Building Data Capacity for Patient-Centered Outcomes Research: Perspectives from the OHDSI community

Patrick Ryan, PhD

Vice President, Observational Health Data Analytics, Janssen Research and Development

Assistant Professor, Adjunct, Department of Biomedical Informatics, Columbia University Medical Center

OHDSI to advance the science of observational research

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

- OHDSI is an open science community, founded by scientists at J&J, Columbia University, Stanford, UCLA, Regenstrief Institute, and Iqvia
- OHDSI is an **open multi-stakeholder**, interdisciplinary collaborative with a mission to improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care
- OHDSI is driving development and adoption of **open community data standards, open source analysis software**, and **open science best practices** within regulators, academia, industry, payors, and health systems

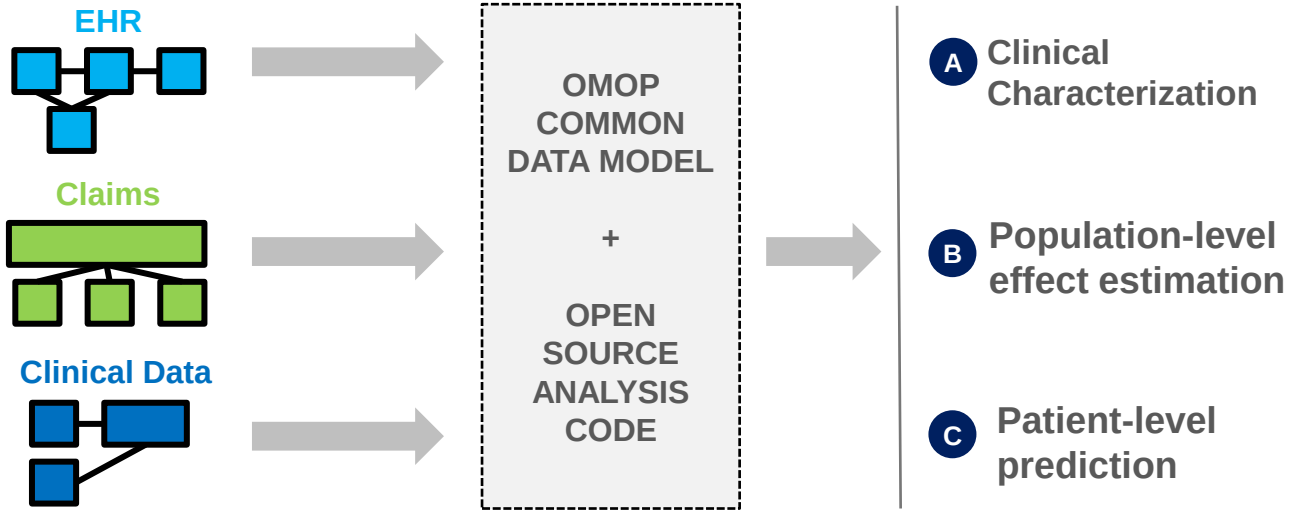
GLOBAL OHDSI NETWORK - *Industry, Academia, Government*

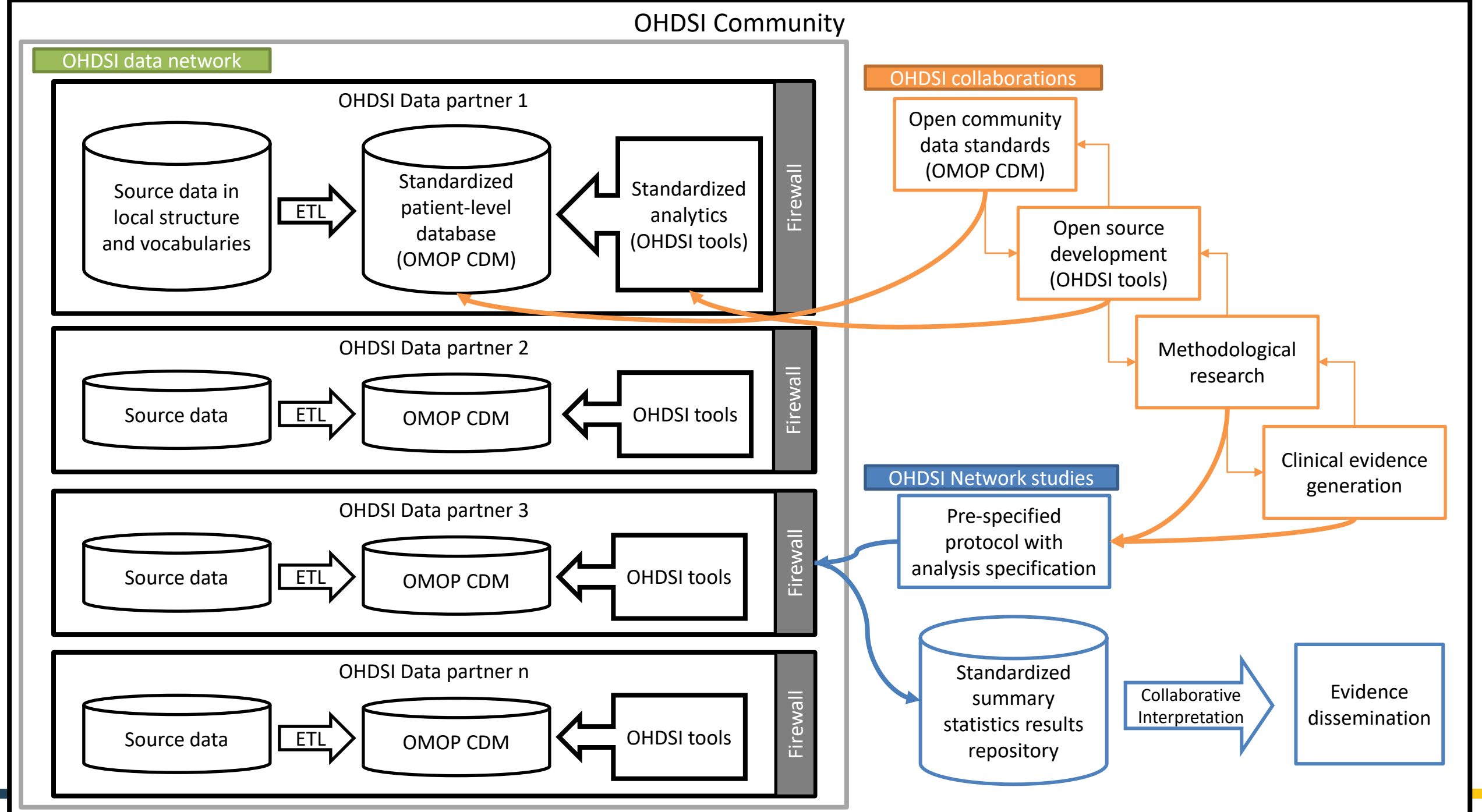
>2,100
Researchers in epidemiology, statistics, informatics, health policy, and clinical sciences

>150
Databases Included

>2B
Patient records Represented

OMOP DRIVING RWE





Hripcsak G, Schuemie MJ, Madigan D, Ryan PB, Suchard MA. Drawing Reproducible Conclusions from Observational Clinical Data with OHDSI. Yearb Med Inform. 2021 Apr 21.

Data standards are a means to an end, not an end in itself

- What *data standards* could make the PCOR data infrastructure more useful for research and other data needs?
- What *data standards* are likely to become more relevant looking forward?
- What needs to be prioritized?

- What **evidence** would be useful to improve health policy and health care, which could be reliably generated by the PCOR data infrastructure?
- How can *data standards* enable real-world **analytics** to meet the relevant **evidence** needs moving forward?
- What needs to be prioritized?

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

- Data standards are necessary to enable replicability, generalizability, and robustness
- Data standards without standardized analytics are not sufficient to ensure reliable evidence
- Standards limited to US data are harmful for US population; public health questions require global data to generate global evidence

Delineating the roles of standards to enable evidence

How can *data standards* enable real-world **analytics** to meet the relevant **evidence** needs moving forward?

What **evidence** would be useful to improve health policy and health care?

Data standards to enable data exchange

HL7 FHIR

Data standards to harmonize data structure and enable analytics

OMOP CDM

Analytics standards to generate and disseminate evidence

HADES

Vocabulary standards to harmonize data content and enable analytics

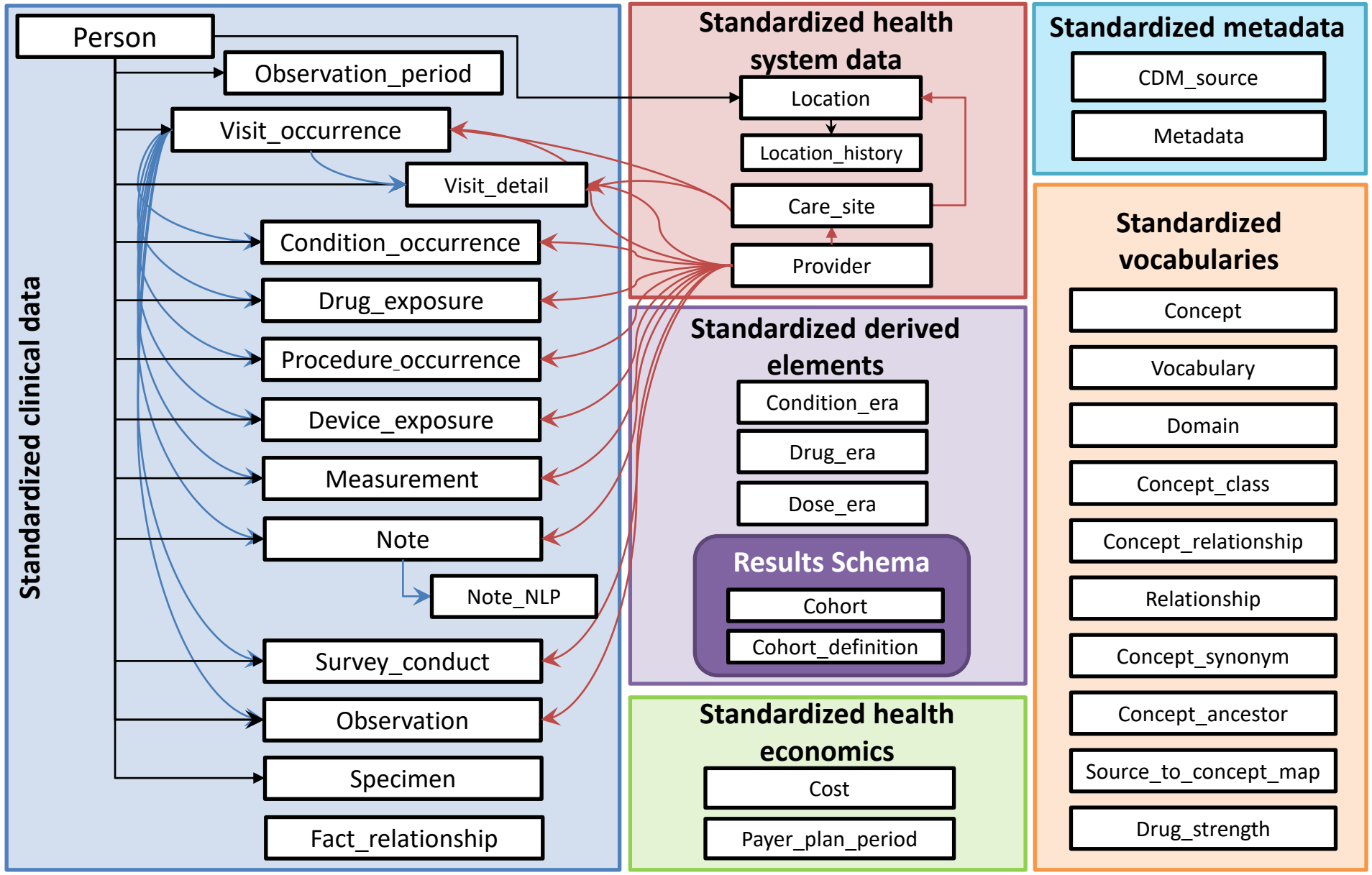
ATHENA

Characterization:
Who are the patients exposed to medical interventions?
How often do outcomes occur amongst those patients?

Estimation:
Is the risk of outcome causally related to exposure?
How does the risk compare with alternative interventions?

Prediction:
Which risks can be actionably predicted with available data?
Which patients are at highest risk of adverse events?

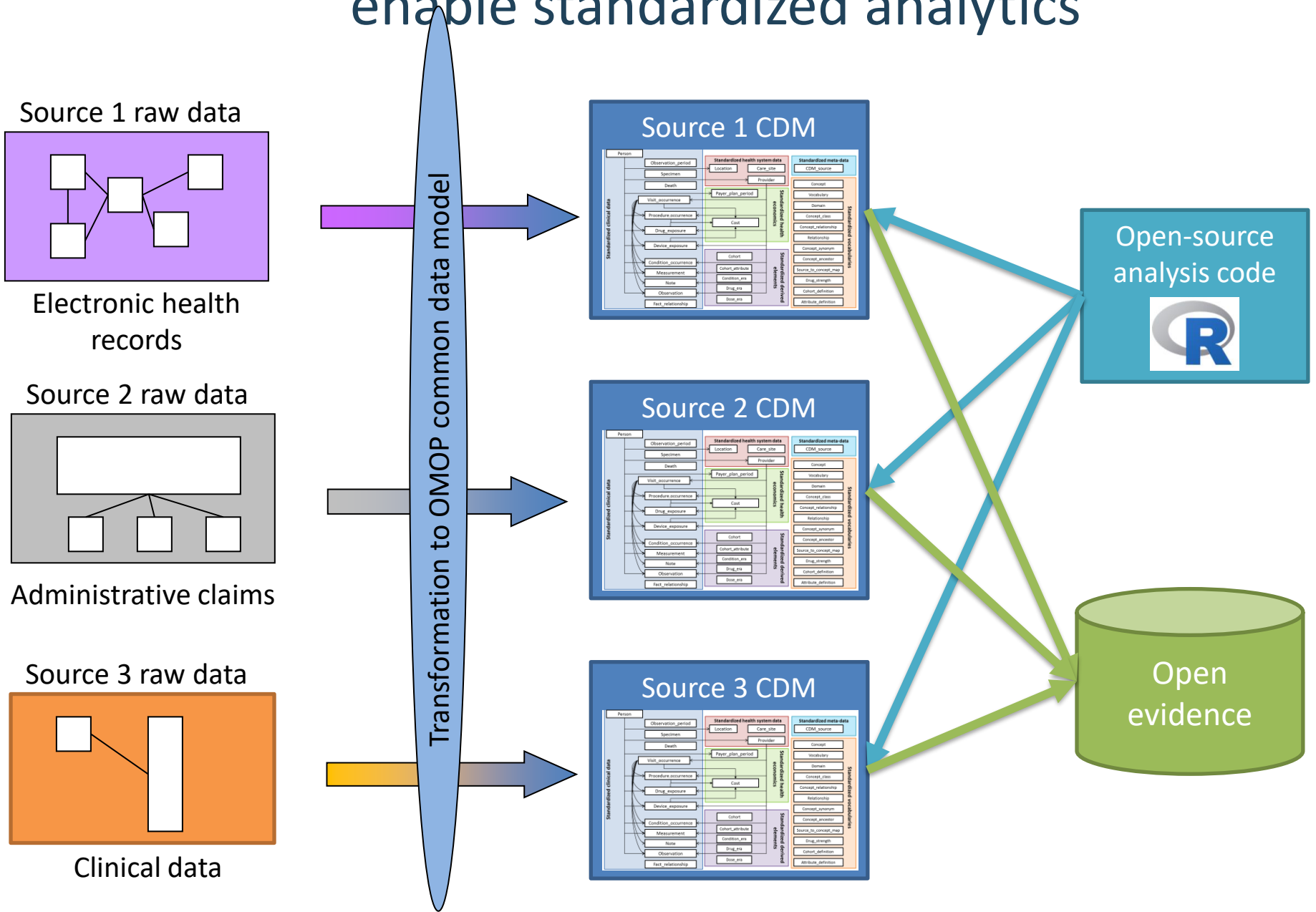
Open community data standard to enable analytics: OMOP Common Data Model



OHDSI's standardized vocabularies to harmonize data content and enable analytics

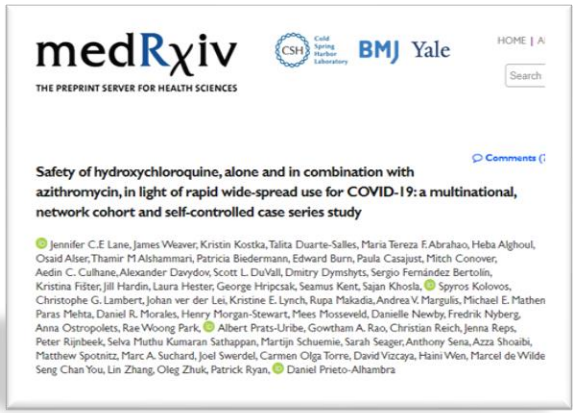
- 164 Vocabularies across 41 domains
  - MU3 standards: SNOMED, RxNorm, LOINC
  - Disparate sources: ICD9CM, ICD10(CM), Read, NDC, Gemscript, CPT4, HCPCS…
- >9.5 million concepts
  - >3.4 million standard concepts
  - >5.4 million source codes
  - >674,000 classification concepts
- >57 million concept relationships
- >86 million ancestral relationships

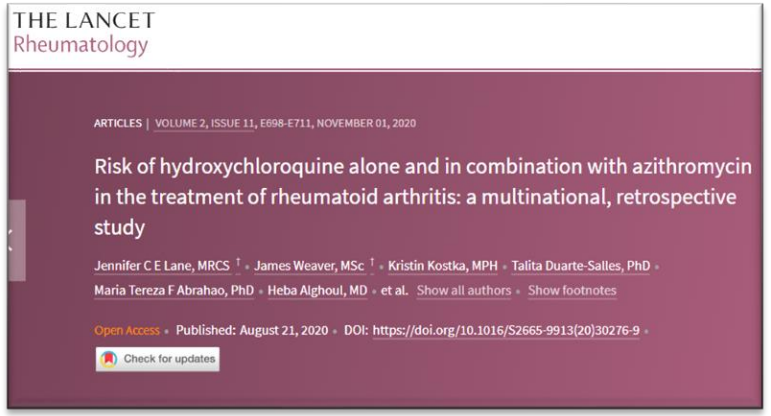
Common data model and standardized vocabularies to enable standardized analytics

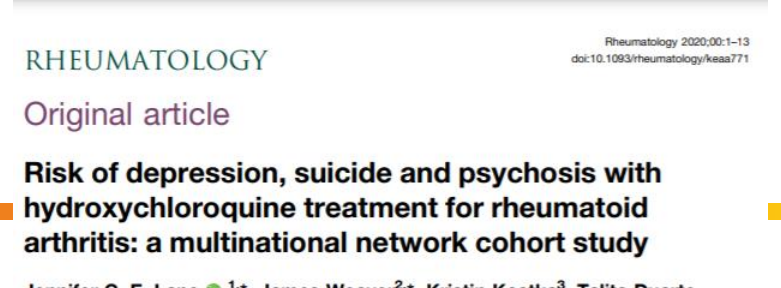


Standards enabling evidence for regulators: Safety of hydroxychloroquine (HCQ)

- Evidence needed around the safety of HCQ alone and in combination with azithromycin (AZ)
- Standardized analytics: comparative cohort design with large-scale propensity adjustment
- Standardized data: 14 databases (claims and EHR) from 6 countries, all using OMOP CDM
- Findings: In RA population, HCQ alone is generally safe but in combination with AZ it shows a doubling of risk of 30-day cardiovascular mortality.







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

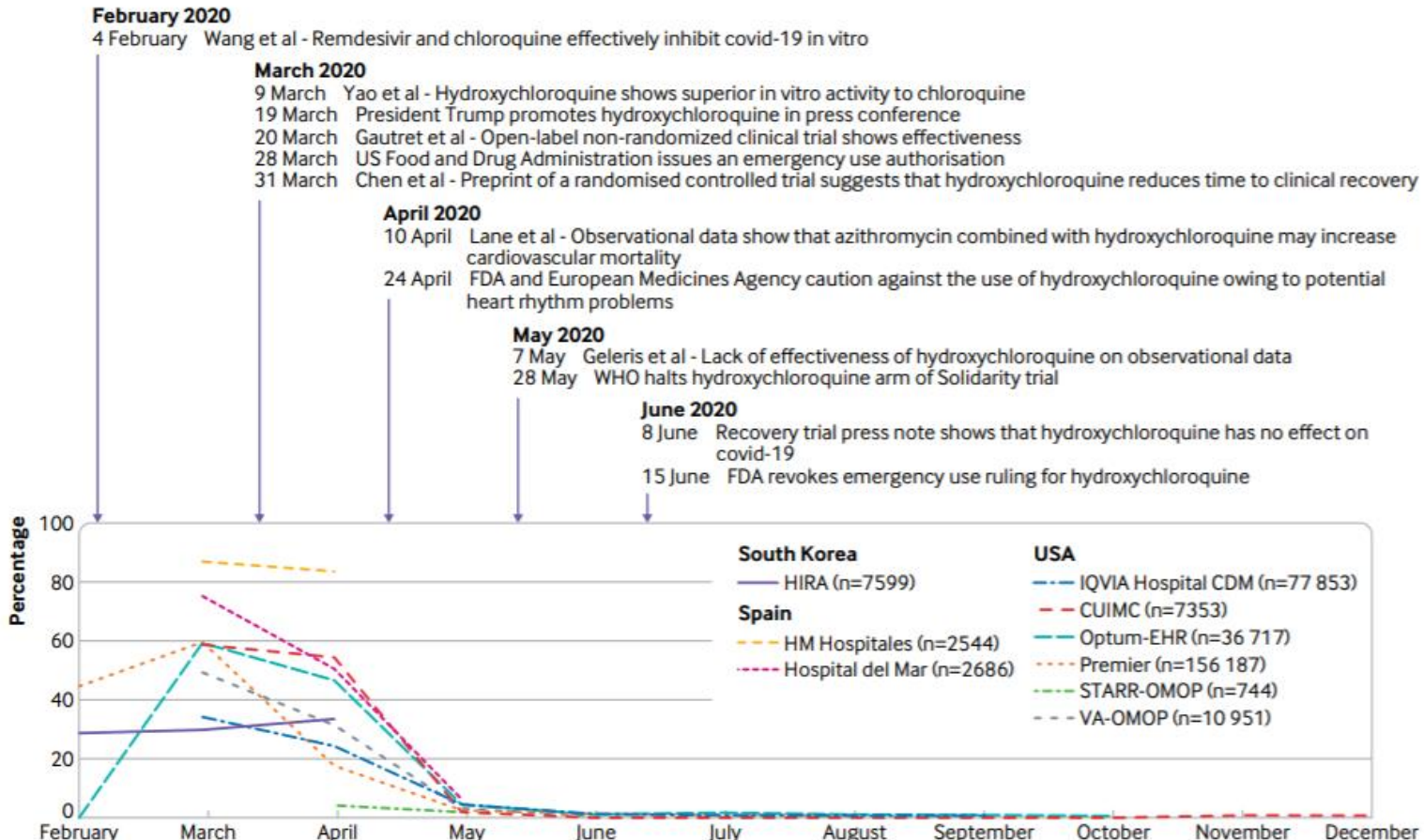
303 264 patients admitted to hospital with covid-19 from January 2020 to December 2020.

MAIN OUTCOME MEASURES

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

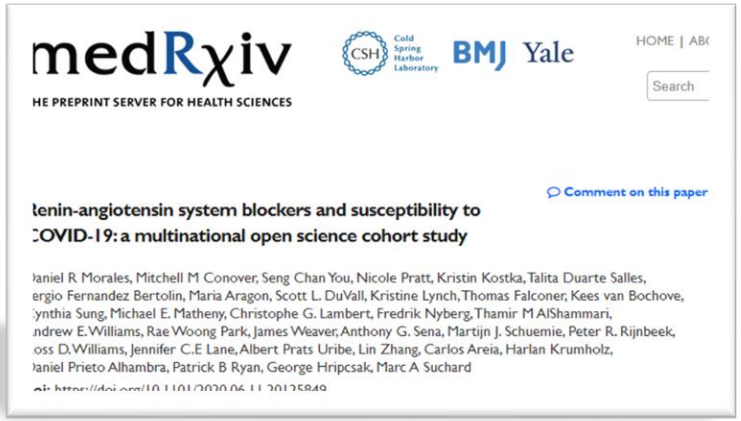
CONCLU

Multiple
the covi

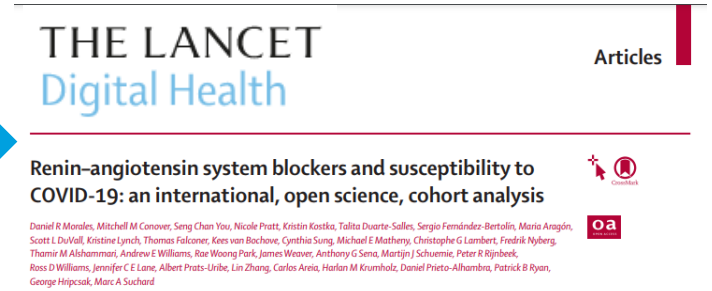


Data standards are part of methodological standards for reliable evidence

- 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.


medRxiv
HE PREPRINT SERVER FOR HEALTH SCIENCES
CSH Cold Spring Harbor Laboratory BMJ Yale
HOME | ABO
Search
Comment on this paper
lenin-angiotensin system blockers and susceptibility to
COVID-19: a multinational open science cohort study
Daniel R Morales, Mitchell M Conover, Seng Chan You, Nicole Pratt, Kristin Kostka, Talita Duarte Salles, ergio Fernandez Bertolin, Maria Aragon, Scott L. DuVall, Kristine Lynch, Thomas Falconer, Kees van Bochove, ynthia Sung, Michael E. Matheny, Christophe G. Lambert, Fredrik Nyberg, Thamir M AlShammari, ndrew E. Williams, Rae Woong Park, James Weaver, Anthony G. Sena, Martijn J. Schuemie, Peter R. Rijnbeek, oss D. Williams, Jennifer C.E Lane, Albert Prats Uribe, Lin Zhang, Carlos Areia, Harlan Krumholz, Daniel Prieto Alhambra, Patrick B Ryan, George Hripcsak, Marc A Suchard


EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH
ENCePP
European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EMA/95098/2010 Rev.8
The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology
(Revision 8)


THE LANCET
Digital Health
Articles
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

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.

Take-home messages

- Data standards should be driven by evidence needs
- Complementary standards required across the evidence generation lifecycle
  - Data standards for exchange
  - Data standards for structure for analytics
  - Vocabulary standards for data content and analytics
  - Analytic standards for evidence generation and dissemination
- Reliable evidence is strengthened by disparate data and analytics standardized within a common shared framework
  - Data network should focus on depth and breadth as required for evidence needs
  - Open science enables adoption of data standards and analytic best practices
  - International collaboration is possible with international standards