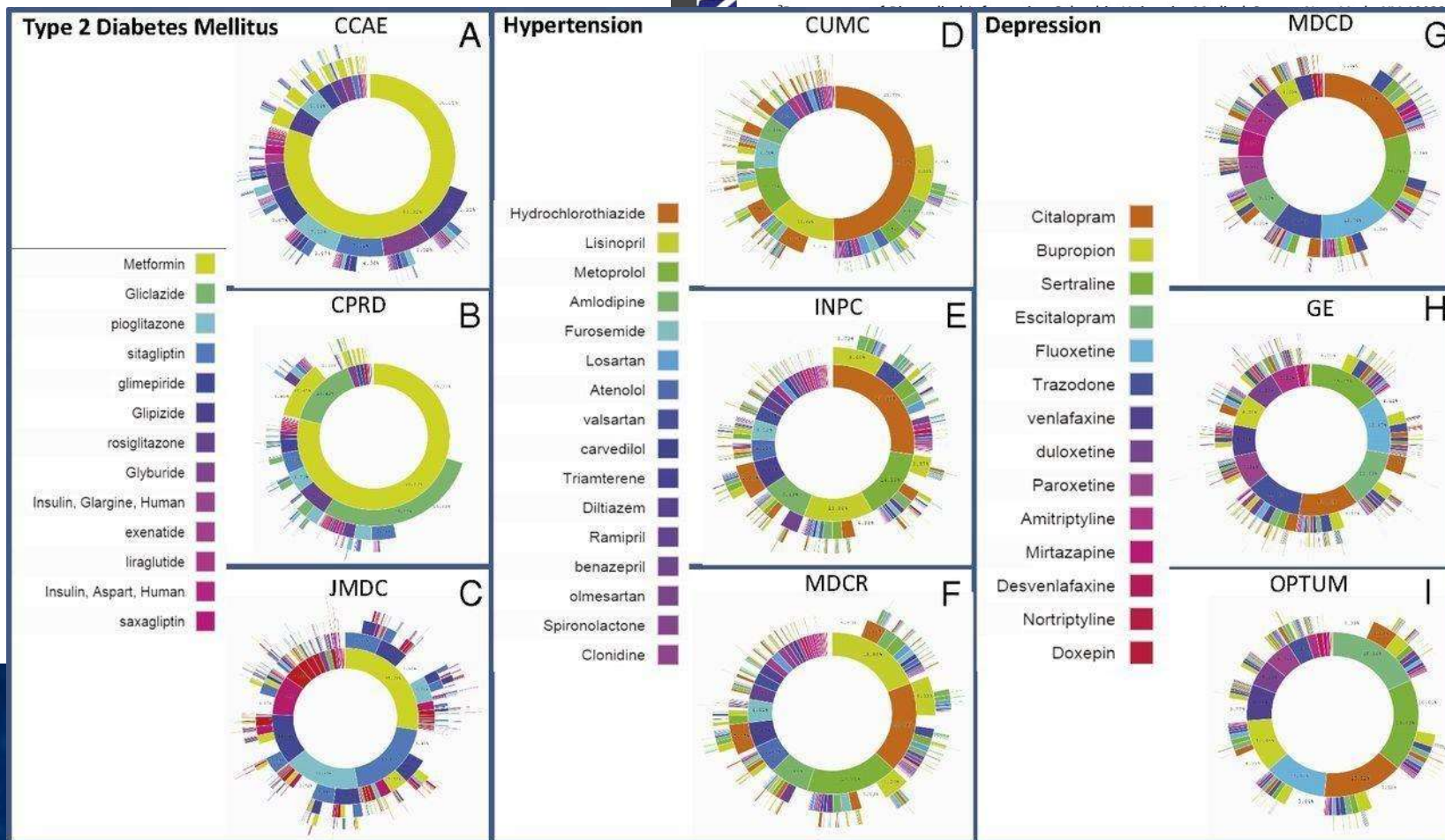
ROC plot

# Does it work?

OHDSI

## Characterizing treatment pathways at scale using the OHDSI network

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



<sup>a</sup>Medical Informatics Services, NewYork-Presbyterian 0032; <sup>b</sup>Epidemiology Analytics, Janssen Research and 0032; <sup>c</sup>Epidemiology Analytics, Janssen Research and 0032; <sup>d</sup>Epidemiology Analytics, Janssen Research and 0032; <sup>e</sup>Epidemiology Analytics, Janssen Research and 0032; <sup>f</sup>Center for Biomedical Informatics 0032; <sup>g</sup>Center for Biomedical Informatics 0032; <sup>h</sup>Center for Biomedical Informatics 0032; <sup>i</sup>Center for Biomedical Informatics 0032; <sup>j</sup>Center for Biomedical Informatics 0032; <sup>k</sup>Center for Biomedical Informatics 0032; <sup>l</sup>Lister 0032; <sup>m</sup>Lister 0032; <sup>n</sup>Lister 0032; <sup>o</sup>Lister 0032; <sup>p</sup>Lister 0032; <sup>q</sup>Lister 0032; <sup>r</sup>Lister 0032; <sup>s</sup>Lister 0032

# COVID-19 efforts – Study-A-thon

In only **88** hours, we have:

- Convened **351** participants brought from **30** countries
- Held **12** Global Huddles, **>100** calls, **>13,000** chat messages
- Engaged **15** concurrent channels
- Reviewed **>10,000** publications
- Drafted **9** protocols
- Released **13** study packages
- Designed **355** cohort definitions
- Assembled a distributed data network with **37** partners signed on to execute studies

## nature communications

Explore our content ▾ Journal information ▾

nature THE LANCET  
Rheumatology

Article |

Dee  
hos  
netw

Edward

Nature

14k Ac

ARTICLES | VOLUME 2, ISSUE 11, E698-E711, NOVEMBER 01, 2020



PDF [1 MB]

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

Jennifer C E Lane, MRCS <sup>†</sup> • James Weaver, MSc <sup>†</sup> • Kristin Kostka, MPH • Talita Duarte-Salles, PhD • Maria Tereza F Abrahao, PhD • Heba Alghoul, MD • et al. [Show all authors](#) • [Show footnotes](#)

[Open Access](#) • Published: August 21, 2020 • DOI: [https://doi.org/10.1016/S2665-9913\(20\)30276-9](https://doi.org/10.1016/S2665-9913(20)30276-9)



# APHRODITE phenotypes in multiple sites and countries



Volume 27, Issue 6  
June 2020

## Development and validation of phenotype classifiers across multiple sites in the observational health data sciences and informatics network

Mehr Kashyap ✉, Martin Seneviratne, Juan M Banda, Thomas Falconer, Borim Ryu, Sooyoung Yoo, George Hripcsak, Nigam H Shah

*Journal of the American Medical Informatics Association*, Volume 27, Issue 6, June 2020, Pages 877–883, <https://doi.org/10.1093/jamia/ocaa032>

**Published:** 06 May 2020 **Article history** ▼

<https://doi.org/10.1093/jamia/ocaa032>



## More resources

- The OHDSI Bible:
  - <https://ohdsi.github.io/TheBookOfOhdsi/>
  
- Questions? Need some help joining the OHDSI journey?
  - [jbanda@gsu.edu](mailto:jbanda@gsu.edu)