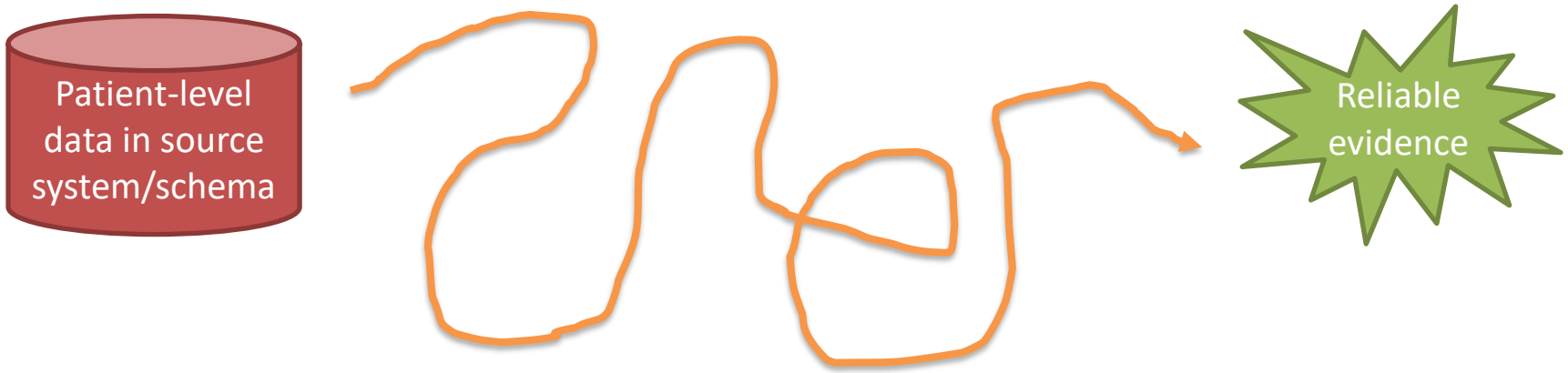
The Journey of OHDSI: Where have we been and where can we go together?

Patrick Ryan, PhD

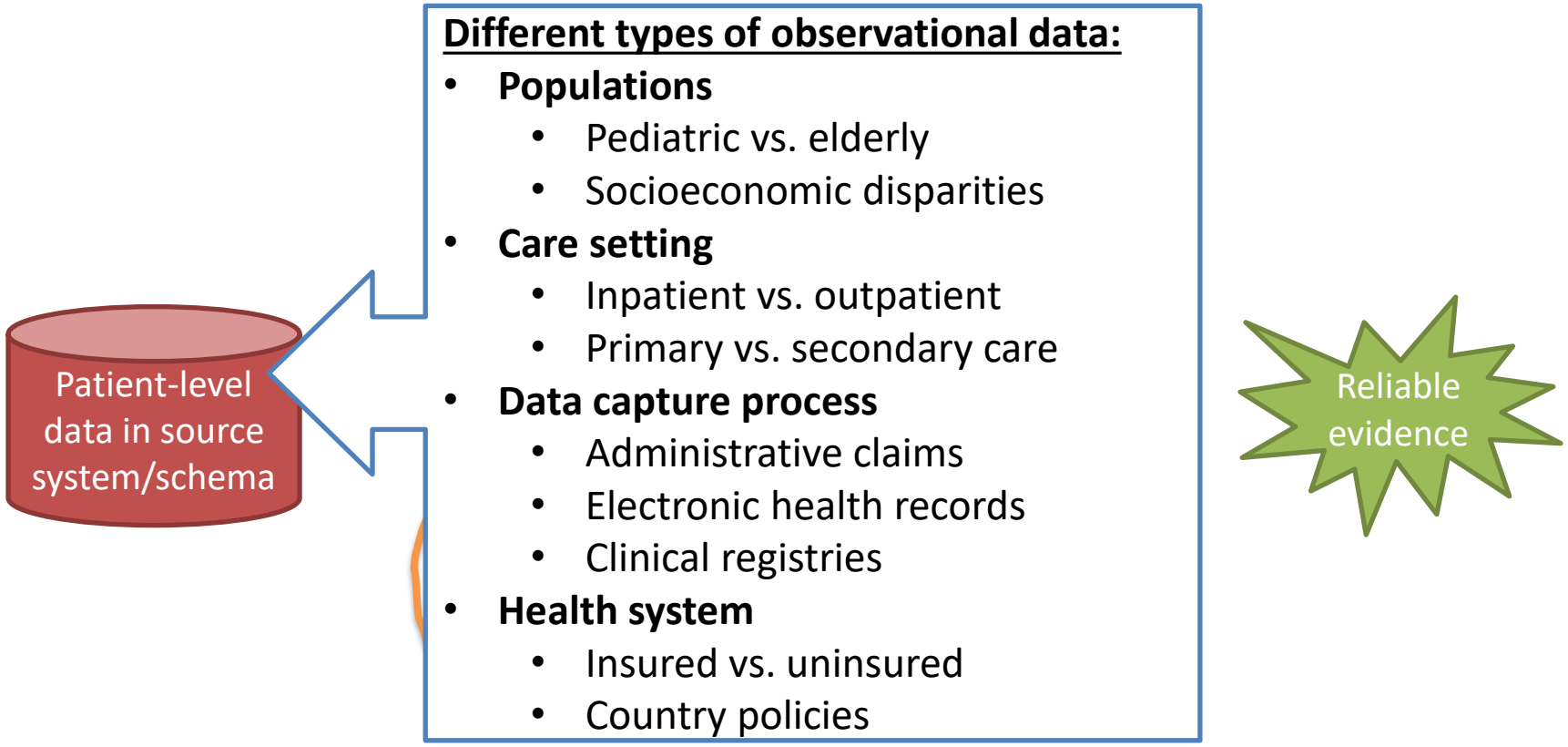
Vice President, Observational Health Data Analytics, Janssen Research and Development

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

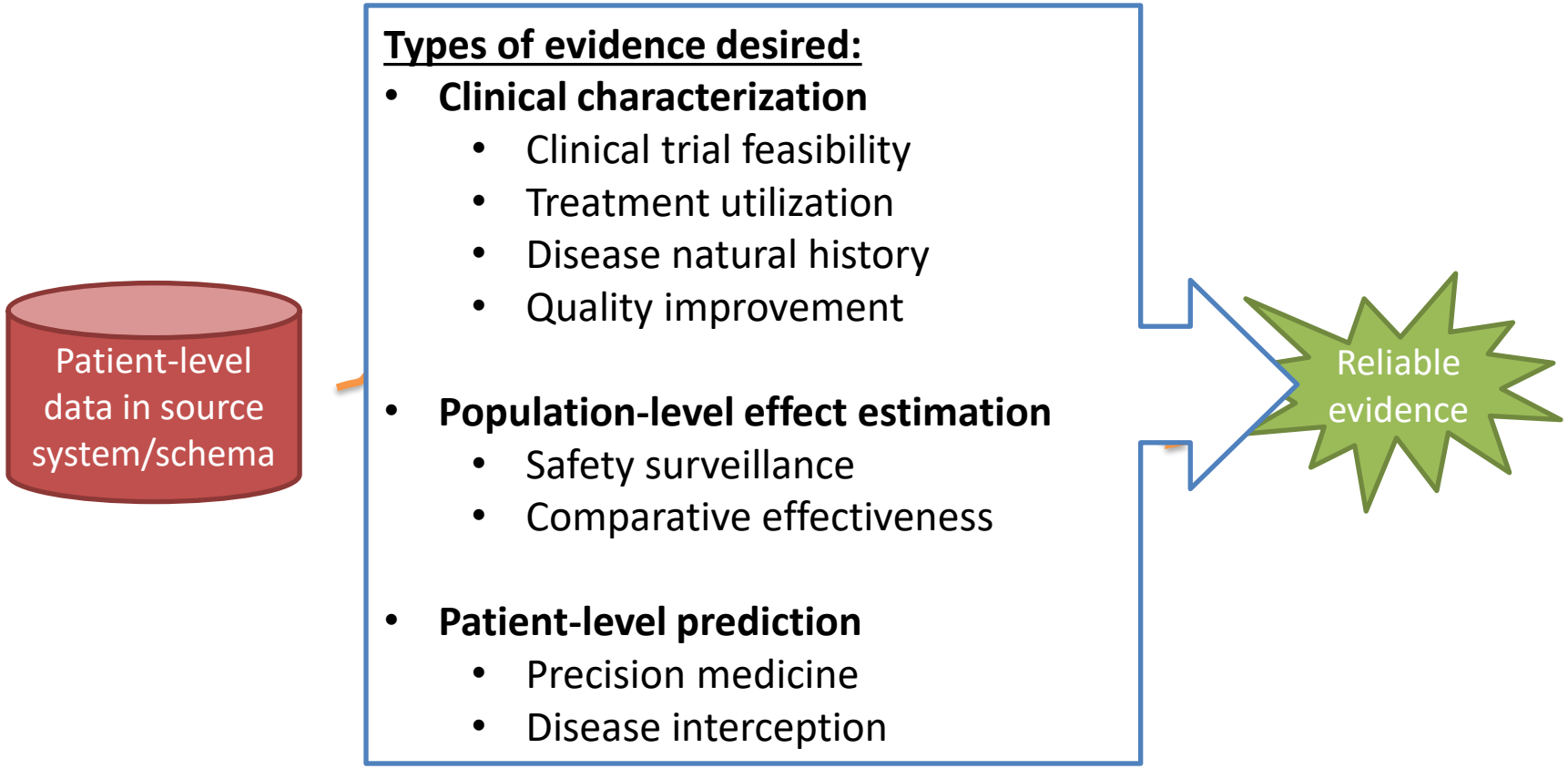
The journey to real-world evidence



The journey to real-world evidence

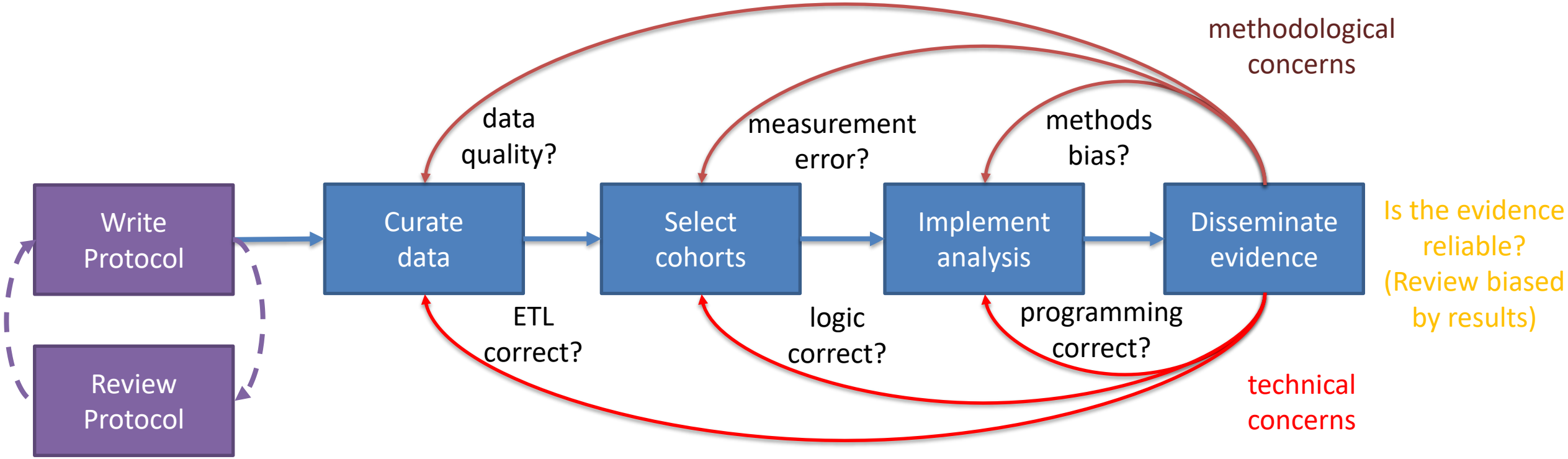


The journey to real-world evidence



Current status quo in observational research makes it challenging to build trust in evidence

Does the study provide an unbiased effect estimate?
Are the findings generalizable to the population of interest?



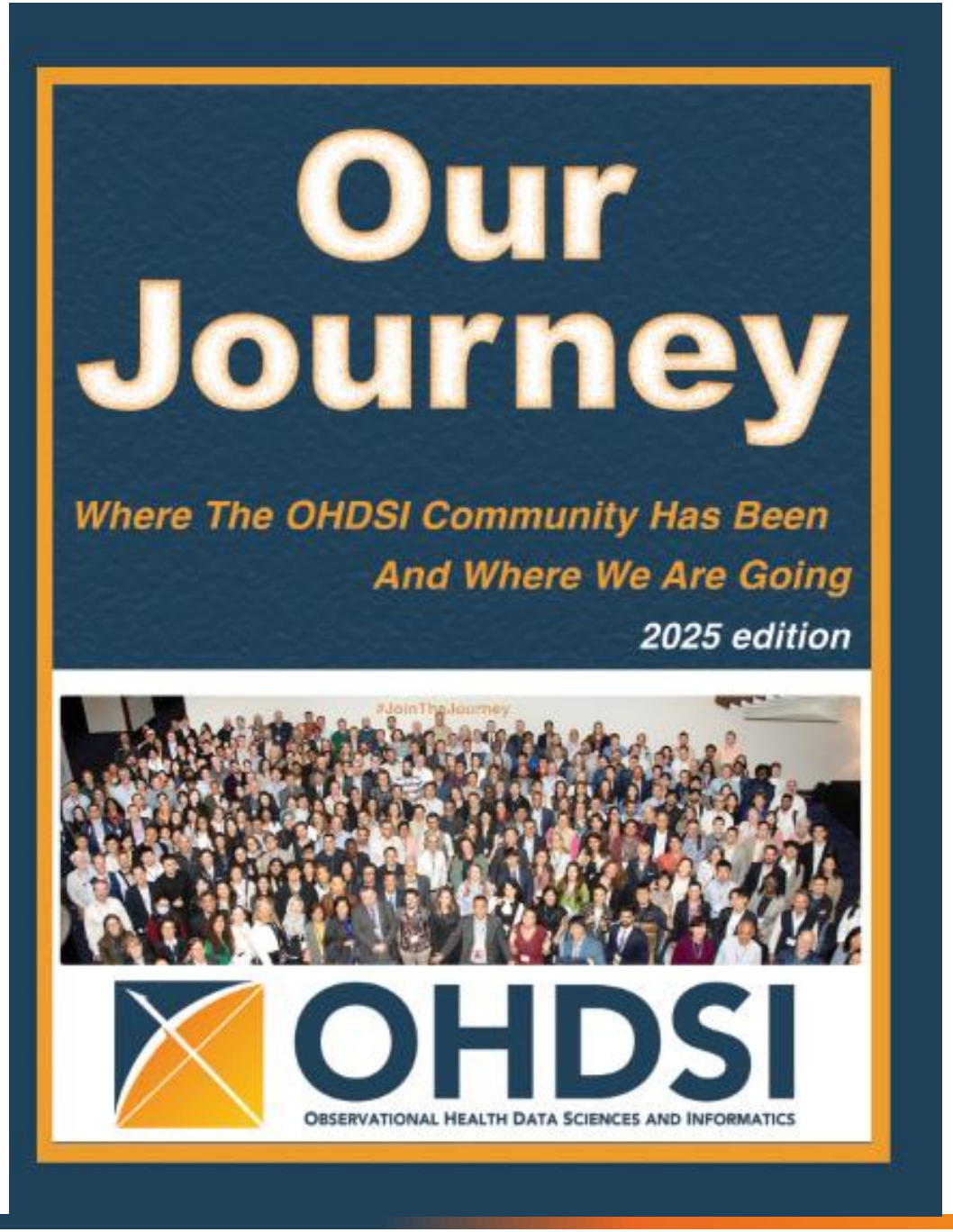
Can the study be fully reproduced?
Does the analysis actually do what the protocol said it would do?

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

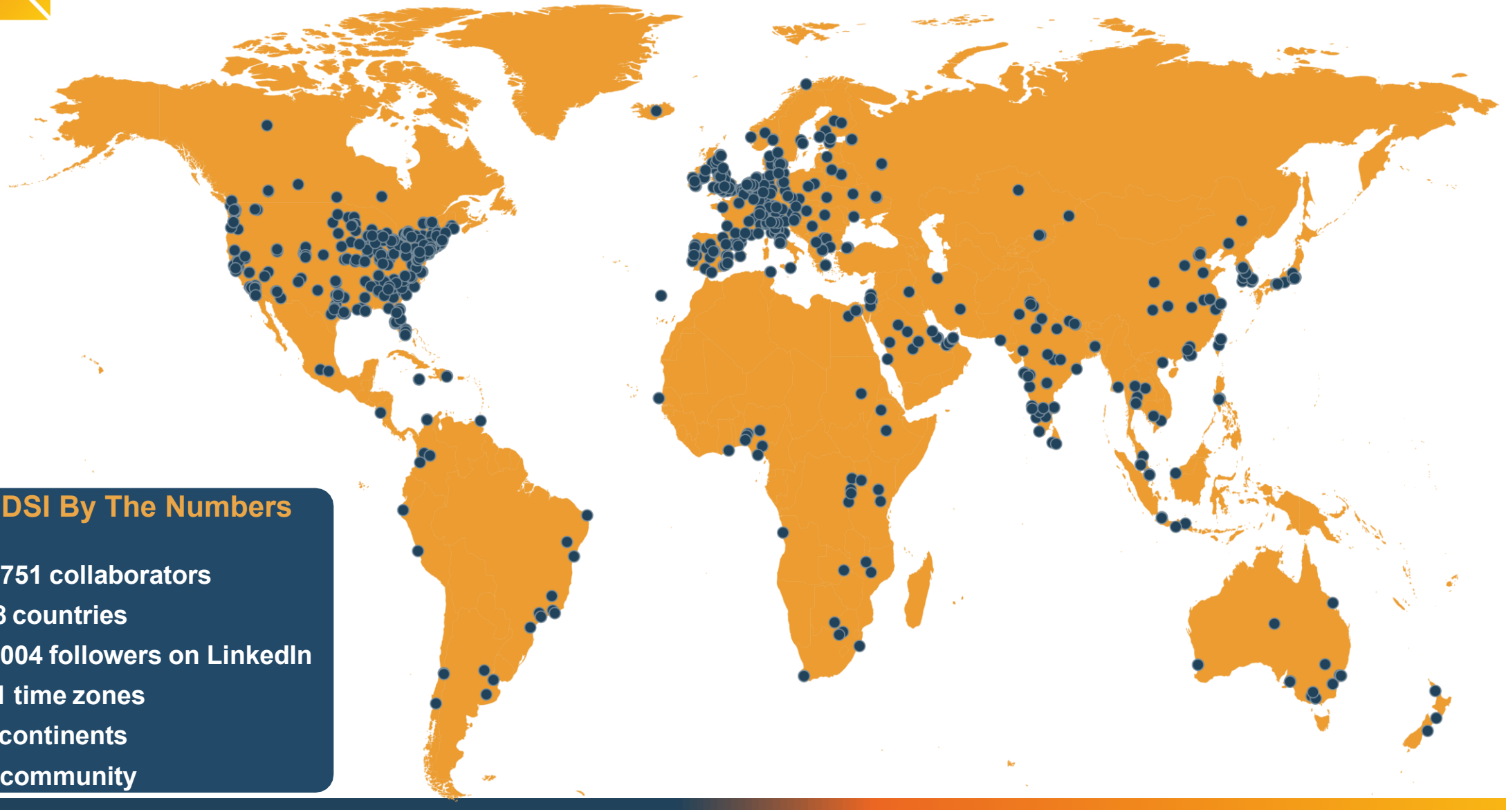
OHDSI's mission

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



https://www.ohdsi.org/wp-content/uploads/2025/10/OurJourney-2025.pdf

OHDSI collaborators



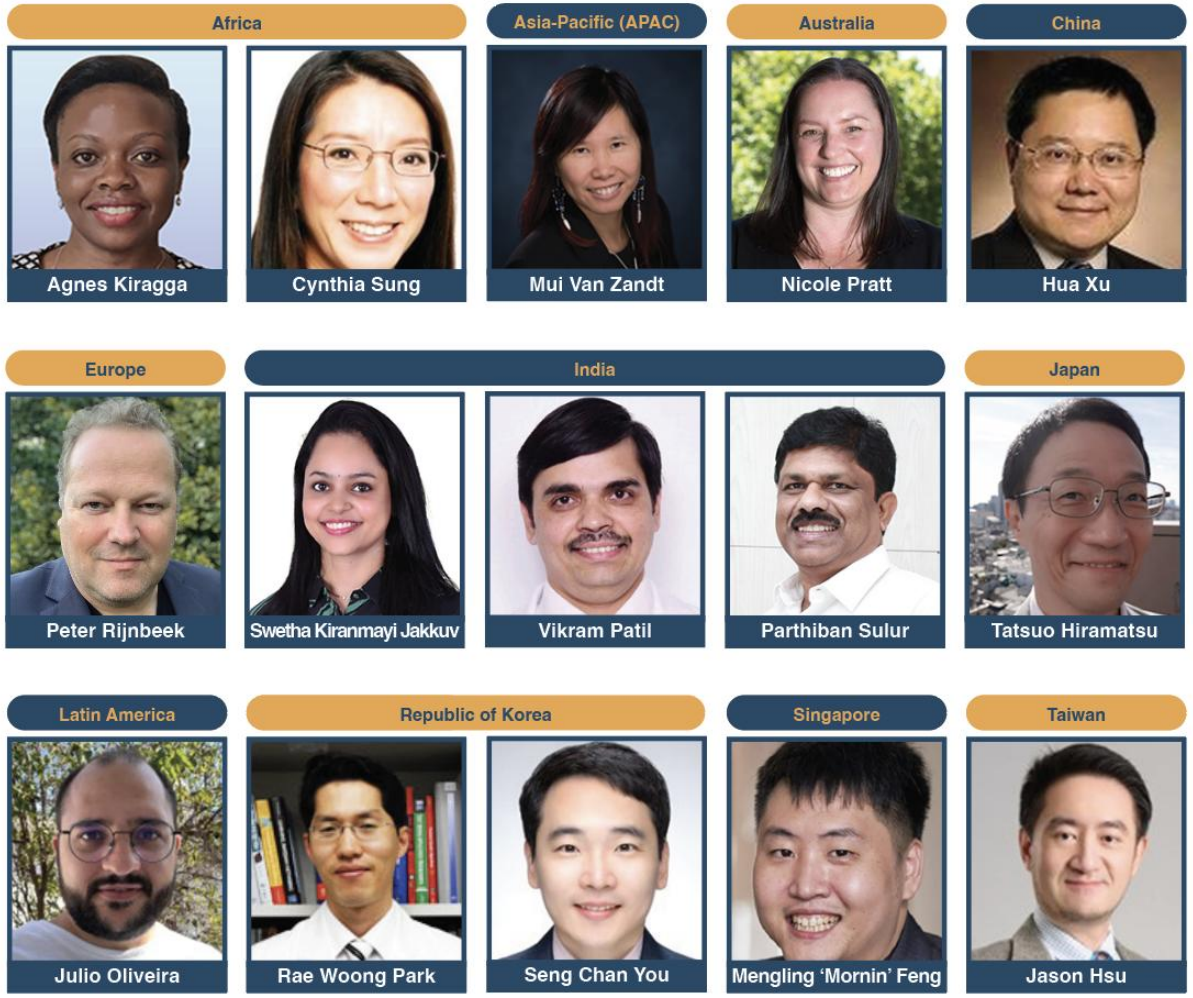
OHDSI By The Numbers

- 4,751 collaborators
- 88 countries
- 9,004 followers on LinkedIn
- 21 time zones
- 6 continents
- 1 community

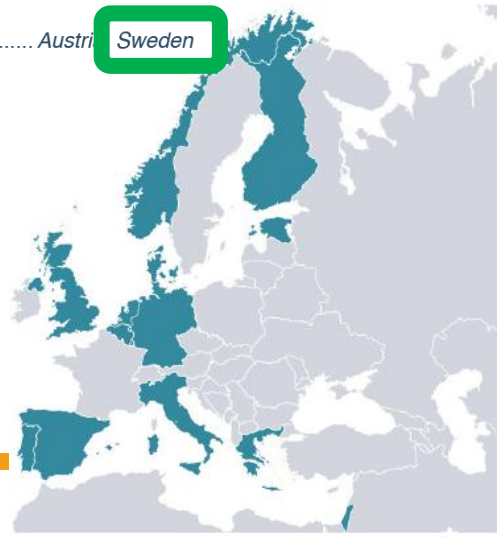
Join the Journey at https://ohdsi.org/

Workgroups led by community

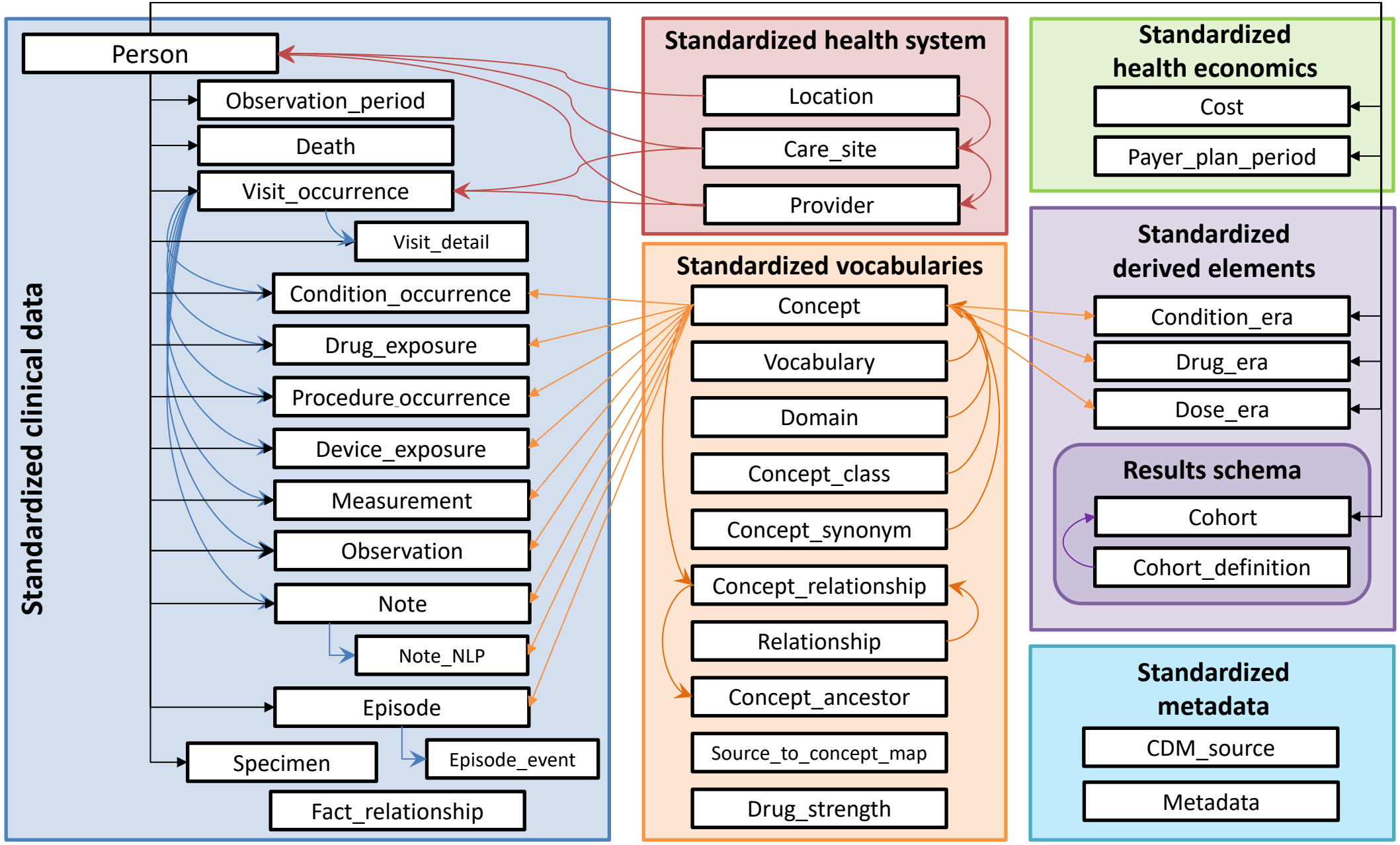
Regional chapters and national nodes



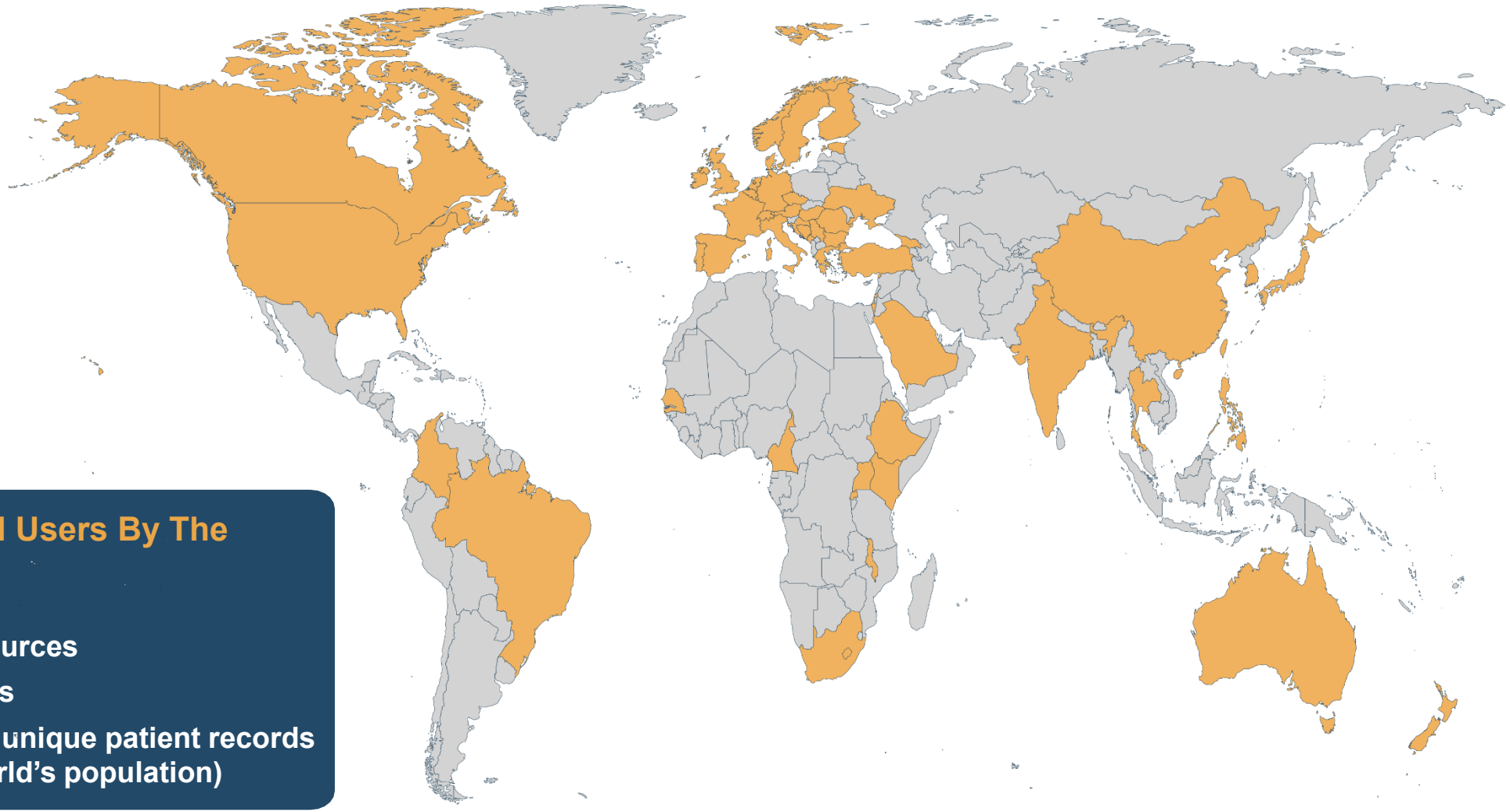
Node	Lead(s)
Belgium	Liesbet Peeters, Annelies Verbiest, Ilse Vermeulen
Denmark	Ismail Gögenur, Martin Høyer Rose, Andreas Weinberger Rosen
Estonia	Raivo Kolde, Sulev Reisberg
Finland	Eric Fey, Gustav Klingstedt
Germany	Ines Reinecke, Michele Zoch
Greece	Anastasia Farmaki, Pantelis Natsiavas,Grigoris Papapostolou
Hungary	Zsolt Bagyura, Ágota Mészáros
Ireland	Aedin Culhane, Mark Lawler, Catherine Mahoney
Israel	Chen Yanover
Italy	Lucia Sacchi, Matteo Gabetta
Luxembourg	Claudine Backes, Andreas Kremer, Maria Quaranta
Netherlands	Renske Los, Aniek Markus
Norway	Espen Enerly, Siri Larønningen
Portugal	Patricia Couceiro, Carmen Nogueira
Spain	Miguel Angel Mayer, Talita Duarte Salles
Switzerland	Olga Endrich, Karen Triep
United Kingdom	Dani Prieto-Alhambra
coming soon	*Austri* *Sweden*



OMOP Common Data Model v5.4



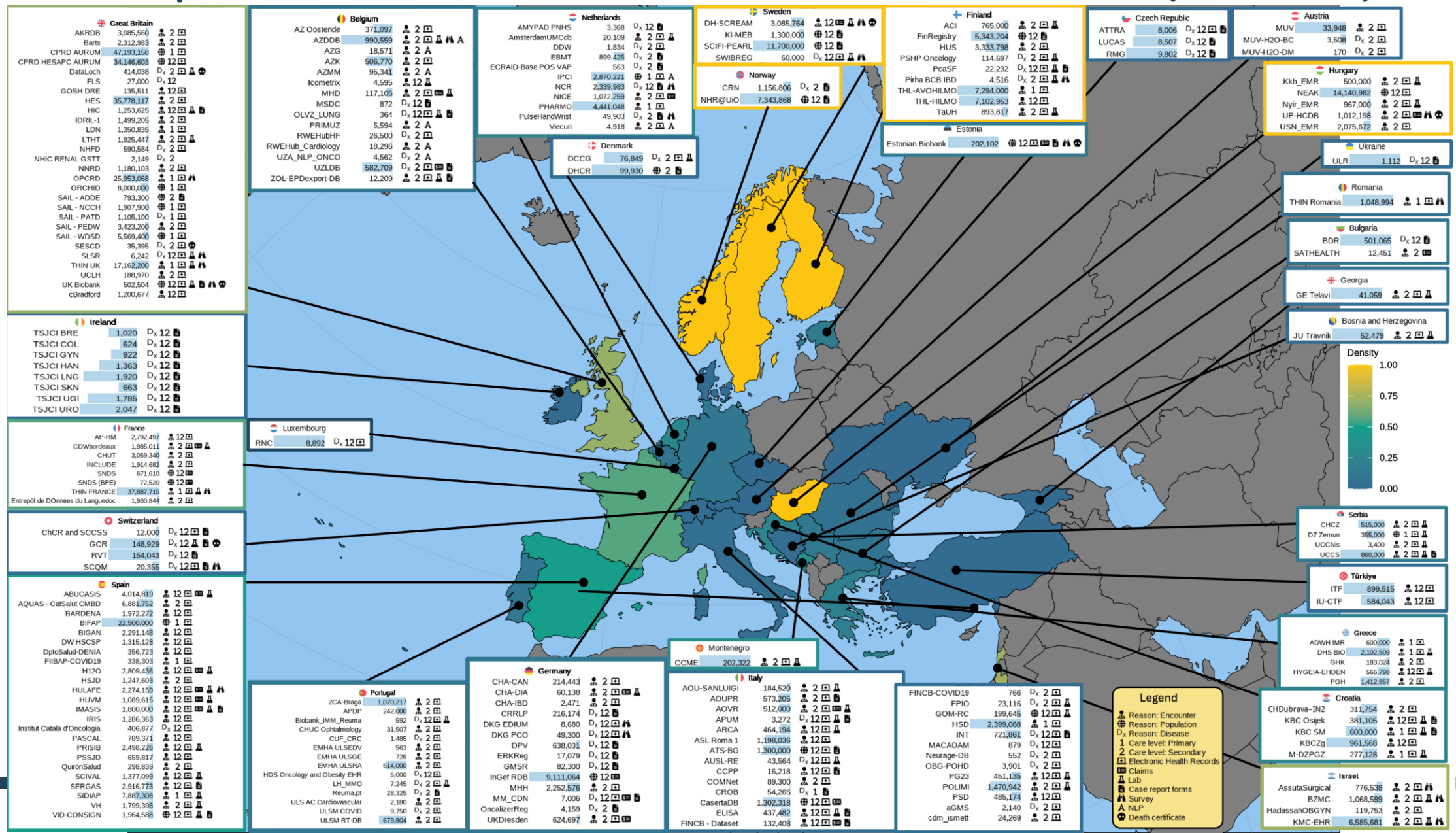
OMOP Common Data Model adoption



OMOP CDM Users By The Numbers

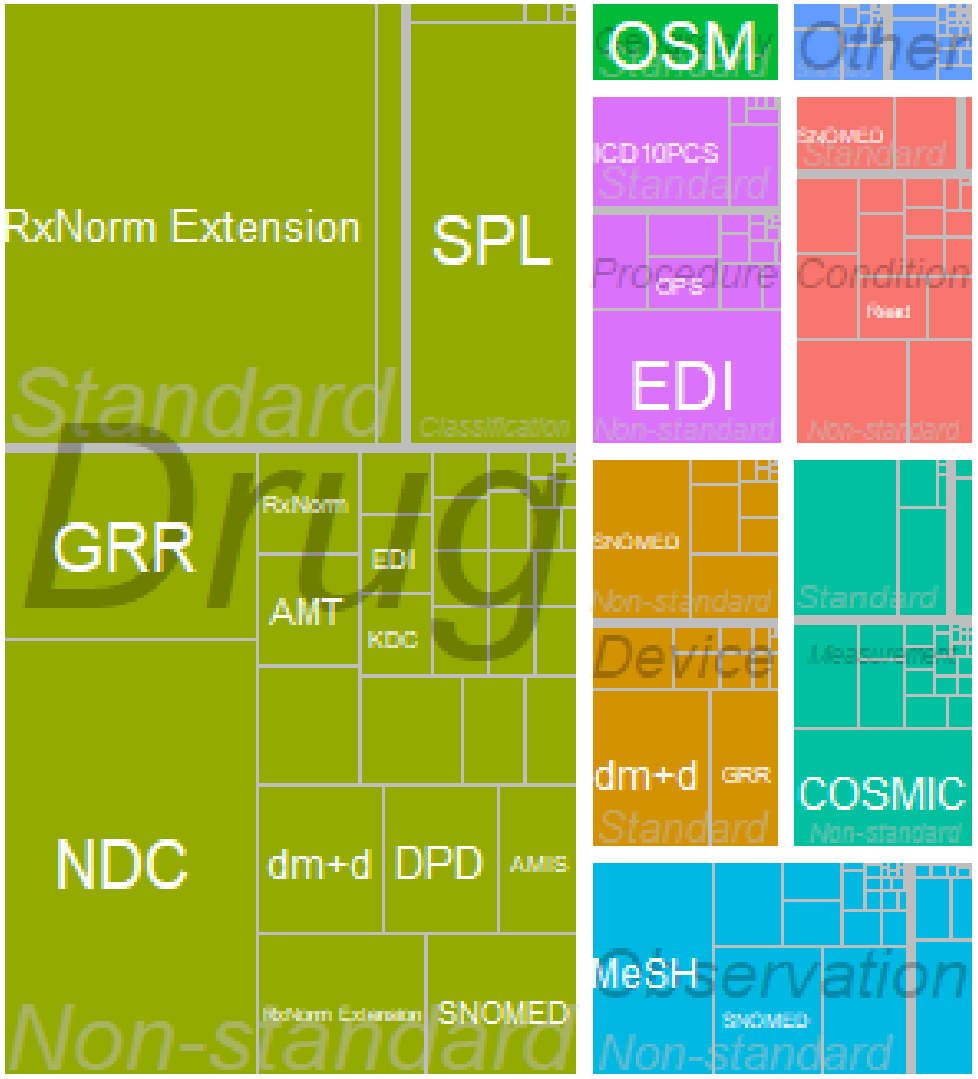
- 544 data sources
- 54 countries
- 974 million unique patient records (12% of world's population)

European Health Data & Evidence Network (EHDEN)



Blacketer
JMIR
2025

OHDSI standardized vocabularies

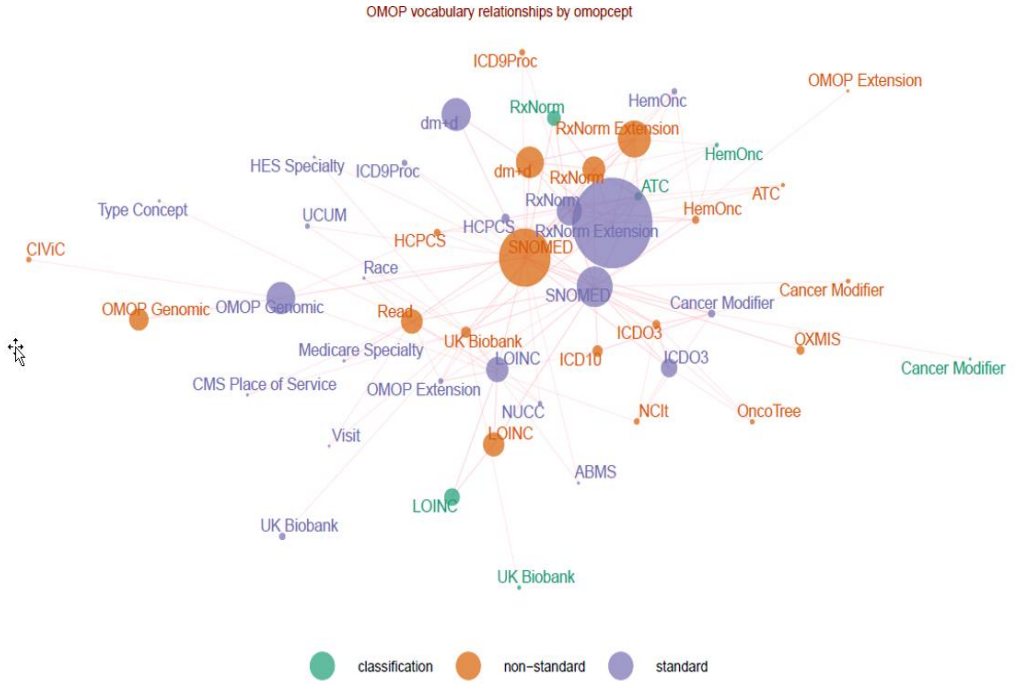


OHDSI Vocabularies By The Numbers

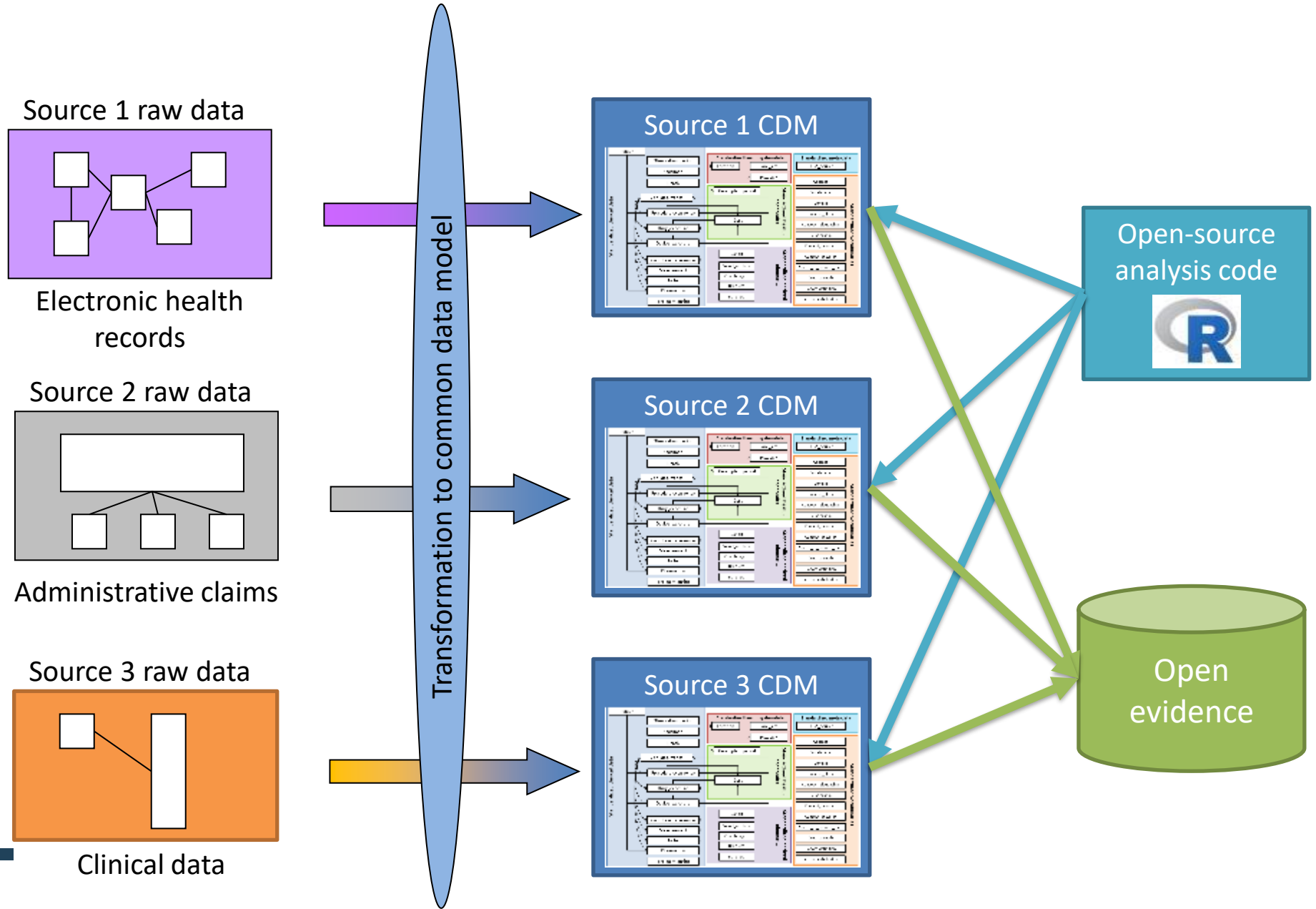
as of August 2025 release

- 11,804,307 concepts
  - 3,784,263 standard concepts
  - 971,914 classification concepts
- 145 vocabularies
- 43 domains
- 87,948,636 concept relationships
- 101,696,159 ancestral relationships
- 6,028,711 concept synonyms

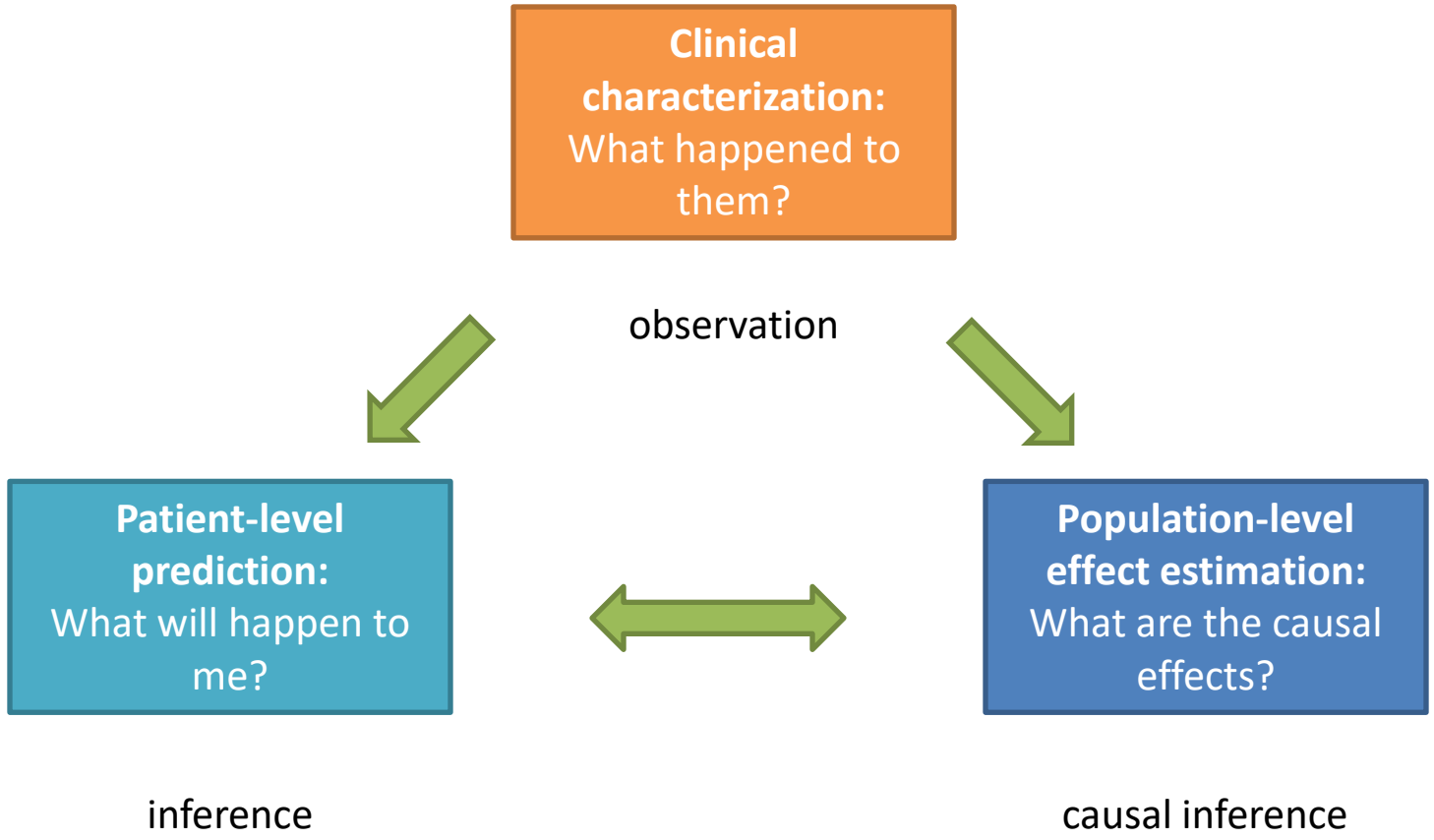
1 Shared Resource to Enable Data Standards



Common data model can enable standardized analytics across a distributed data network



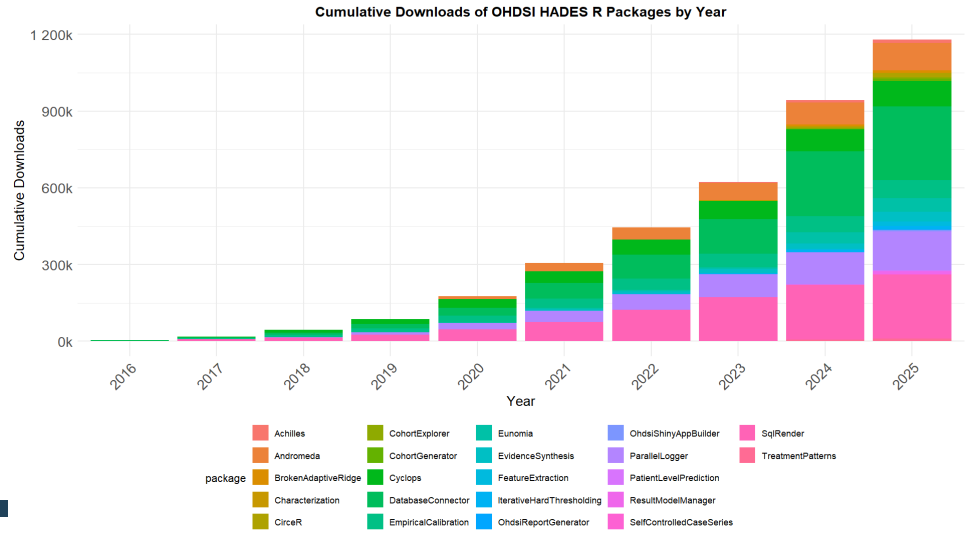
Complementary evidence to inform the patient journey



Analytic use case	Type	Structure	Example
Clinical characterization	**Disease Natural History**	Amongst patients who are diagnosed with **<insert your favorite disease>**, what are the patient's characteristics from their medical history?	Amongst patients with **rheumatoid arthritis**, what are their demographics (age, gender), prior conditions, medications, and health service utilization behaviors?
	Treatment utilization	Amongst patients who have **<insert your favorite disease>**, which treatments were patients exposed to amongst **<list of treatments for disease>** and in which sequence?	Amongst patients with **depression**, which treatments were patients exposed to **SSRI, SNRI, TCA, bupropion, esketamine** and in which sequence?
	Outcome incidence	Amongst patients who are new users of **<insert your favorite drug>,** how many patients experienced **<insert your favorite known adverse event from the drug profile>** within **<time horizon following exposure start>**?	Amongst patients who are new users of **methylphenidate,** how many patients experienced **psychosis** within **1 year of initiating treatment**?
Population-level effect estimation	**Safety surveillance**	Does exposure to **<insert your favorite drug>** increase the risk of experiencing **<insert an adverse event>** within **<time horizon following exposure start>**?	Does exposure to **ACE inhibitor** increase the risk of experiencing **Angioedema** within **1 month after exposure start**?
	Comparative effectiveness	Does exposure to **<insert your favorite drug>** have a different risk of experiencing **<insert any outcome (safety or benefit) >** within **<time horizon following exposure start>**, relative to **<insert your comparator treatment>**?	Does exposure to **ACE inhibitor** have a different risk of experiencing **acute myocardial infarction** while **on treatment**, relative to **thiazide diuretic**?
Patient level prediction	**Disease onset and progression**	For a given patient who is diagnosed with **<insert your favorite disease>**, what is the probability that they will go on to have **<another disease or related complication>** within **<time horizon from diagnosis>?**	For a given patient who is **newly diagnosed with atrial fibrillation**, what is the probability that they will go onto to have **ischemic stroke** in **next 3 years**?
	Treatment response	For a given patient who is a new user of **<insert your favorite chronically-used drug>**, what is the probability that they will **<insert desired effect>** in **<time window>**?	For a given patient **with T2DM who start on metformin** , what is the probability that they will **maintain HbA1C<6.5%** after **3 years**?
	Treatment safety	For a given patient who is a new user of **<insert your favorite drug>,** what is the probability that they will experience **<insert adverse event >** within **<time horizon following exposure>**?	For a given patients who is a **new user of warfarin**, what is the probability that they will have **GI bleed** in **1 year**?

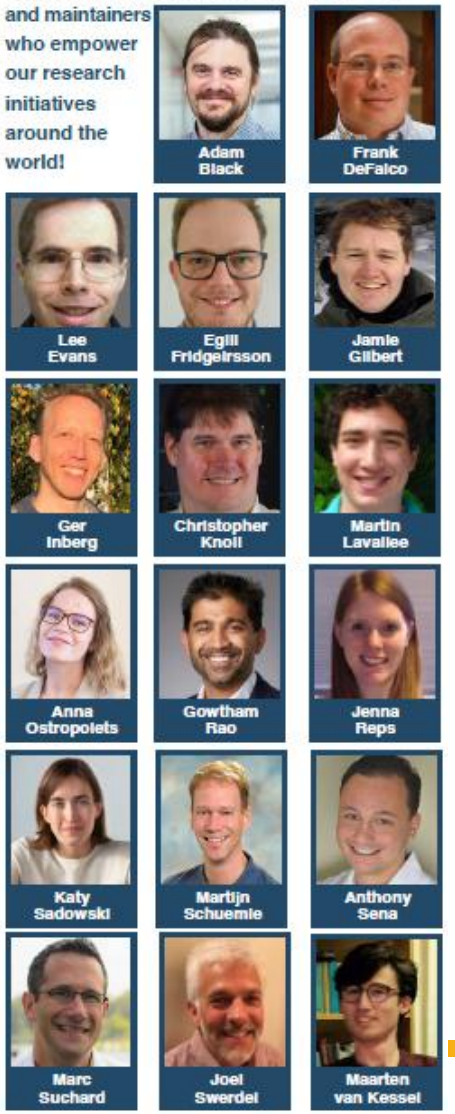
Open-source software development

- HADES is an ecosystem of 41 R packages to support standardized analytics for the OMOP CDM and across OHDSI network
- OHDSI CRAN packages (n=22) have been downloaded >1m times

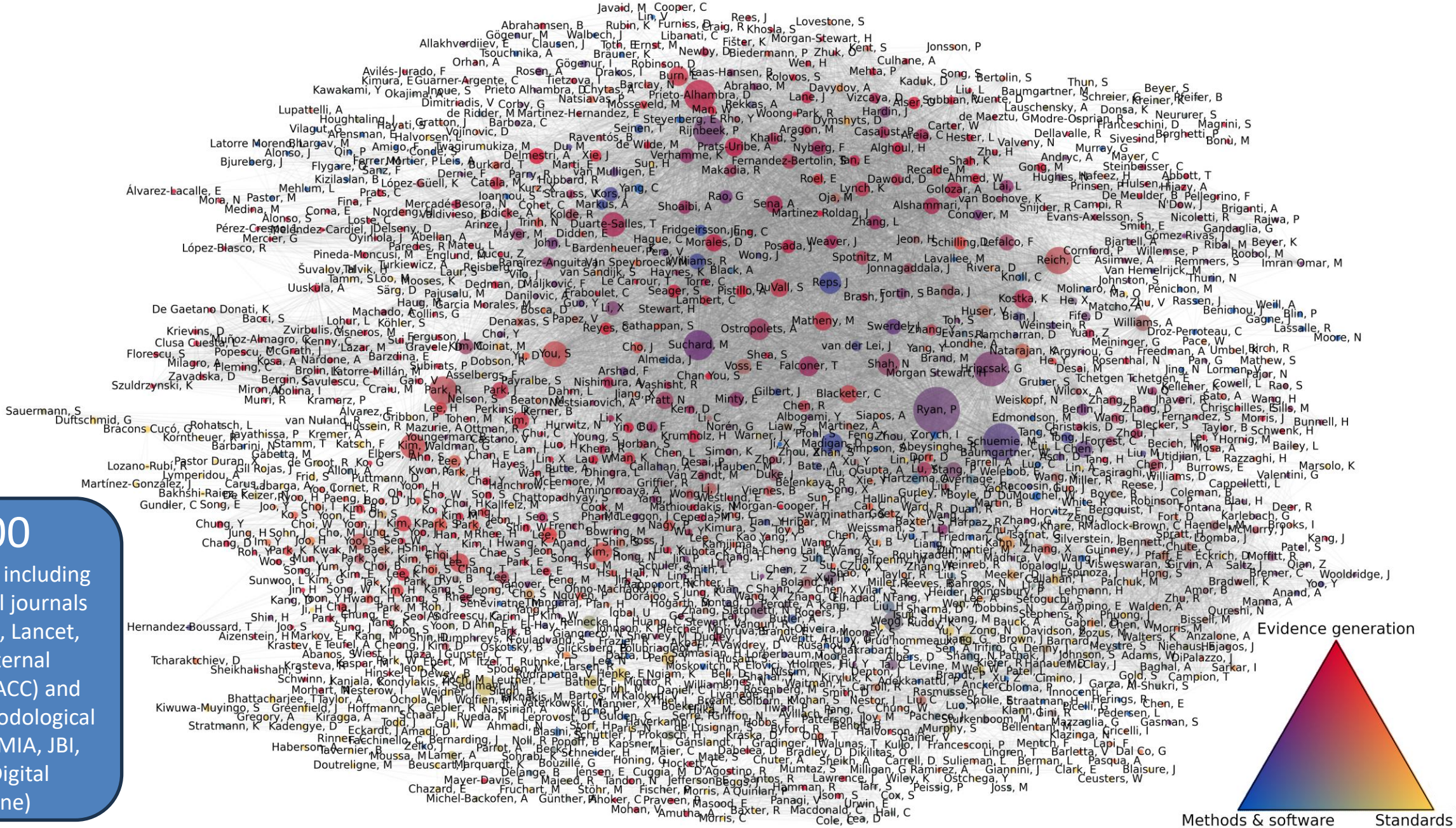


Package	Version	Maintainer(s)	Availability
Achilles	v8.7.2	Frank DeFalco	CRAN
Andromeda	v0.6.1	Martijn Schuemie	CRAN
BigKnn	v1.3.2	Martijn Schuemie	GitHub
BrokenAdaptiveRidge	v1.3.1	Marc Suchard	CRAN
Cyclops	v2.1.0	Martin Lavallee	GitHub
Characterization	v2.2.0	Jenna Reps	CRAN
CirceR	v1.3.3	Chris Knoll	CRAN
CohortDiagnostics	v3.4.2	Jamie Gilbert	GitHub
CohortExplorer	v0.1.0	Gowtham Rao	CRAN
CohortGenerator	v0.12.2	Anthony Sena	CRAN
CohortIncidence	v4.1.0	Chris Knoll	GitHub
CohortMethod	v5.5.0	Martijn Schuemie	GitHub
Cyclops	v3.6.0	Marc Suchard	CRAN
DatabaseConnector	v6.4.0	Martijn Schuemie	CRAN
DataQualityDashboard	v2.7.0	Katy Sadowksi	GitHub
DeepPatientLevelPrediction	v2.2.0	Egill Fridgeirsson	GitHub
EmpiricalCalibration	v3.1.4	Martijn Schuemie	CRAN
EnsemblePatientLevelPrediction	v1.9.2	Jenna Reps	GitHub
Eunomia	v2.1.0	Frank DeFalco	CRAN
EvidenceSynthesis	v1.0.0	Martijn Schuemie	CRAN
FeatureExtraction	v3.11.0	Ger Inberg	CRAN
Hydra	v0.4.0	Anthony Sena	Deprecated
IterativeHardThresholding	v1.0.3	Marc Suchard	CRAN
Keeper	v0.2.1	Anna Ostropolets	GitHub
MethodEvaluation	v2.4.0	Martijn Schuemie	GitHub
OhdsiReportGenerator	v1.1.1	Jenna Reps	CRAN
OhdsiSharing	v0.2.2	Lee Evans	GitHub
OhdsiShinyAppBuilder	v1.0.0	Jenna Reps	CRAN
OhdsiShinyModules	v1.3.0	Jenna Reps	GitHub
ParallelLogger	v3.5.0	Martijn Schuemie	CRAN
PatientLevelPrediction	v6.5.0	Egill Friogeirsson & Jenna Reps	CRAN
PhenotypeLibrary	v1.16.0	Gowtham Rao	GitHub
Phevaluator	v2.2.15	Joel Swerdel	GitHub
ResultModelManager	v0.5.12	Jamie Gilbert	CRAN
ROhdsiWebApi	v1.3.3	Gowtham Rao	GitHub
SelfControlledCaseSeries	v6.1.0	Martijn Schuemie	CRAN
SelfControlledCohort	v1.6.0	Jamie Gilbert	GitHub
ShinyAppBuilder	v2.3.0	Jenna Reps	Deprecated
SqlRender	v1.16.3	Martijn Schuemie	CRAN
Strategus	v1.4.1	Anthony Sena	GitHub
TreatmentPatterns	v3.1.1	Maarten van Kessel	CRAN

The open-source tools that empower OHDSI research are not only available to the community, but they are DEVELOPED by the community. We thank the many developers and maintainers who empower our research initiatives around the world!



OHDSI collaborations in scholarship

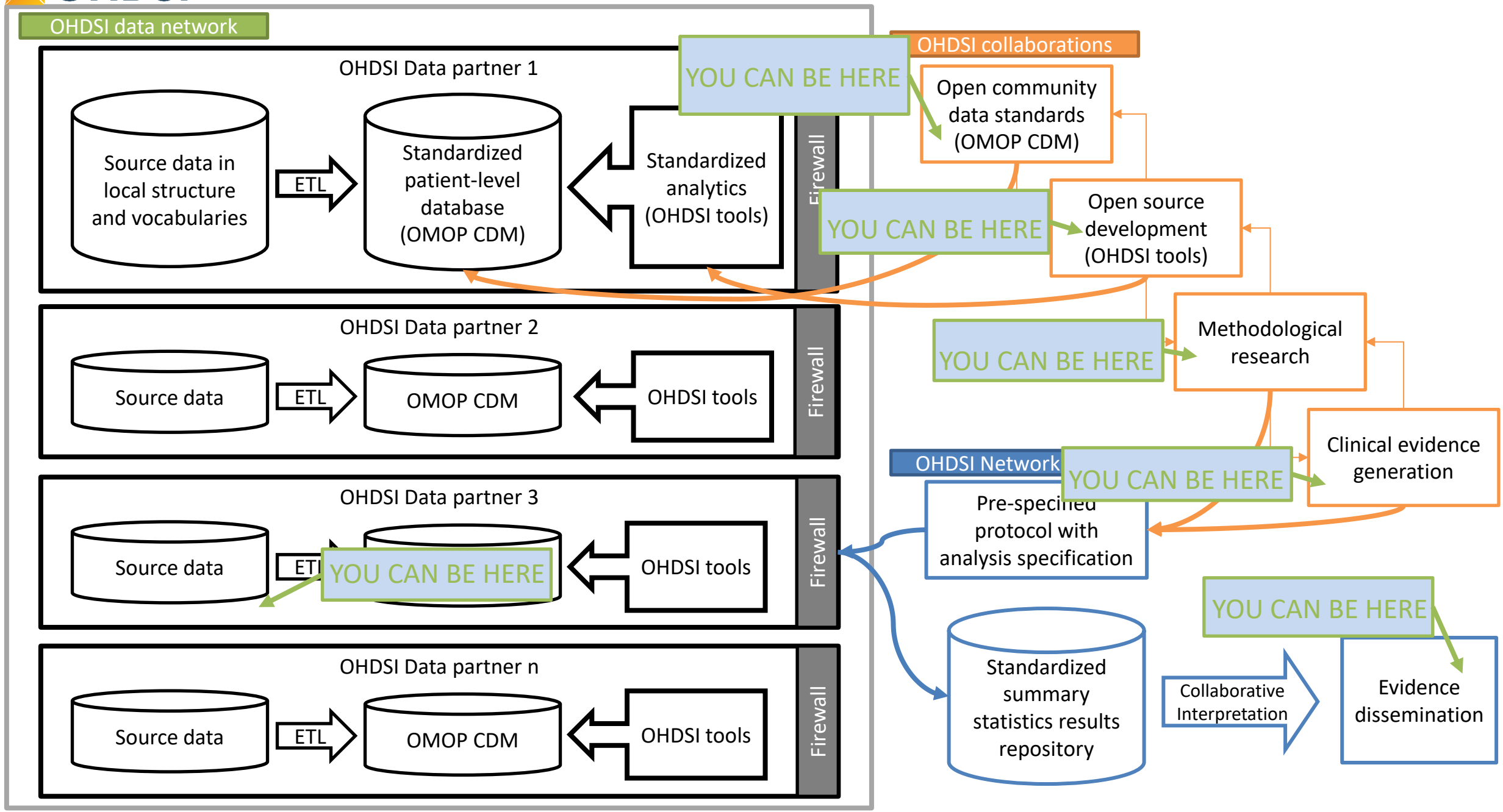


>900

publications, including in top clinical journals (JAMA, BMJ, Lancet, JAMA Internal Medicine, JACC) and leading methodological journals (JAMIA, JBI, Nature Digital Medicine)

OHDSI

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



Why OHDSI and the world needs Sweden

The Legatum Prosperity Index™ Ranks 1–56

2013 rank	2022 rank	2023 rank	Country	Health	Education	Natural Environment
4	1	1	Denmark	16	5	5
1	3	2	Sweden	9	8	1
2	2	3	Norway	7	10	9
5	4	4	Finland	15	2	2
3	5	5	Switzerlan	10	4	6
6	6	6	Netherlan	11	6	39
8	7	7	Luxembou	12	36	10
14	9	8	Iceland	8	11	13
9	8	9	Germany	13	19	12
7	10	10	New Zeala	25	18	11

European Journal of Epidemiology (2025) 40:563–579
https://doi.org/10.1007/s10654-025-01226-9

DATA RESOURCE

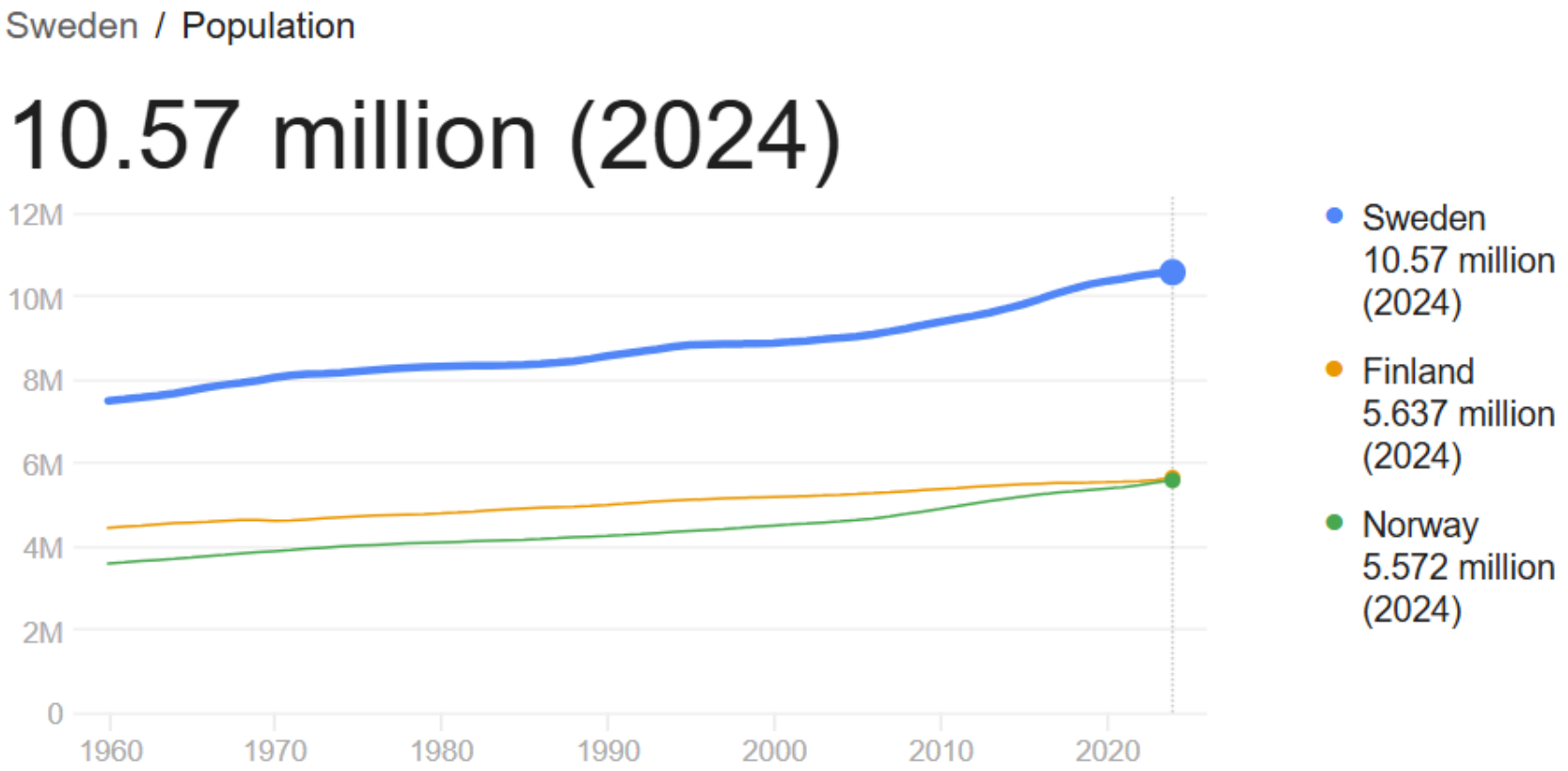
The healthcare system in Sweden

Jonas F. Ludvigsson[1,2,3] · David Bergman[1] · Catharina Ihre Lundgren[4,5] · Kristina Sundquist[6] · Jean-Luc af Geijerstam[7] · Anna H. Glenngård[8] · Marie Lindh[9] · Johan Sundström[10,11] · Johan Kaarme[12,13] · Jialu Yao[1]

Received: 10 March 2025 / Accepted: 23 March 2025 / Published online: 19 May 2025

Abstract
The Swedish population is characterized by high life expectancy and low avoidable mortality rates. This review outlines the Swedish healthcare system, which offers universal access to all residents and has a long tradition of reforms for social equity. Responsibility for healthcare is shared between the state, the regions, and the municipalities. The Ministry of Health and Social Affairs provides the overall healthcare framework; additionally, several governmental agencies are directly involved in healthcare and public health initiatives. The 21 regions organize, finance, and provide most primary, secondary, and tertiary care, as well as health information channels. Resources for primary care are less plentiful than in many other countries. The 290 municipalities deliver care to elderly people and those with functional impairment. The Swedish healthcare system is primarily tax-funded, with 86% of total healthcare expenditures from public expenses and <1% from voluntary health insurance. The gross domestic product (GDP) share of healthcare expenditures, 10.5% in 2022, is above the EU average. The level of unmet needs in the population is low, due to universal coverage and caps on user charges except for dental care. Sweden's healthcare system performs well on care quality and patient satisfaction, but suffers from workforce shortage and care fragmentation. Limitations in care coordination can be attributed to a siloed digital infrastructure and care governance, a low number of hospital beds per capita, and a compensation system that often does not incentivize coordination. Despite these challenges, life expectancy is high and avoidable mortality rates are low in Sweden.

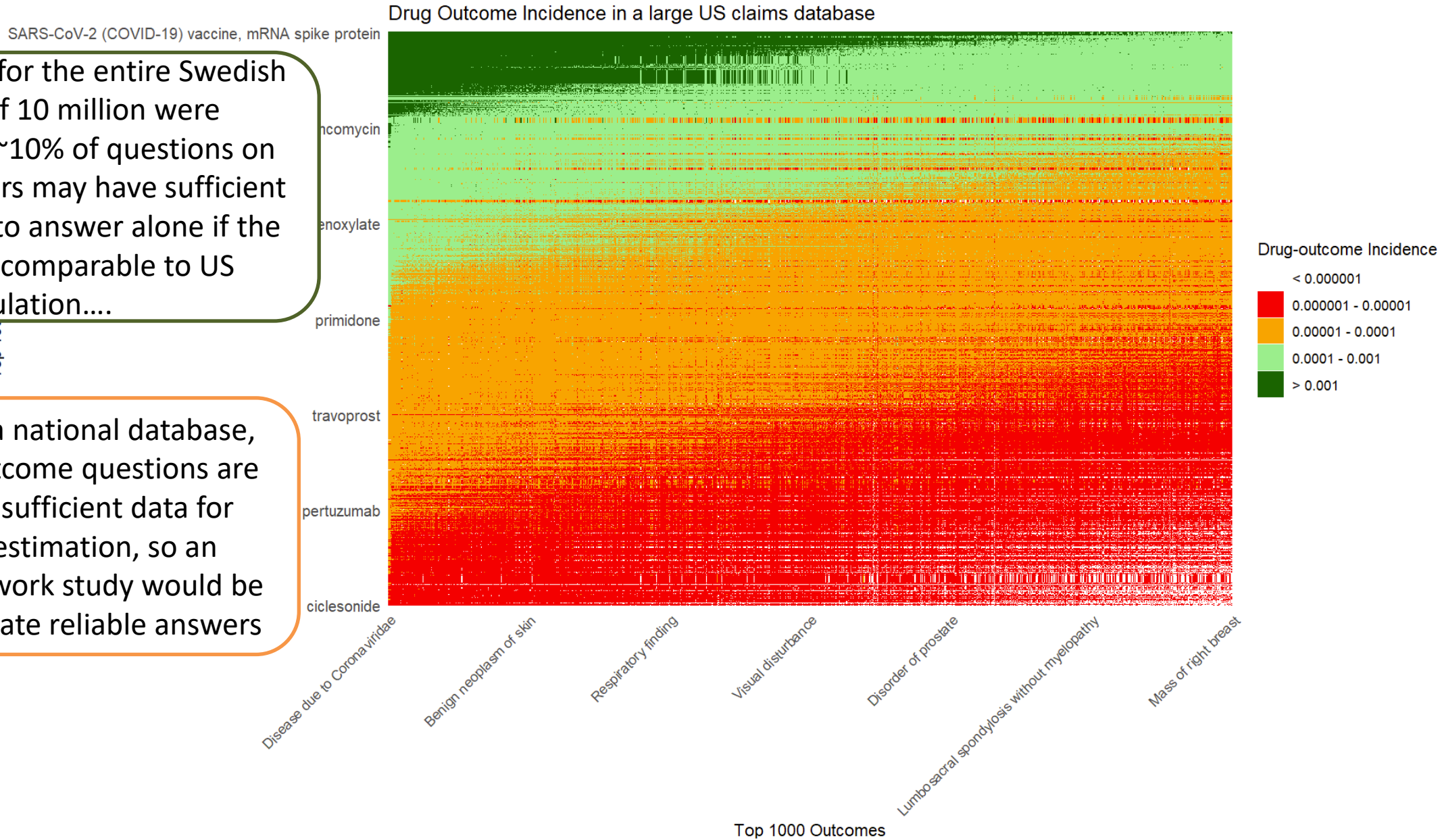
Why Sweden needs OHDSI and the rest of the world



Why Sweden needs OHDSI and the rest of the world

If full health data for the entire Swedish population of 10 million were accessible, then ~10% of questions on drug-outcome pairs may have sufficient statistical power to answer alone if the incidence are comparable to US population….

...but even with a national database, >70% of drug-outcome questions are likely to have insufficient data for prediction or estimation, so an international network study would be required to generate reliable answers



Join the Journey!

https://www.ohdsi.org/join-the-journey/

Looking forward to collaborating with all of you!

Email me: ryan@ohdsi.org