Will phase IV trials become obsolete?

Transforming routinely-collected health data into real-world evidence

Annelies Verbiest MD PhD

annelies.verbiest@uza.be

UZA medical oncology service

UZA Data Innovation Lab

OHDSI slides courtesy of Renske Los & Maxime Moinat

Presented at EUFEMED-Healixia conference

22-05-2025

Disclosures

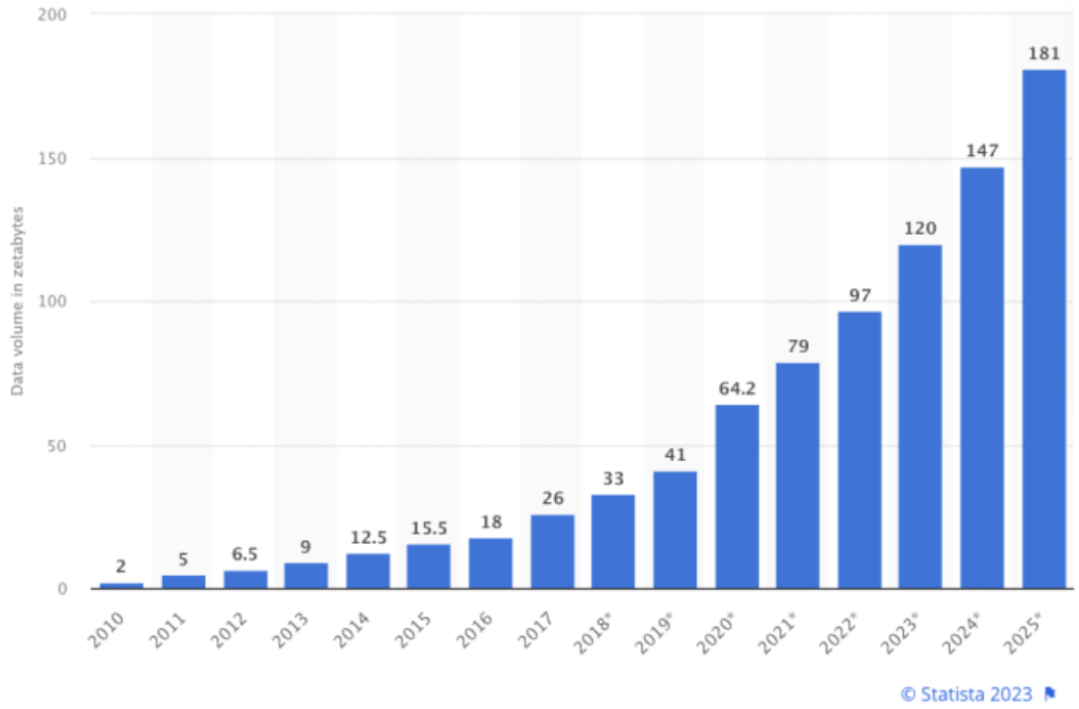
None related to this talk

- Speaker fees: Novartis, Ipsen
- Advisory board & congress: Ipsen, Astra Zeneca

All fees institutional

The healthdata explosion

- The healthcare sector has the fastest growing datasphere
  - ➢ Clinical data, radiology, pathology, -omics, monitoring, wearables…
- Investment in qualitative data and insights lags behind that of other sectors



What if…

The care for every citizen/patient is personalized, based on all known information about the patient and the outcomes of billions of patients around the world?

Population-Level Effect Estimation

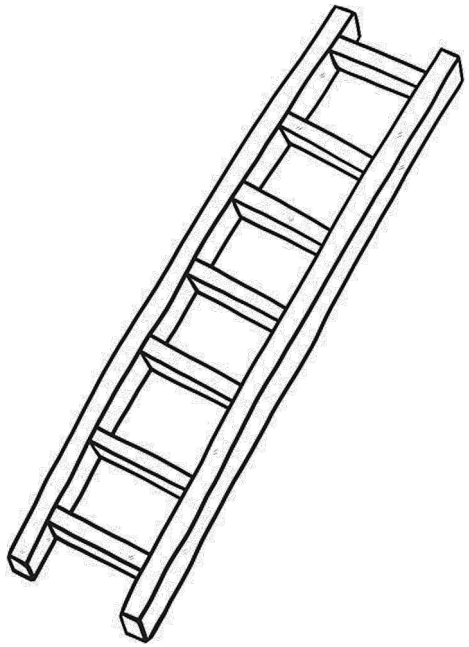
- What are the causal effects?

Patient-Level Prediction

- What will happen to me?

Clinical Characterisation

- What happened to the patients?



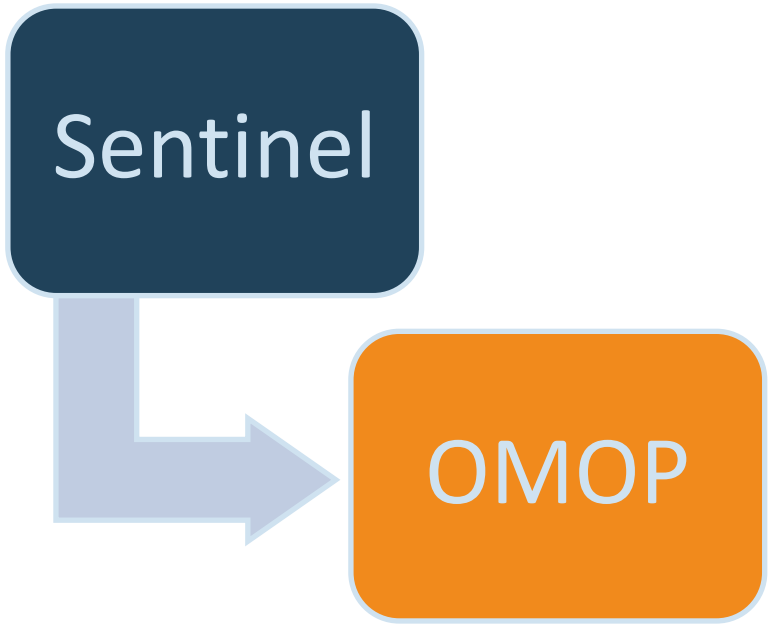
Raw data

Fit-for-purpose data

Standardized analytics

History

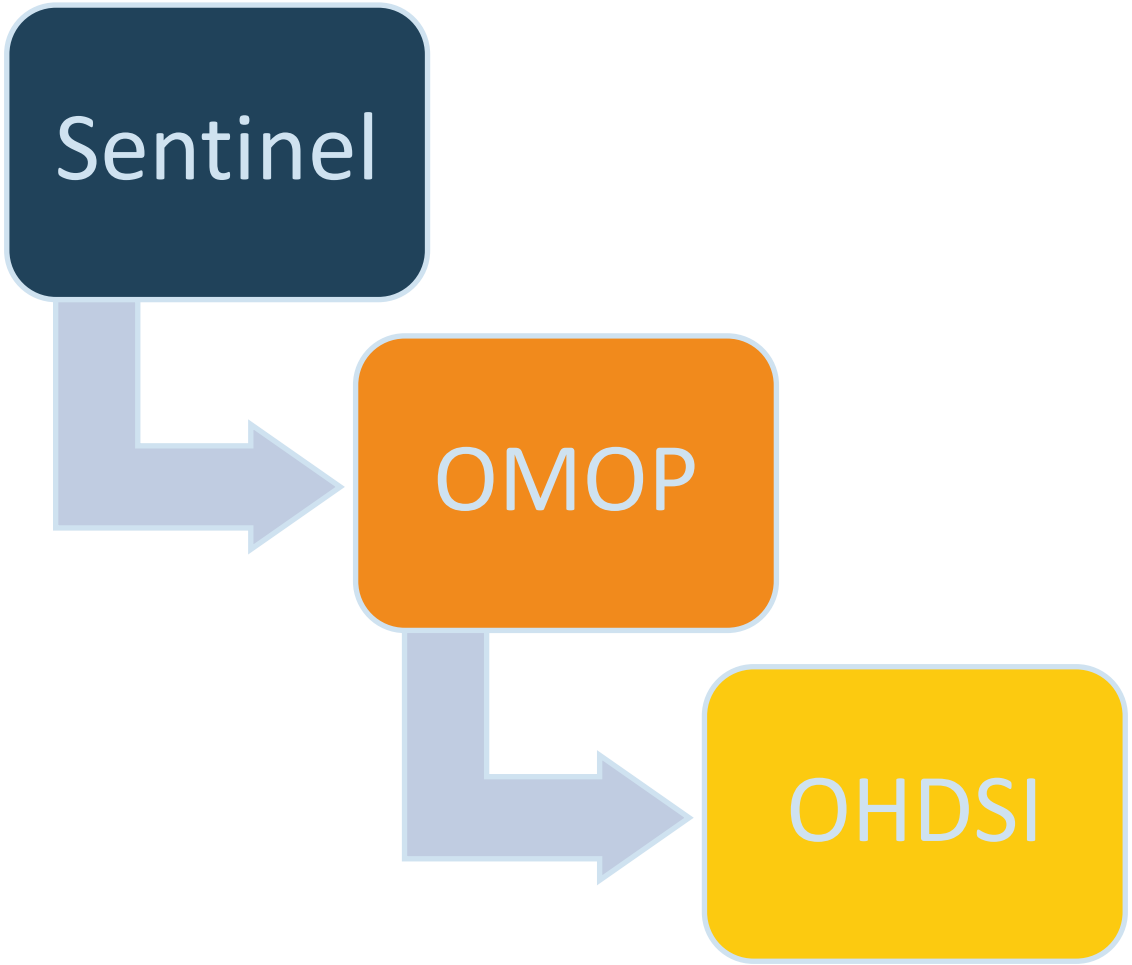
- Since 2000 rapidly increasing number of drug side effects
- FDA Amendments Act □ Risk Identification and Analysis System



OMOP experiment (2009)

- Empirically reproduce a set of known (medical) facts using positive and negative test cases. The assumption was that this could be achieved with the best database and the best methods.
- 10 databases; 14 methods
  - Different database, same method, potentially different outcome
  - Analysis choices are crucial

History



OHDSI (est. 2014)

Vision

A world in which observational research produces a comprehensive understanding of health and disease.

Mission

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

This is OHDSI!

*OHDSI is a **fun** way to collaborate with amazing people across the globe to collectively advance science and improve the lives of patients around the world.*

Philosophy

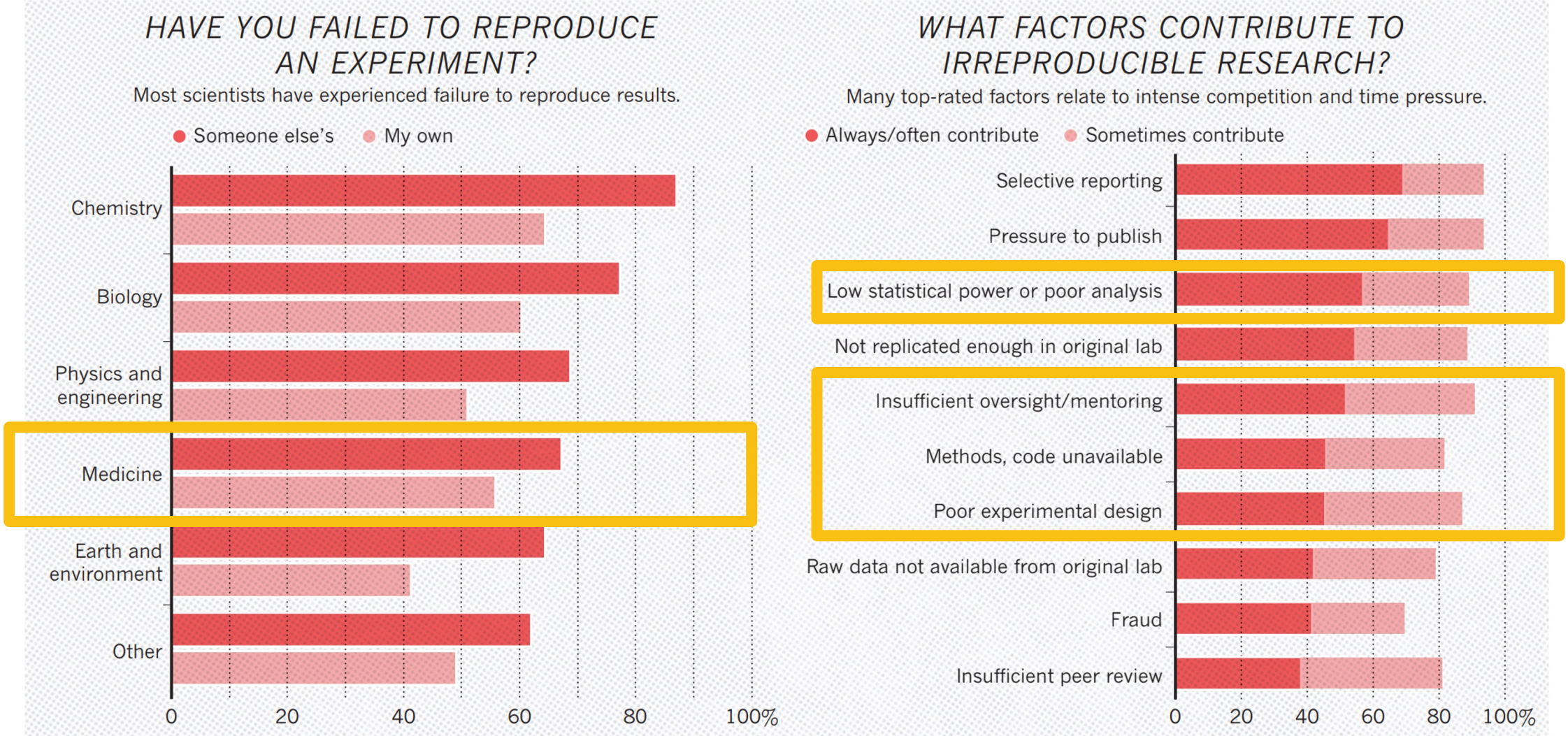
- Interdisciplinary teams
- International collaboration
- All tools should remain public and freely available

More info?
- **The Book of OHDSI**
chapter 1.2
- **EHDEN Academy course**
"OMOP CDM and Standardised Vocabularies"

Values

- **Innovation**: Observational research is a field which will benefit greatly from disruptive thinking. We actively seek and encourage fresh methodological approaches in our work.
- **Reproducibility**: Accurate, reproducible, and well-calibrated evidence is necessary for health improvement.
- **Community**: Everyone is welcome to actively participate in OHDSI, whether you are a patient, a health professional, a researcher, or someone who simply believes in our cause.
- **Collaboration**: We work collectively to prioritize and address the real world needs of our community's participants.
- **Openness**: We strive to make all our community's proceeds open and publicly accessible, including the methods, tools and the evidence that we generate.
- **Beneficence**: We seek to protect the rights of individuals and organizations within our community at all times.

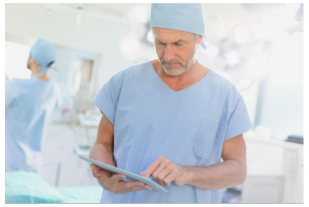
Reproducibility crisis (Nature 2016)



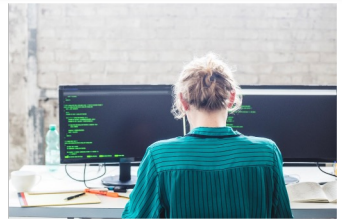
Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016)

Old-school research gone wrong









survey

cohort

PDMS

...

claims

hospital

GP

?

?

?

?

?

?

?



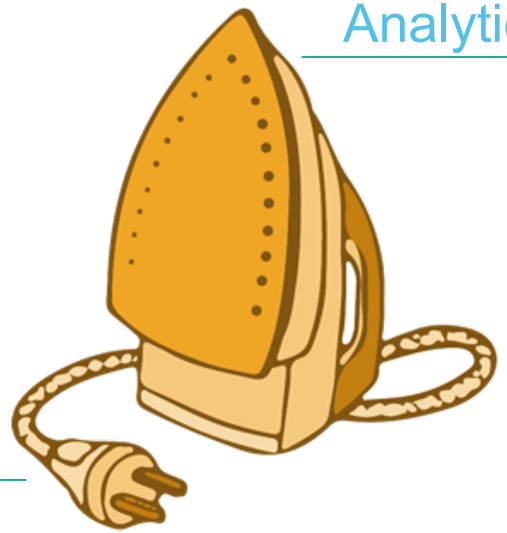




Why Most Published Research Findings Are False

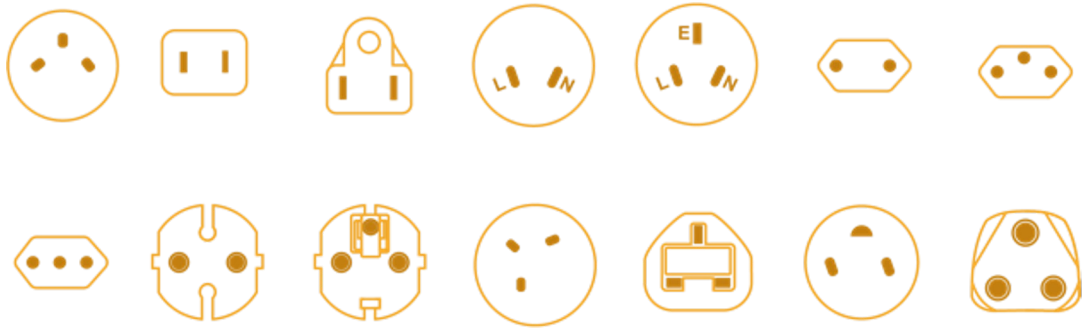
The challenges of Real-World Data

Analytical method



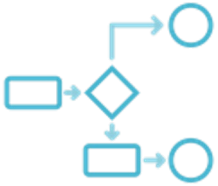
Link to data

The data…



What will it require?

Data interoperability

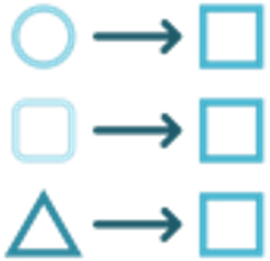


Standardized analytics

Data network

Strong community

Data interoperability



Data interoperability

OMOP CDM

Semantic

Interoperability

Syntactic

Observational research OHDSI style

Strong community

Standardized analytics

survey

cohort

PDMS

...

claims

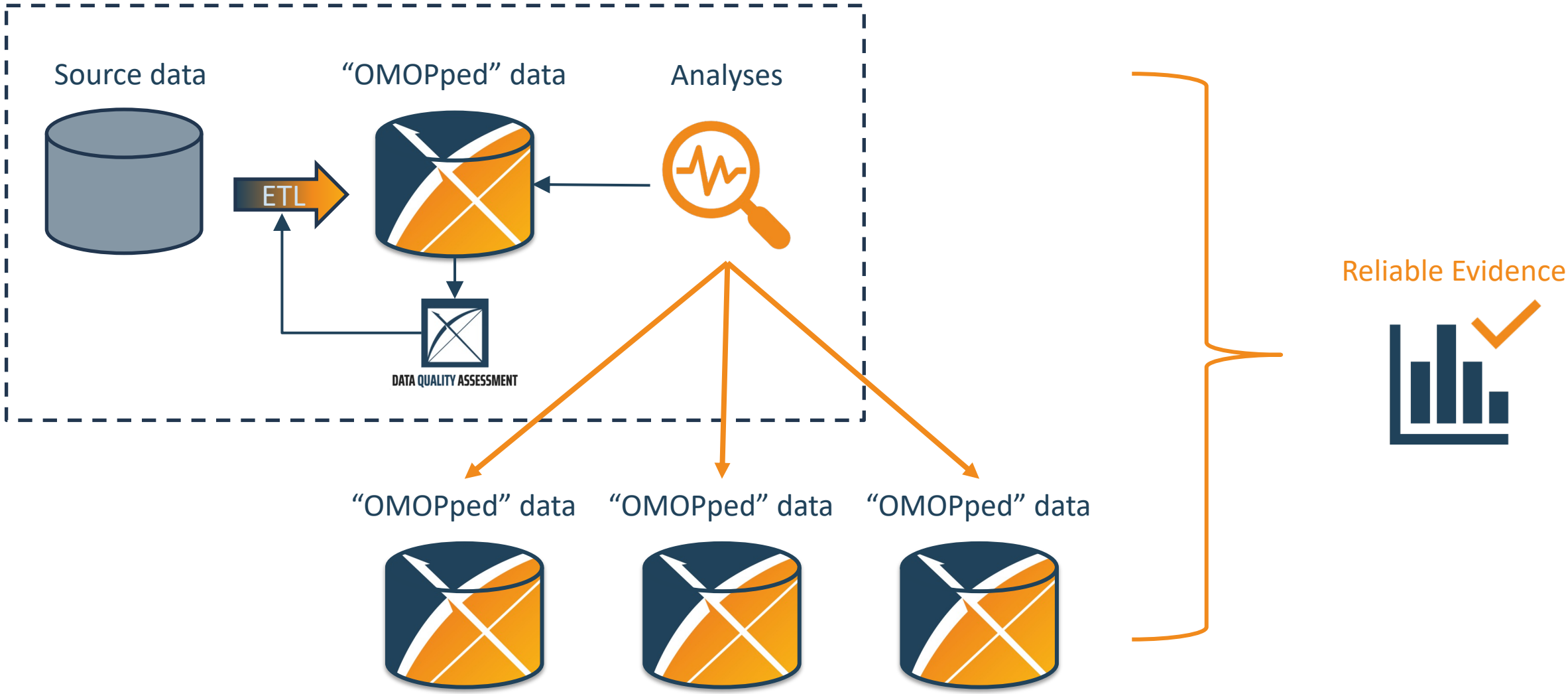
hospital

GP

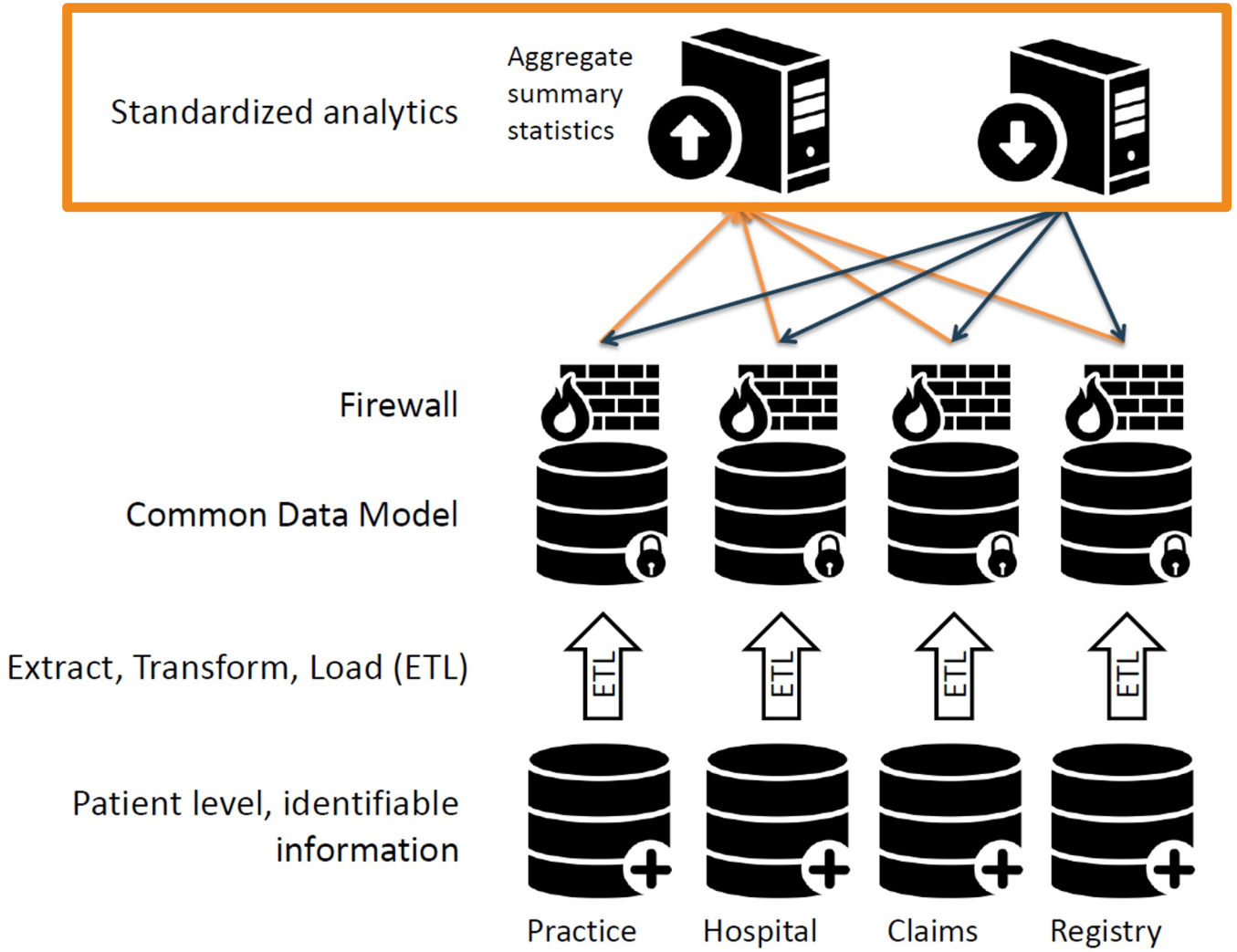
Data aggregation & analysis


Why Most Published Research Findings Are False

Standardized analytics can be reused over data sources

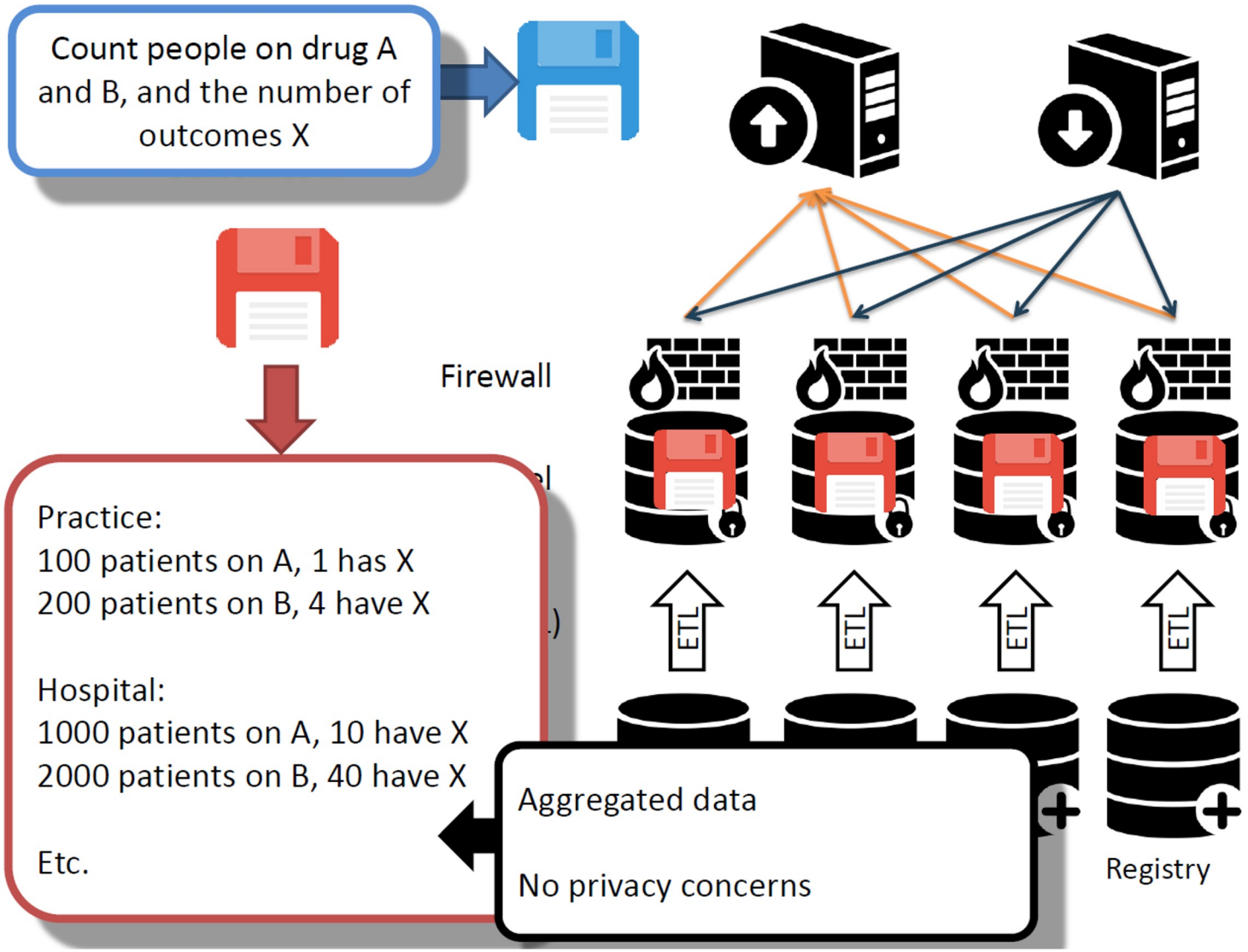


Federated analyses in OHDSI network



Schuemie MJ. How to extract transform and load observational data. https://www.ohdsi.org/wp-content/uploads/2014/07/Beijing2015.pdf?_ga=2.178811554.749634320.1678273784-1300990784.1664885317 Last accessed 09-MAR-23

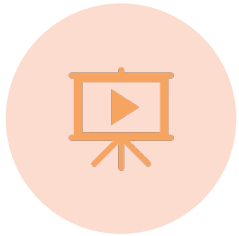
Federated analyses in OHDSI network



Where can I learn more?

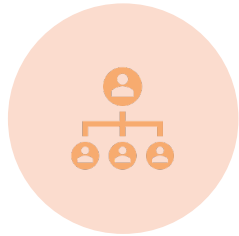
EHDEN ACADEMY

OHDSI WEBSITE / OHDSI TEAMS



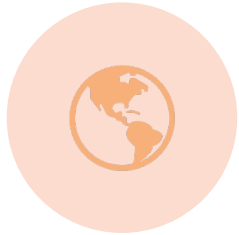
YOUTUBE

FORUMS



GITHUB

COMMUNITY CALL



OHDSI GLOBAL SYMPOSIUM

BOOK OF OHDSI

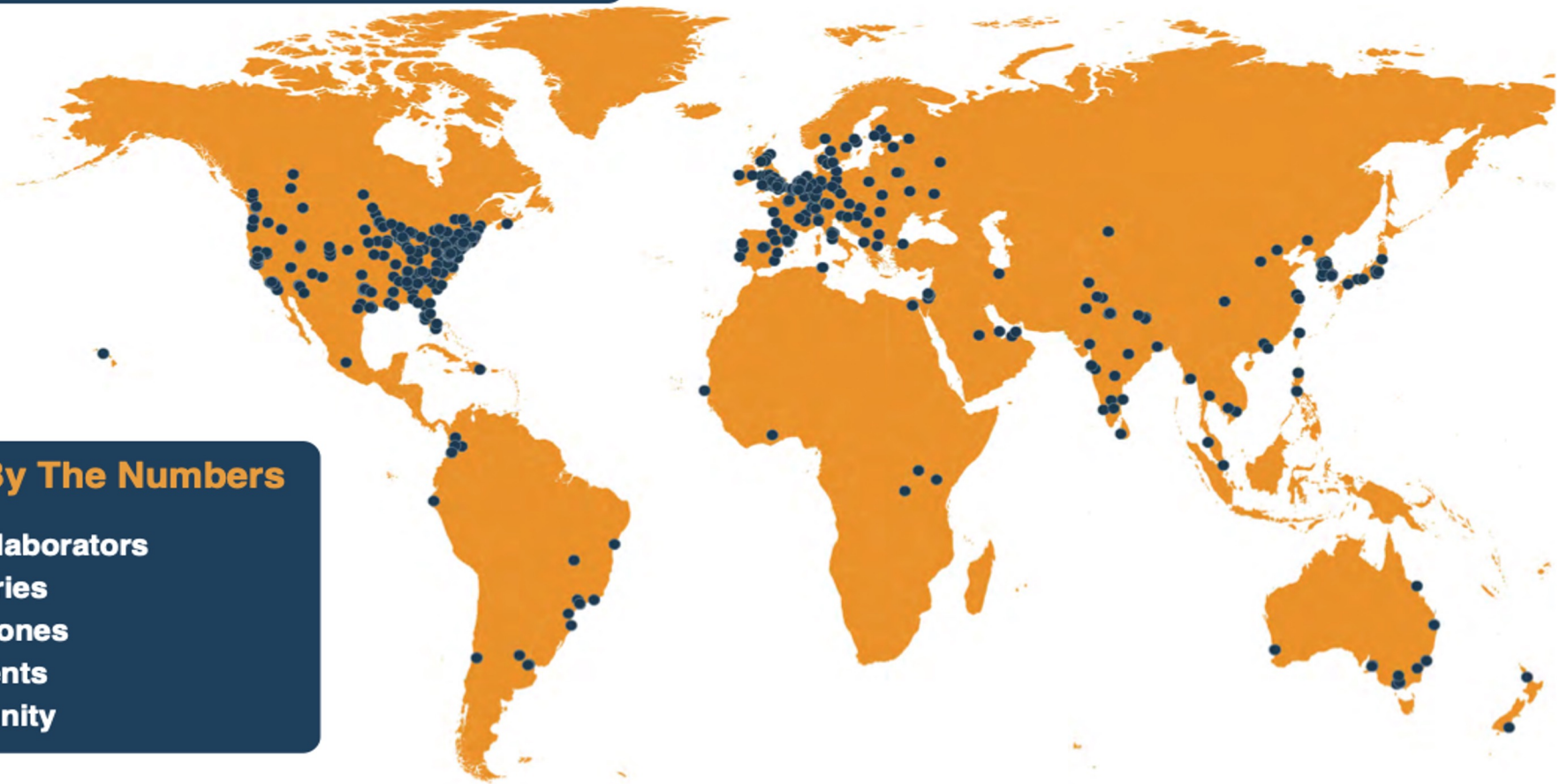
OUR JOURNEY MAGAZINE

Map of Collaborators

The OHDSI community brings together volunteers from around the world to establish open community data standards, develop open-source software, conduct methodological research, and apply scientific best practices to both answer public health questions and generate reliable clinical evidence.

OHDSI COLLABORATORS

Our community is ALWAYS seeking new collaborators. Do you want to focus on data standards or methodological research? Are you passionate about open-source development or clinical applications? Do you have data that you want to be part of global network studies? Do you want to be part of a global community that truly values the benefits of open science? Add a dot to the map below and JOIN THE JOURNEY!



OHDSI By The Numbers

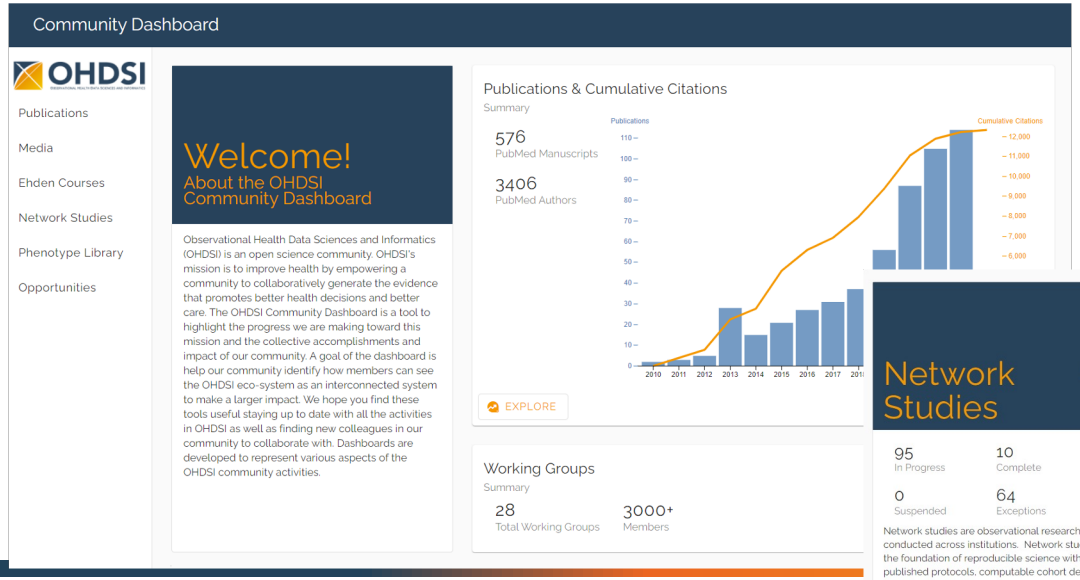
- 2,367 collaborators
- 74 countries
- 21 time zones
- 6 continents
- 1 community

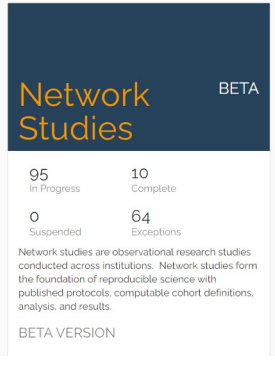
#JoinTheJourney

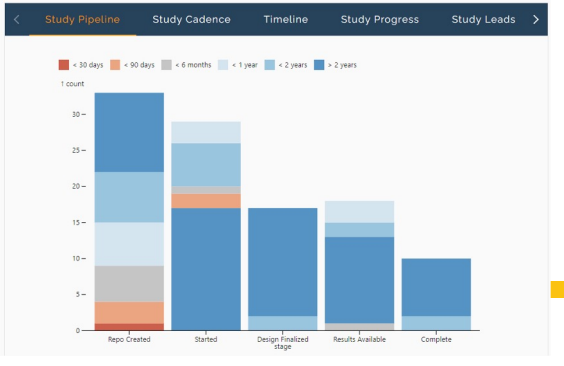
OHDSI Community

www.ohdsi.org

- dash.ohdsi.org







What do we do?

Population-Level Effect Estimation

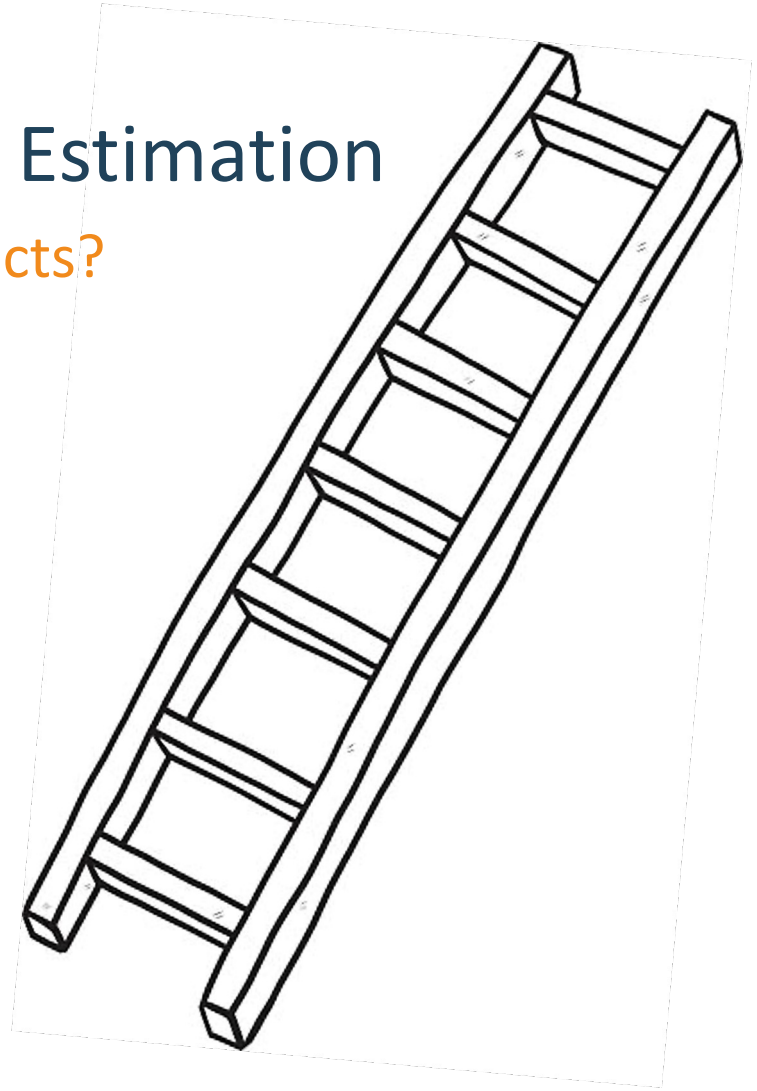
- What are the causal effects?

Patient-Level Prediction

- What will happen to me?

Clinical Characterisation

- What happened to the patients?

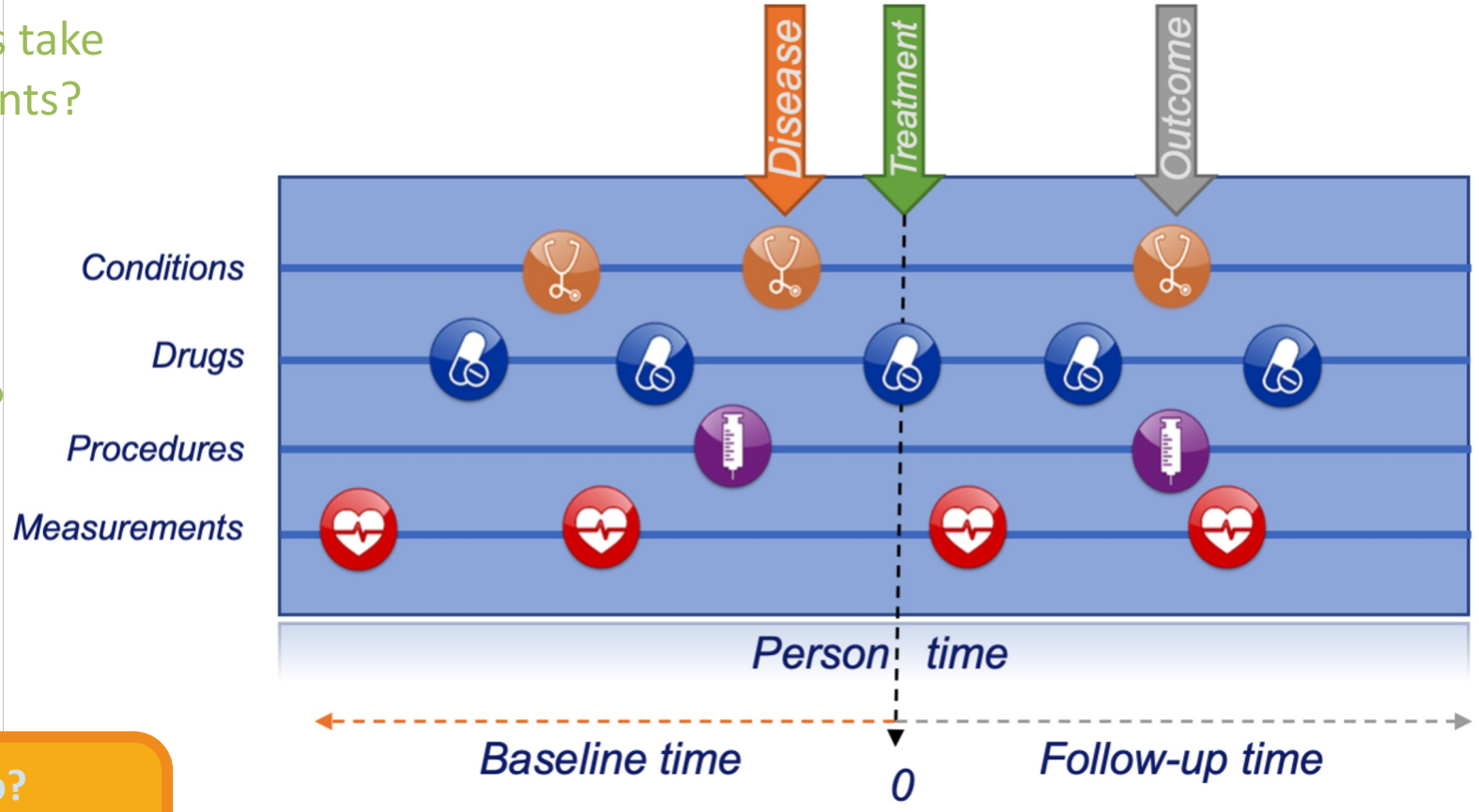


Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE 2019;32(1): 42-49, DOI: 10.1080/09332480.2019.1579578.

Questions asked across the patient journey

Which patients take which treatments?

How many patients experienced complications?



What is the probability that I will develop a given disease?

What is the probability that I will experience an (adverse) outcome?

More info?
- **The Book of OHDSI** chapter 7

Does treatment cause an outcome?

Does one treatment lead to a better outcome than the alternative?

Treatment Patterns Across Disease Domains

- Objective: provide high-level insight in real-world treatment practices across different care settings, countries and continents

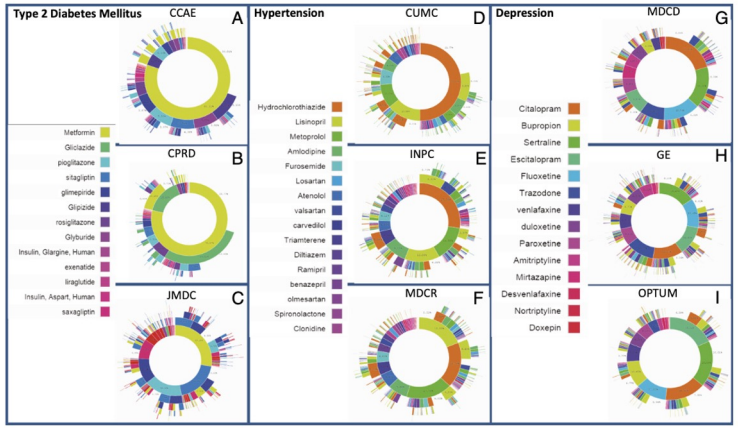


Fig. 3. For each disease, diabetes (*A–C*), hypertension (*D–F*), and depression (*G–I*), the inner circle shows the first relevant medication that the patient took, the second circle shows the second medication, and so forth. Three data sources are shown for each disease; the data source abbreviations are defined in Table 2.

Hripcsak et al. (2016)

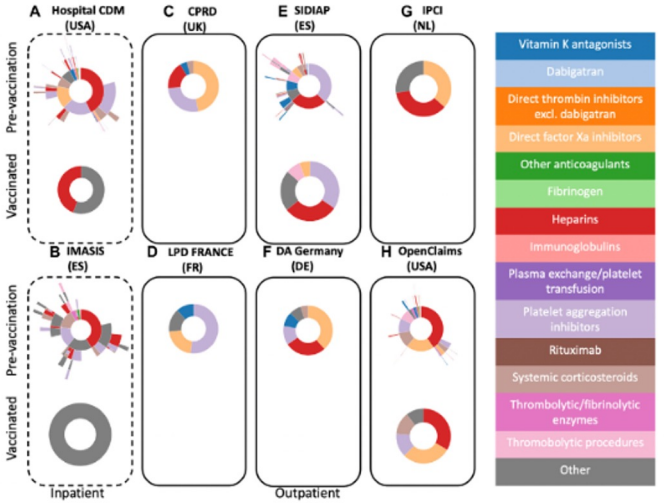
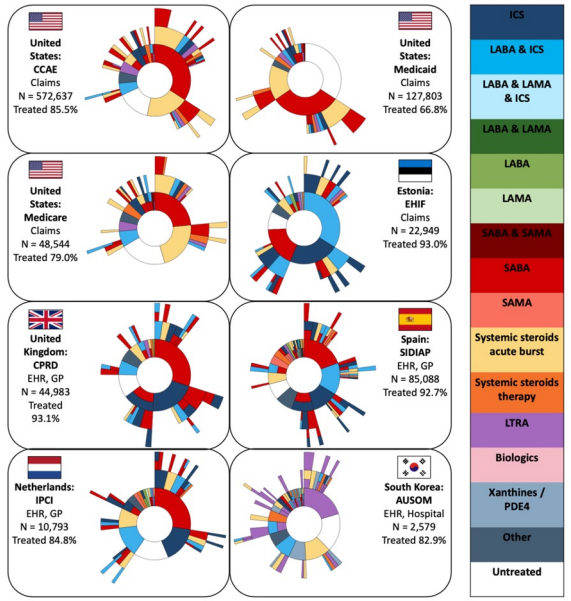


FIGURE 1
Sunburst plots visualizing treatment pathways for TTS patients in pre-vaccination cohorts (top) *versus* vaccinated cohorts (bottom). Inpatient databases are depicted with a dashed line frame, whilst outpatient ones have a solid frame.

Markus et al. (2023)



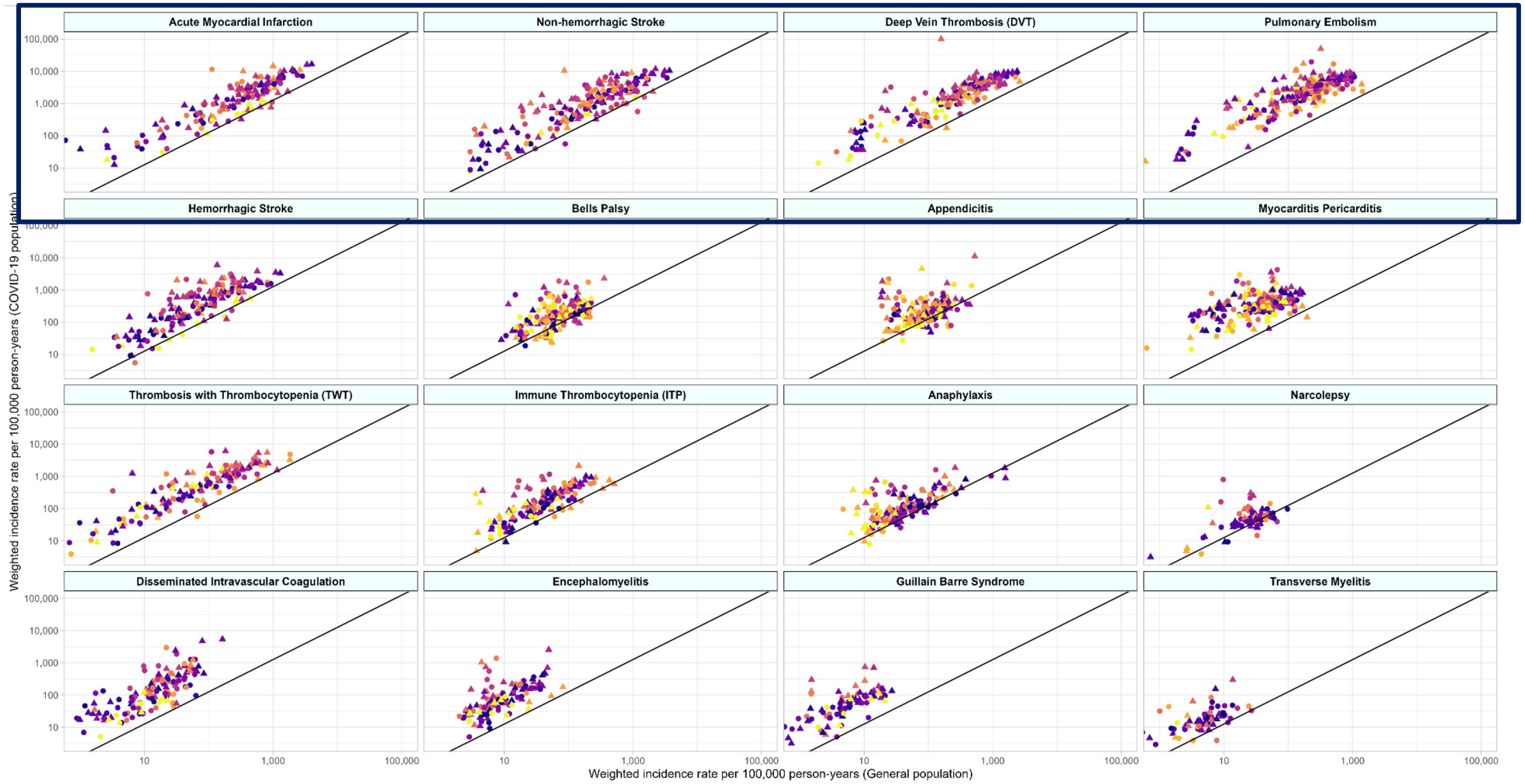
Markus et al. (2024)

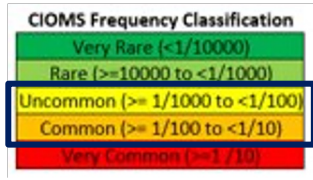
- Large-scale characterization helps to generate hypotheses for follow-up studies to address current gaps in clinical practice

https://darwin-eu.github.io/TreatmentPatterns/

More info?
- The Book of OHDSI chapter 11.3 + 11.9
- EHDEN academy course in progress

Comparing the ‘Subjects with COVID-19' Stratified Incidence Rates for AESIs to Those for ‘Pre-Pandemic Background Population’





Voss et al. (2023)

26 databases
11 countries

OHDSI LEGEND studies

Journal of the American Medical Informatics Association, 0(0), 2020, 1268–1277
doi: 10.1093/jamia/ocaa124
Research and Applications

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]

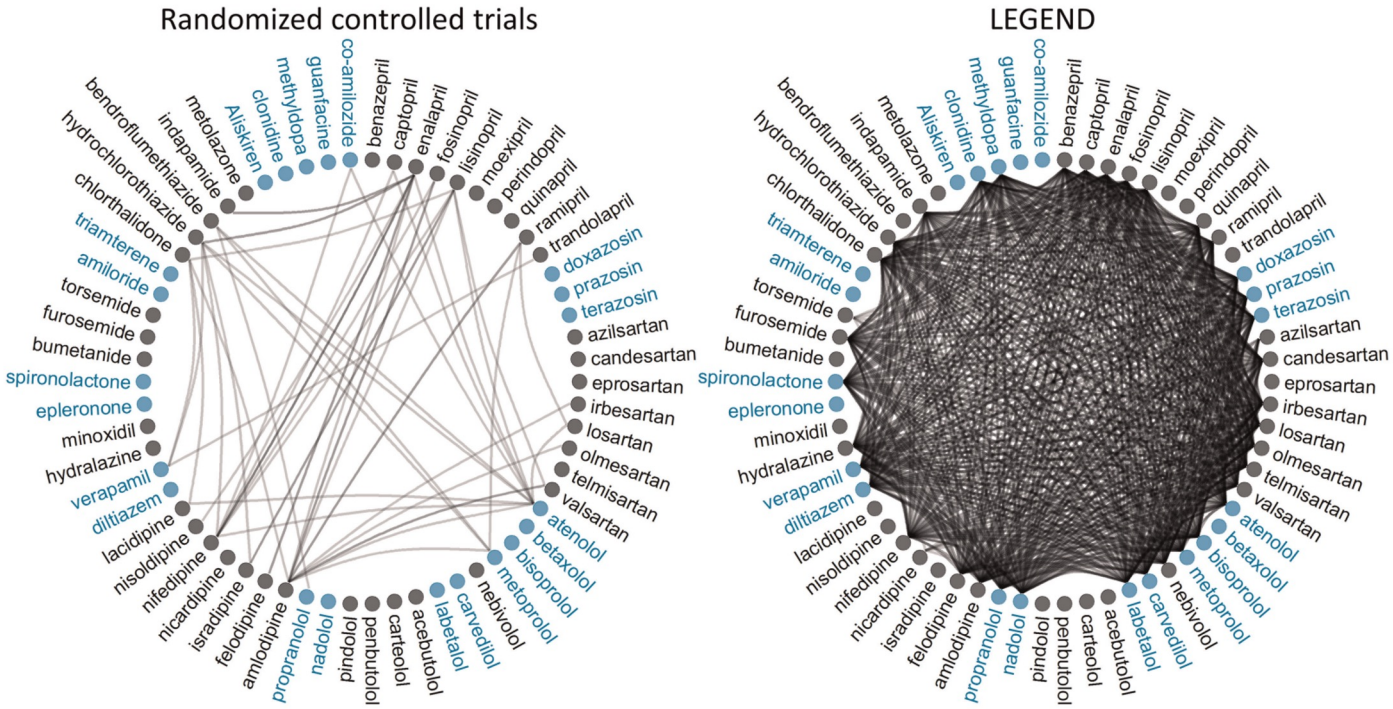


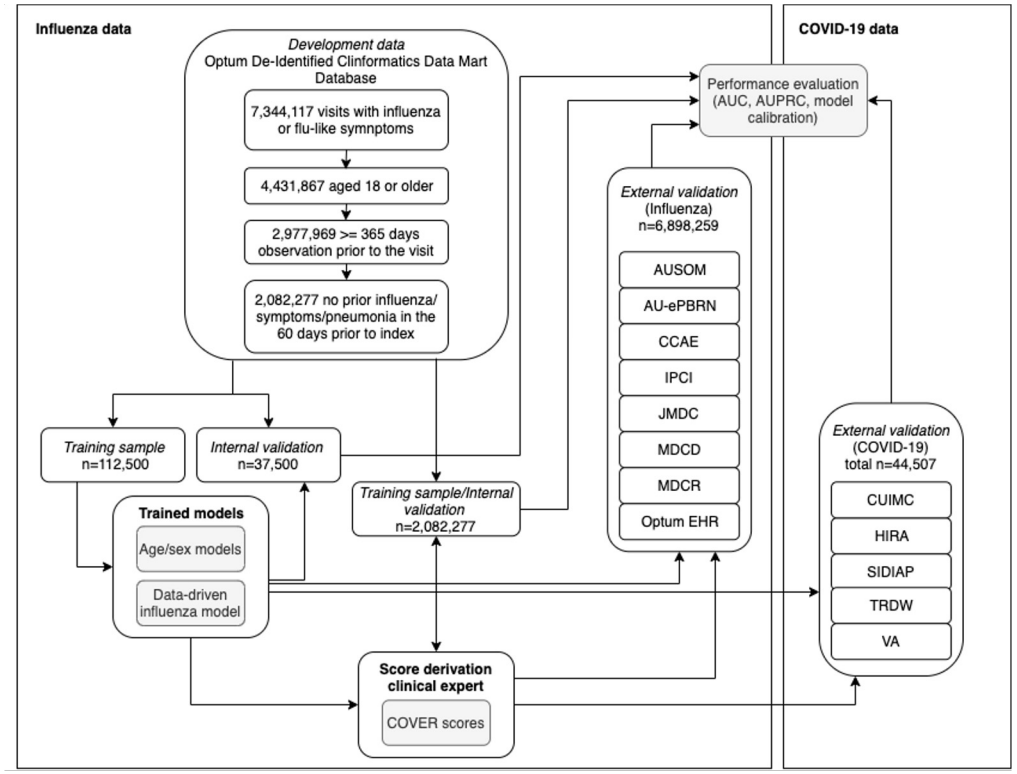
Figure 3. Comparisons of single-drug hypertension treatments in randomized controlled trials (left) and in LEGEND (right). Each circle represents an ingredient. Color groupings indicate drug classes. A line between circles indicates the 2 drugs are compared in at least 1 study.

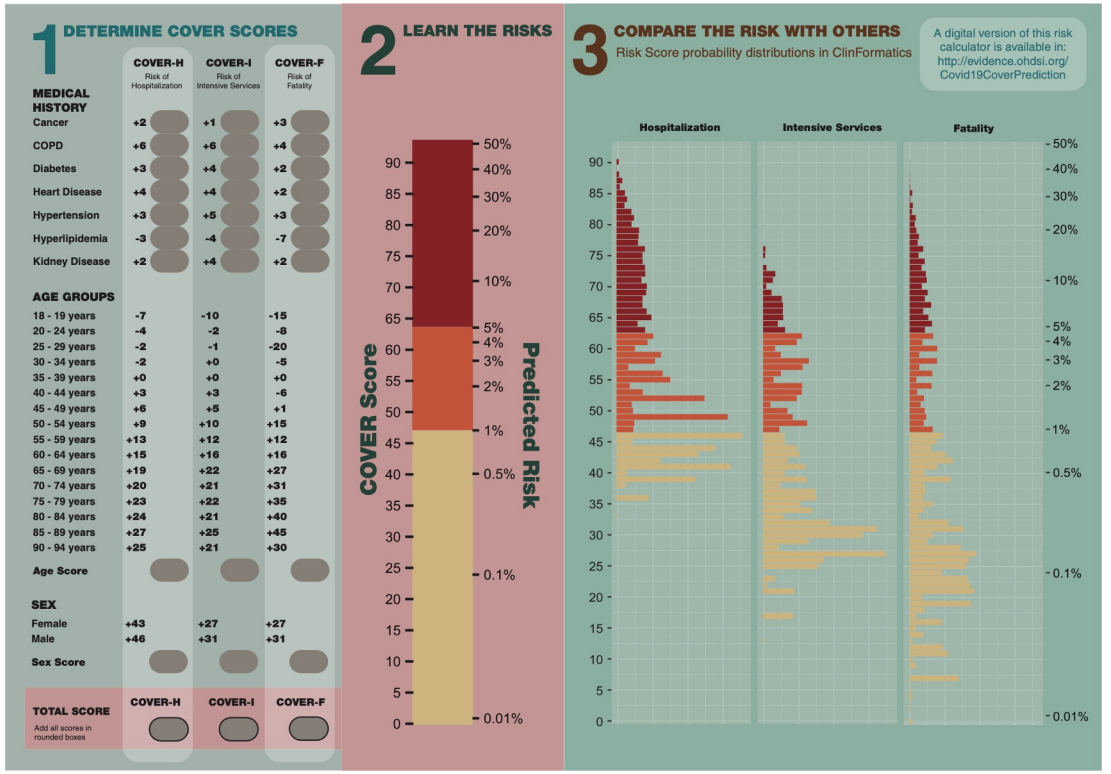
More info?
- **OHDSI website**
https://www.ohdsi.org/legend-oct2021-update/
- **Various publications**
e.g. 10.1093/jamia/ocaa103

Results: From 21.6 million unique antihypertensive new users, we generate 6 076 775 effect size estimates for 699 872 research questions on 12 946 treatment comparisons. Through propensity score matching, we achieve balance on all baseline patient characteristics for 75% of estimates, observe 95.7% coverage in our effect-estimate 95% confidence intervals, find high between-database consistency, and achieve transitivity in 84.8% of triplet hypotheses. Compared with meta-analyses of RCTs, our results are consistent with 28 of 30 comparisons while providing narrower confidence intervals.

SEEK COVER

- Objective: develop and externally validate **COV**ID-19 **E**stimated **R**isk scores that quantify a patient's risk of hospital admission, hospitalization requiring intensive services or fatality.



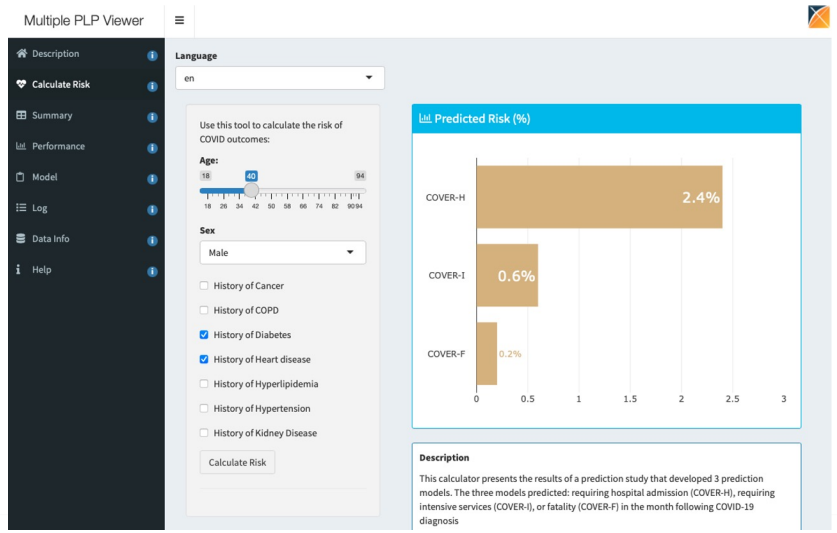


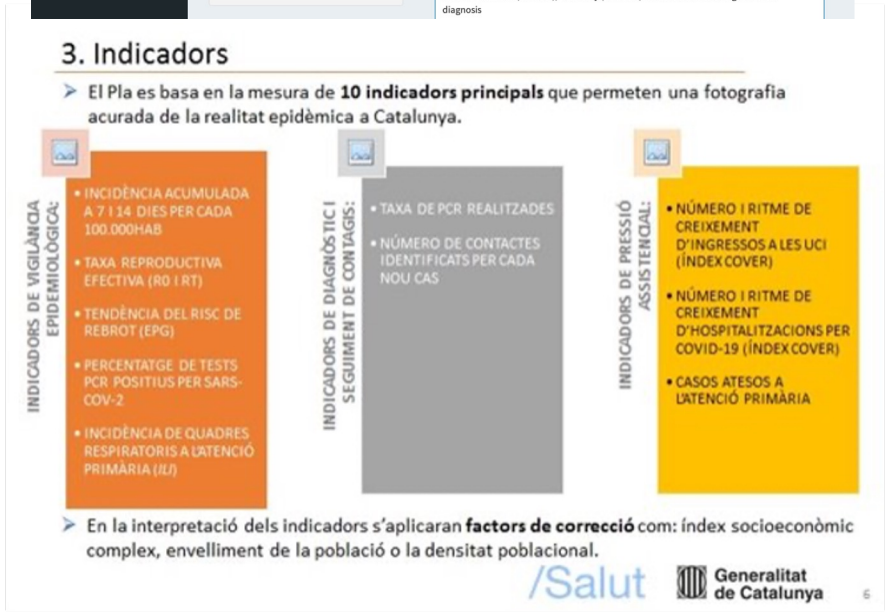
Williams and Markus et al. (2022)

SEEK COVER

- COVER interactive website to provide live risk scores.

- Impact: Health minister of Catalonia Spain explicitly mentioned the COVER index as one of the indicators they used to measure the impact of a given outbreak.





Only 5 years ago…

Open access Research

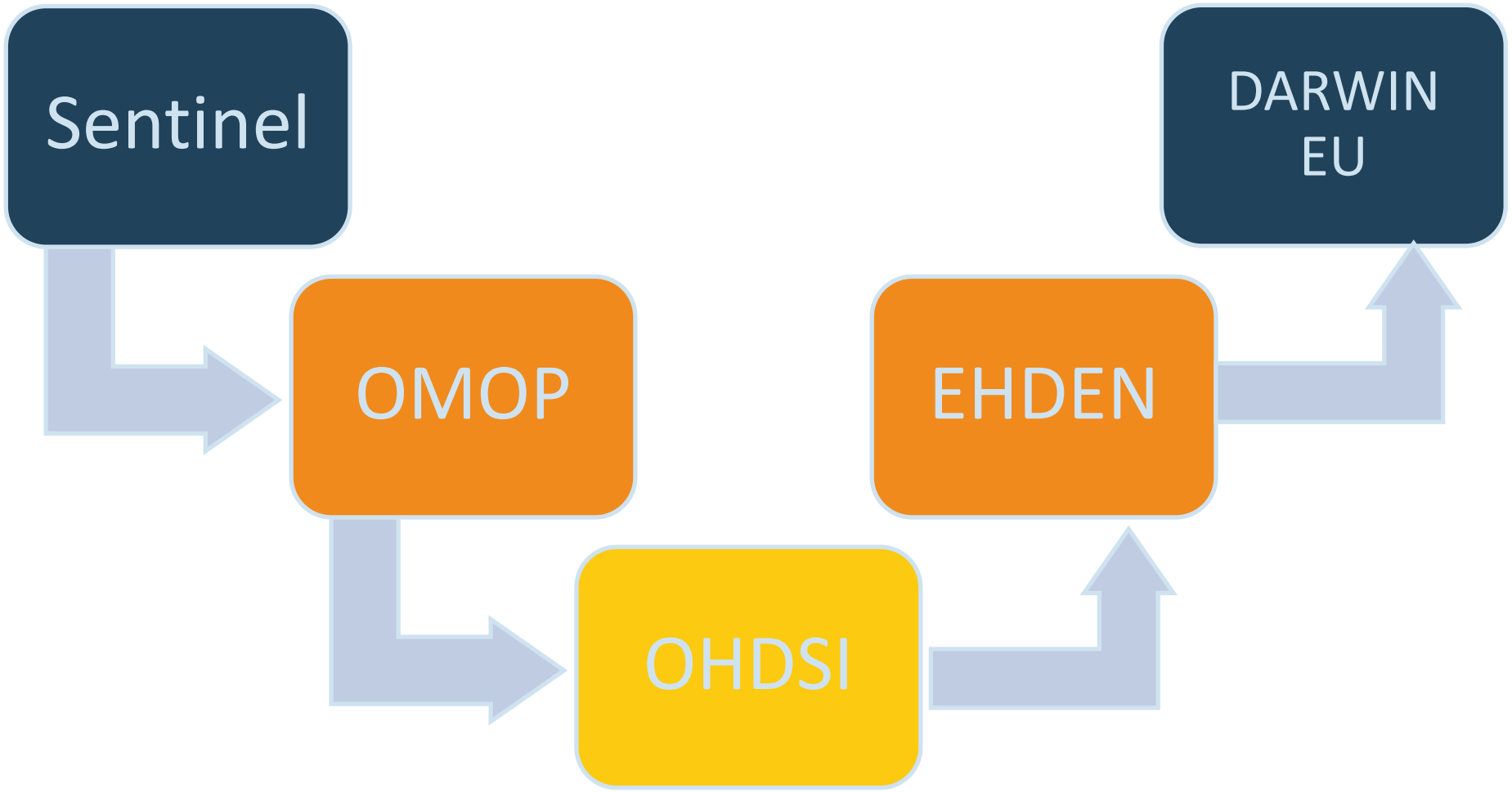
BMJ Open

Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation

Alexandra Pacurariu,[1] Kelly Plueschke,[1] Patricia McGettigan,[1,2] Daniel R Morales,[1,3] Jim Slattery,[1] Dagmar Vogl,[1] Thomas Goedecke,[1] Xavier Kurz,[1] Alison Cave[1]

Results A total of 34 EHDs were selected after applying key criteria relevant for regulatory purposes. The most represented regions were Northern, Central and Western Europe. The most frequent types of data source were electronic medical records (44.1%) and record linkage systems (29.4%). The median number of patients registered in the 34 data sources was 5 million (range 0.07–15 million) while the median time covered by a database was 18.5 years. Paediatric patients were included in 32 databases (94%). Completeness of information on drug exposure was variable. Published validation studies were found for only 17 databases (50%). Some level of access exists for 25 databases (73.5%), and 23 databases (67.6%) can be linked through a personal identification number to other databases with parent–child linkage possible in 7 (21%) databases. Eight databases (23.5%) were already transformed or were in the process of being transformed into a common data model that could facilitate multidatabase studies.

Conclusion A Few European databases meet minimal regulatory requirements and are readily available to be used in a regulatory context. Accessibility and validity information of the included information needs to be improved. This study confirmed the fragmentation, heterogeneity and lack of transparency existing in many European EHDs.


Sentinel
OMOP
OHDSI
EHDEN
DARWIN EU

DARWIN EU

Key Figures

The European Medicines Agency (EMA) and the European Medicines Regulatory Network established a coordination centre to provide timely and reliable evidence on the use, safety and effectiveness of medicines for human use, including vaccines, from real world healthcare databases across the European Union (EU). This capability is called the **Data Analysis and Real World Interrogation Network (DARWIN EU®)**. It is a flagship of the European Health Data Space

National nodes

- Belgium
- Denmark
- Estonia
- Finland
- Germany
- Greece
- Hungary
- Israel
- Italy
- Luxembourg
- Netherlands
- Norway
- Portugal
- Spain
- United Kingdom

https://www.ohdsi-europe.org/index.php/national-nodes

Meet our OHDSI BE Core Group Members

OHDSI BELGIUM

Liesbet M. Peeters
(UHasselt)

Annelies Verbiest
(UZ Antwerpen)

Peter De Jaeger
(RADar / AZ Delta)

Nicky Van Der Vekens
(AZ Maria Middelares)

Giovanni Briganti
(CHU Mons)

Freija Descamps
(edence Health)

Dries Hens
(Lynxcare)

Bart Vannieuwenhuyse
(pharma.be / Janssen)

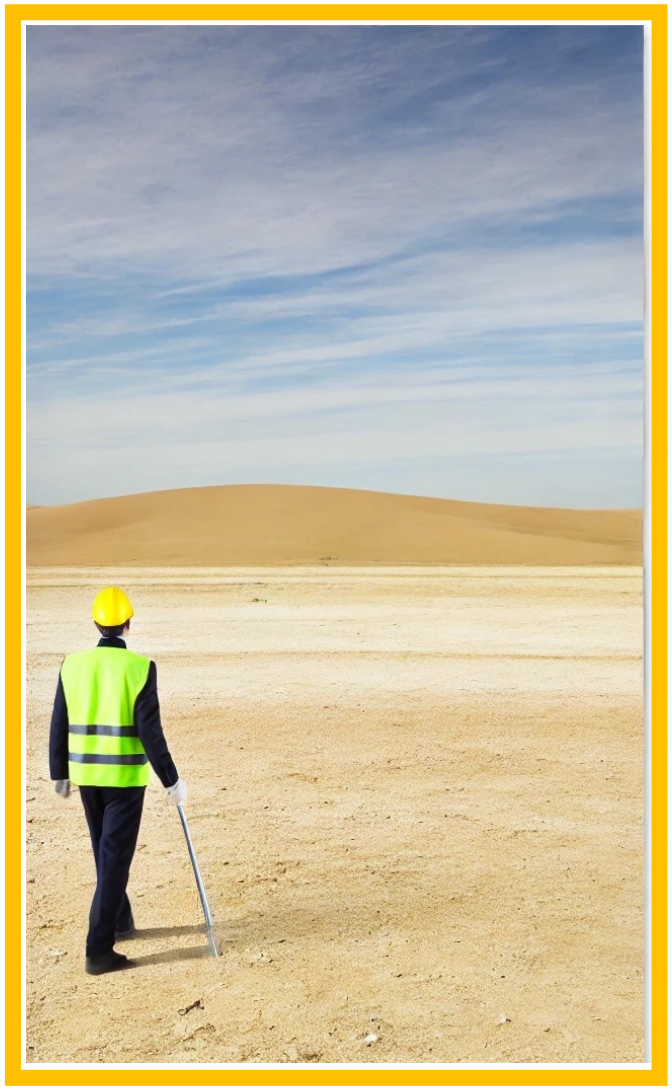
Highlights - OHDSI BE growing community

OHDSI BELGIUM

Since official launch of OHDSI Belgium in June 2023

- **170** interested individuals on mailing list
  Not yet registered? - reach out to ilse.vermeulen@uhasselt.be
- **70** institutions in Belgium
- **15** data partners (EHDEN-network)

Follow our OHDSI Belgium LinkedIn page for news and updates



Raw data

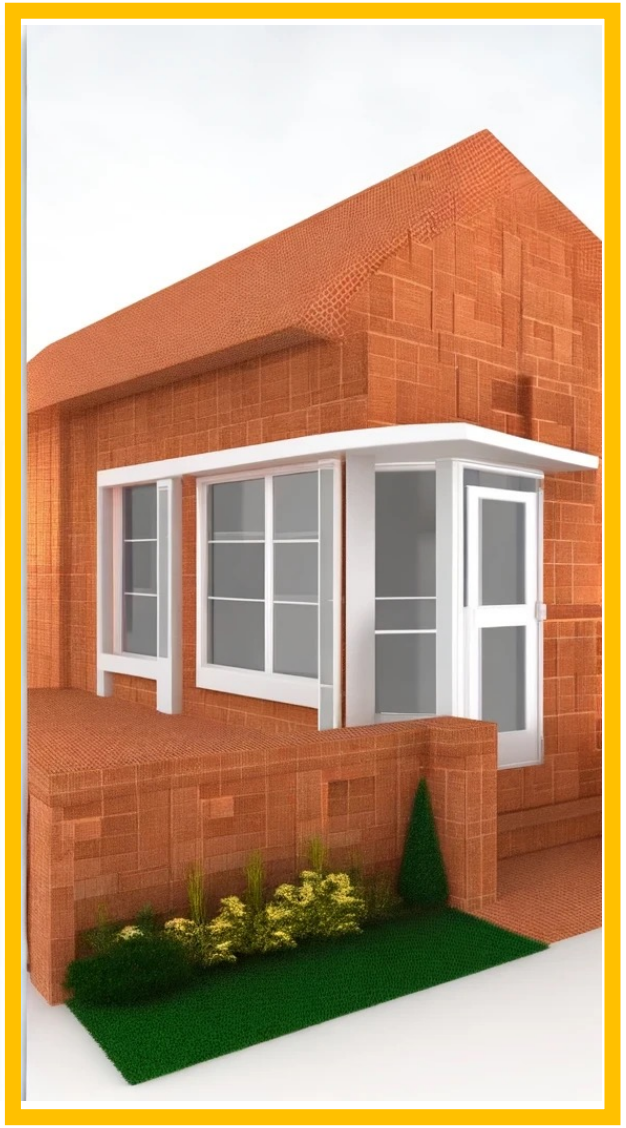
Fit-for-purpose data

Standardized analytics

Belgium has a higly digitized healthcare system

Raw data

Fit-for-purpose data



Standardized analytics

Can Belgian partners run a federated study?

Birth of the OHDSI FALCON network

Federated Alliance of Large-scale Cancer Observational Network

- Poised to address the unique challenges of cancer RWE
  - Representing cancer in a datamodel
  - Small populations in personalized medicine
- On it’s way to become the largest and most diverse federated oncology RWE network
- FALCON-Lung
  - 9 months from conception to delivery

FALCON – Lung

Most diverse & soon largest lung cancer RWE network

- 50+ partners ran data quality assessment
- 22 implemented data quality improvement (“patch”)
- 20 onboarded study (others in progress): 11 countries, hospital-pharma-public/private partners
- 17 currently with results
    - **4 of 17 are Belgian centers** (UZ Antwerp, CHU Liège, GHDC, CHC).
    - Other Belgian centers taking first steps

Kimmo Porkka – Helsinki University Hospital
Asieh Golozar – Nemesis Health & OHDSI oncology

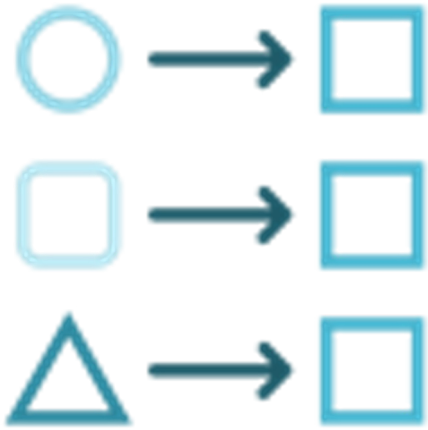
Raw data

Fit-for-purpose data

Standardized analytics

Can we deliver research-grade standardized data?

Value proposition for hospitals



Only once!

Research

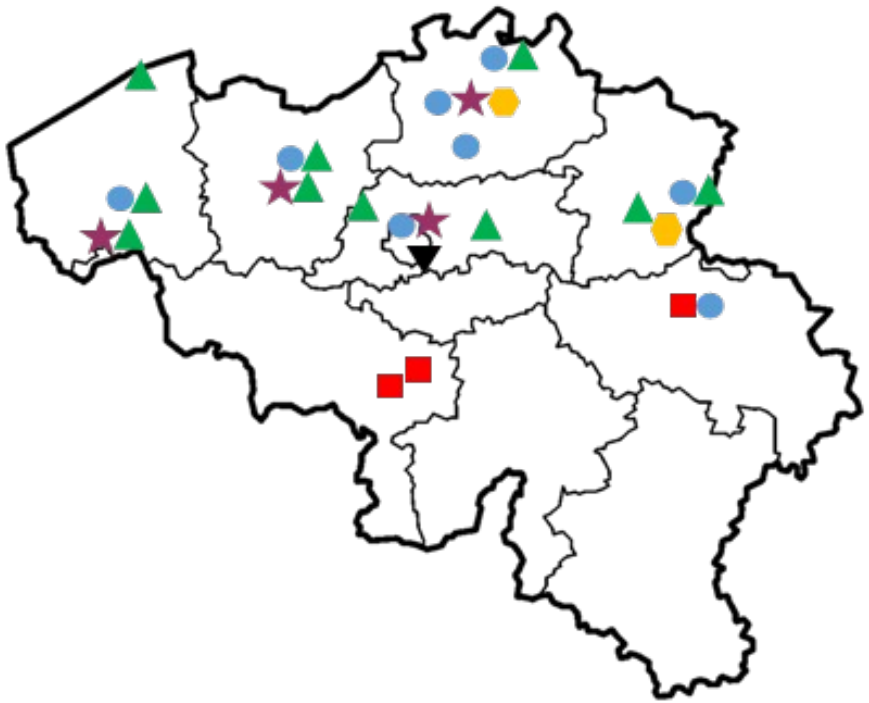
BI

Exchange

Quality improvement

Yet our hospitals still really need good incentives!

Figure 1: Belgian hospitals are organizing themselves to generate research-grade data from routinely-collected healthrecord data, both autonomously and in collaboration with medtech companies. The networks below all use OMOP-CDM as data-warehouse for clinical analytics. Year of inauguration and membership status as of January 2025 are shown.



- **FHIN ngo**: AZ Delta, UZA, UZ Gent, UZ Brussel, CHU Liège, AZ Klina, AZ Imelda, ZOL. 2024. www.fhin.be
- **Levilo Health ngo**: UZ Leuven, ZOL, OLV Aalst, AZ Delta, UZ Gent, AZ Klina, AZ Oostende, AZ Groeninge, AZ Maria Middelares, AZ Turnhout, AZ Jessa. 2024.
- **Lynxcare RWE hub Onco:** UZA, AZ Maria Middelares, AZ Groeninge, UZ Brussel. 2022. www.lynx.care
- **INAH:** CHC, CHU Charleroi, CHDC. 2018. www.inah.be
- **DIGIONE IQVIA:** UCL St Luc (+ international partners). 2022.
- **OHDSI Belgium national node:** >70 institutions, coordinated by UZA, UHasselt. 2023.

OHDSI Europe Symposium 2025