Federated Frontiers:

Unleashing Europe's Real-world Health Data Through Common Models and Studyathons

Robert Snijder
26 NOV 2024
PSDM

- **Disclaimer**
- The views and opinions expressed in this presentation are solely my own and do not necessarily reflect the official policies, positions, or opinions of **Astellas**, **OHDSI Pioneer**, or **EHDEN**.
- This presentation is intended for informational purposes only and should not be interpreted as representing these organizations' strategies, practices, or endorsements.

Topics

European Data Landscape

- Understanding the intricate network of data sources and regulations.

Common Data Model

- Harmonizing data for interoperability and insights.

Federated Data

- Leveraging decentralized data while preserving privacy.

EHDEN

- The European Health Data & Evidence Network and its role in shaping research.

PIONEER

- The forefront of prostate cancer research and data analysis.

Studyathons

- Collaborative, intensive study events driving forward medical research.

OPERATIONAL DEFINITION OF REAL WORLD DATA

- **Real World Data (RWD)**: Health-related data generated and collected outside of controlled, experimental trials. Sources of RWD may include but not limited to: electronic health records (EHR), administrative medical claims records, lab results, images, omics, social media, surveys, cohorts and registries.

Barriers to utilization of RWD in Europe

1

Stringent Privacy Laws:

- GDPR Compliance: Strict guidelines on data protection and privacy.
- Varied Interpretation: Different EU countries may interpret and apply these regulations differently.

2

Diverse Healthcare Systems:

- Varying Structures: Differences in healthcare delivery, funding, and administration across EU nations.
- Inconsistent Data Collection: Diversity in data collection methods and standards, leading to challenges in data harmonization.

3

National Data Protectionism:

- Restricted Access: Some countries, like France, limit data access to national entities, posing barriers for international research and collaboration.
- Legal and Bureaucratic Hurdles: Navigating through each country's unique legal framework to access data.

4

Implications for Research and Development:

- Delays in Drug Development: Challenges in accessing and utilizing RWD can lead to delays in drug research and development.
- Limited Cross-Border Collaboration: Restrictions on data sharing impede collaborative efforts across EU countries.

Current European Health Data Space – OMOP & Federated Data Networks

Overcoming Barriers

Several initiatives are critical in mitigating these challenges, including:

- EHDEN
- IMI Pioneer
- OHDSI
- DARWIN

Innovative Projects and Methods

Projects like BD4BO under IMI are pioneering the exploration of EU data for patient benefit. Tools aiding in this endeavor include:

- Federated data solutions
- Common Data Models like OMOP
- Novel analysis techniques such as federated learning and distributed regression

astellas

The standardization of RWD

Common Data Model (CDM) in Real World Data

Examples of RWD CDM

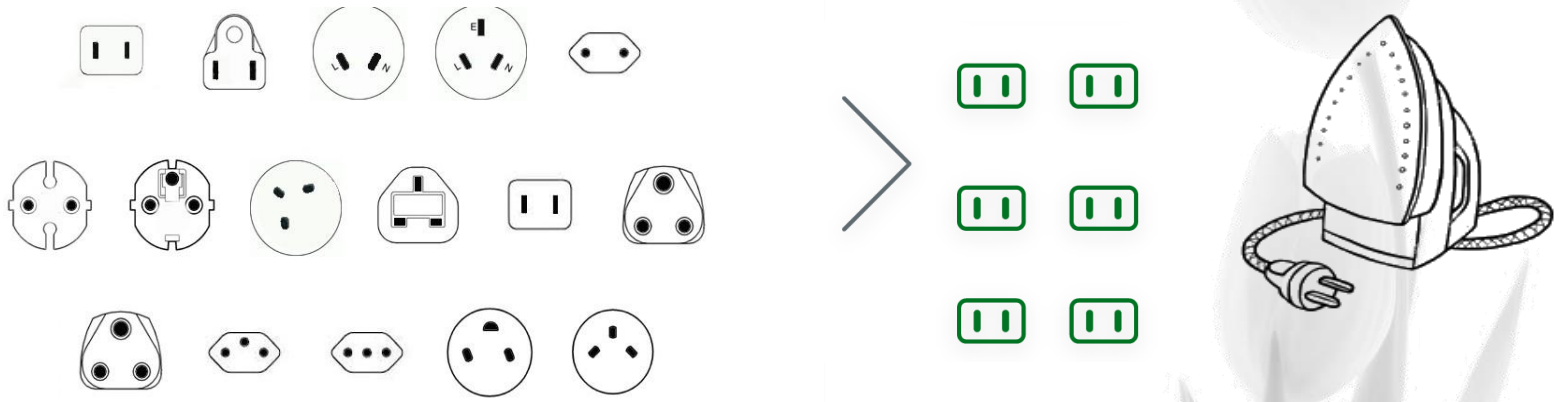
Sentinel

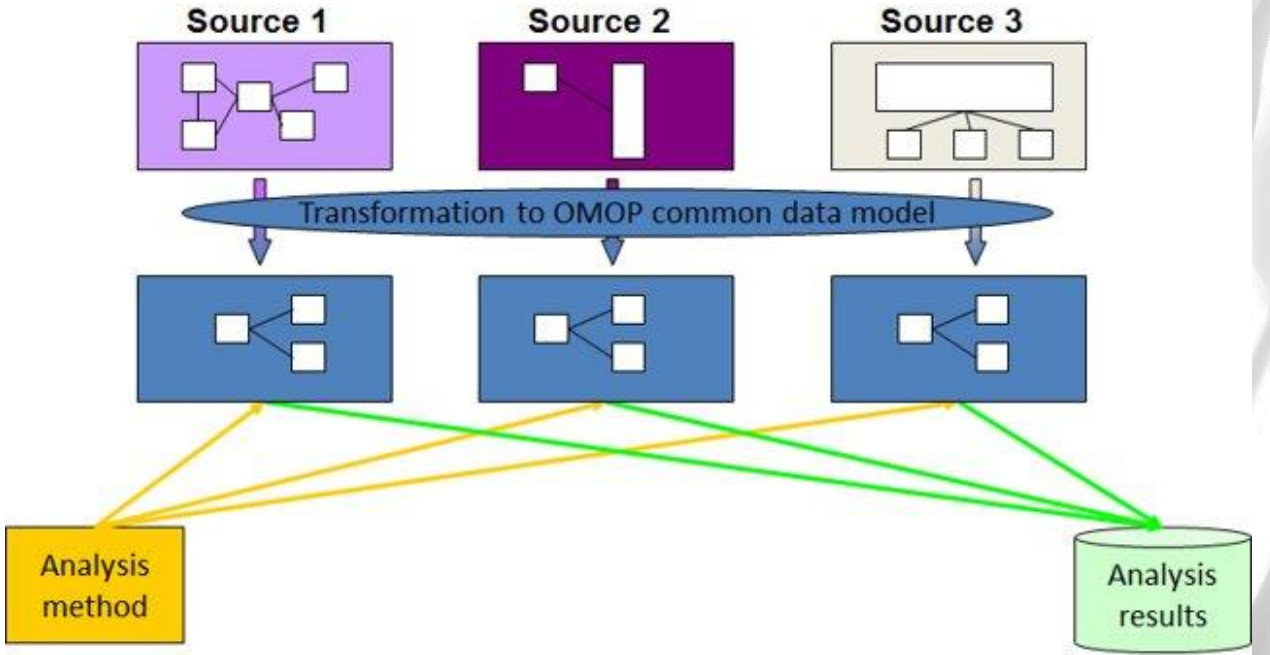
The Sentinel Operations Center (SOC) coordinates the network of Sentinel Data Partners and leads development of the Sentinel Common Data Model (SCDM), a standard data structure that allows Data Partners to quickly execute distributed programs against local data

The **Observational Medical Outcomes Partnership (OMOP) Common Data Model** (CDM) is an open community data standard, designed to standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence.

astellas

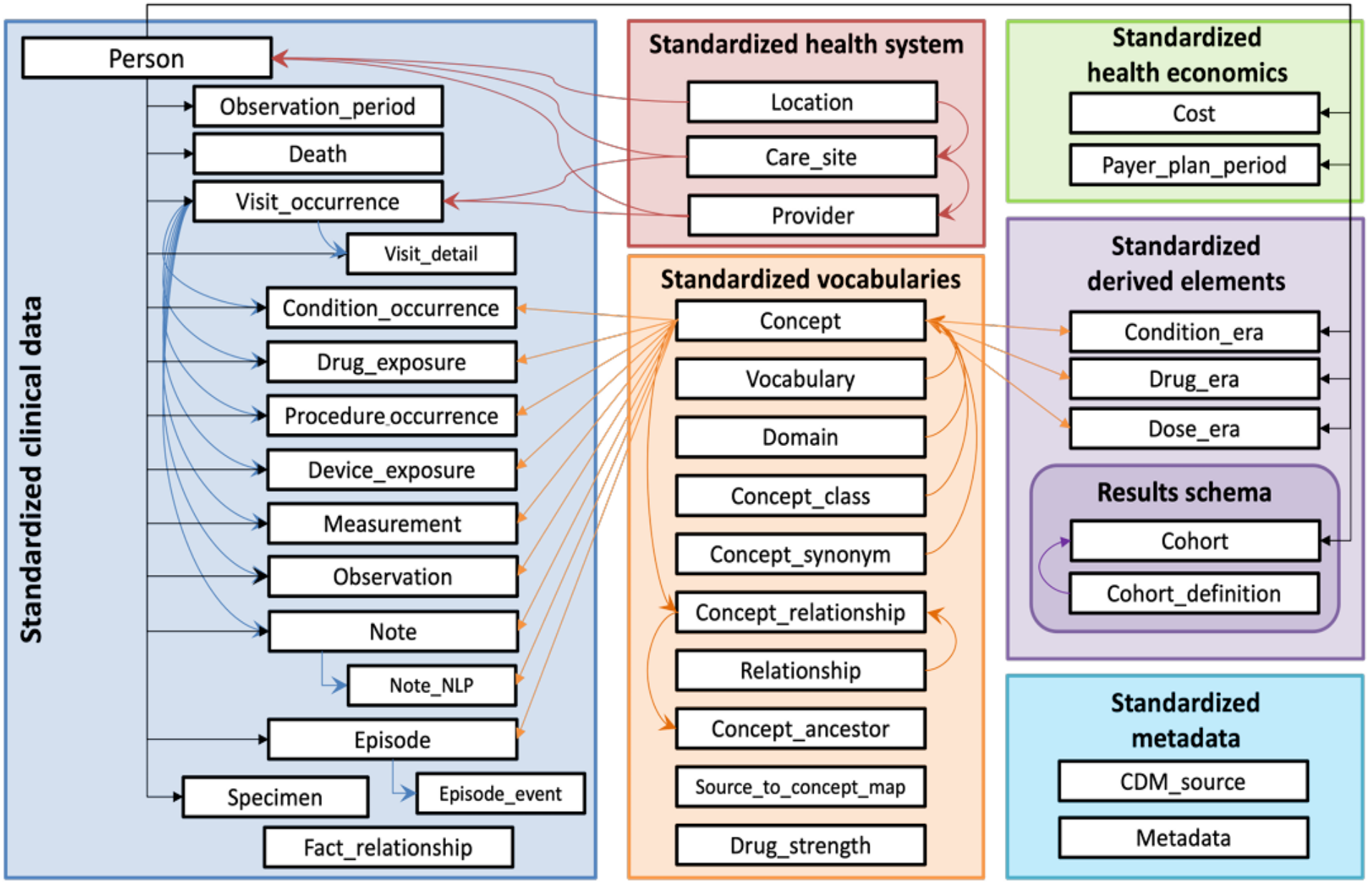
CDM Leads to Consistency in Data Structures





astellas

OMOP Common Data Model



Other Benefits of Common Data Model

Data Integration and Interoperability: Without a common framework, integrating data from various sources can be laborious and error-prone. A CDM helps to unify data, fostering seamless integration and interoperability among different data systems.

Scalability and Flexibility: A CDM provides a scalable and flexible foundation for data infrastructure. As the volume of data or the number of data sources grows, a CDM can help manage this complexity and support efficient expansion.

Data Governance and Quality: A CDM enhances data governance by providing standard definitions and formats, thereby reducing ambiguity and improving data quality. This also simplifies compliance with various data regulations.

Improved Collaboration: A CDM aids in collaboration among teams or departments within an organization, as well as between different organizations. Everyone is working with the same set of definitions and structures, making it easier to share, understand, and use the data.

astellas

Federated Data

astellas

What is it?

Federated data is a decentralized data-sharing approach where data is kept in its original location and accessed through a common interface.

Federated data analysis in medical Real World Evidence (RWE) research refers to the approach where data from multiple sources are aggregated to answer research questions.

Data owners maintain control over their data, and data access is restricted to authorized users.

astellas

Advantages of federated data

Data privacy is an integral part of the design

Data privacy is maintained, as individual-level data never leave the data owner's control.

Data owners have control over their data and can decide who has access to it.

Federated data allows for large-scale data analysis, as it allows for compliance with local data use regulations.

astellas

Challenges of federated data

Varied formats and definitions from multiple sources make data integration a formidable challenge.

Inconsistent quality and missing data across sources undermine the reliability of analyses.

Managing and governing data becomes increasingly complex within a federated environment.

astellas

Examples of federated data initiatives in healthcare

EHDEN
EUROPEAN HEALTH DATA & EVIDENCE NETWORK

The European Health Data and Evidence Network (EHDEN) is a federated data initiative that aims to harmonize health data across Europe.

DARWIN EU

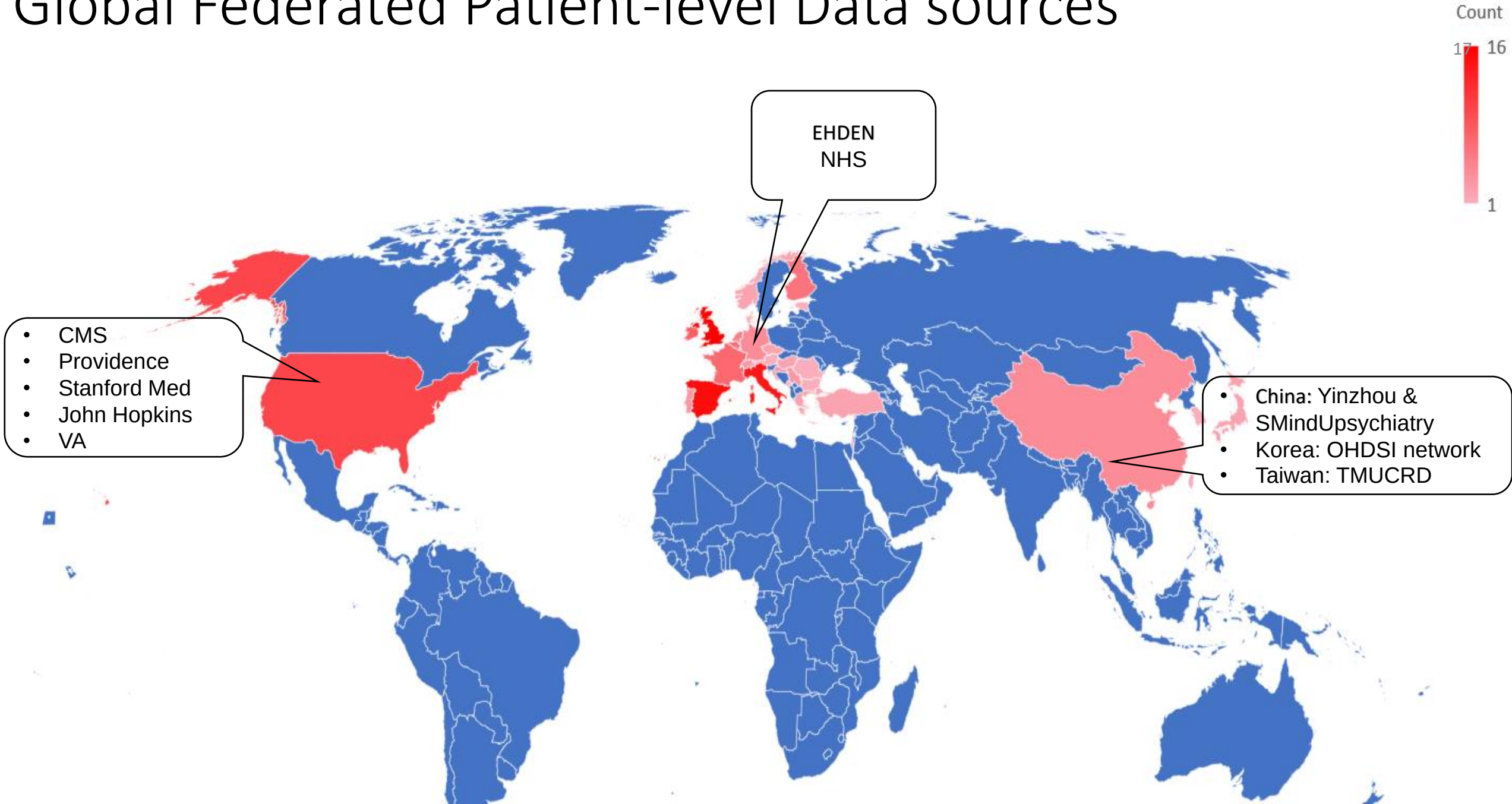
DARWIN EU® is a federated network of data, expertise, and services that supports better decision-making throughout the product lifecycle by generating reliable evidence from real-world healthcare data.

PIONEER

The Pioneer IMI project is a European project that aims to establish a framework for real-world evidence generation and evaluation for medical PCa products. It is a collaborative effort between industry, academia, and regulatory authorities.

astellas

Global Federated Patient-level Data sources



astellas

EHDEN
EUROPEAN HEALTH DATA & EVIDENCE NETWORK

Vision

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

Mission

The 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 standardized to a common data model

Start date: 1 Nov 2018
End date: 30 Apr 2024
Duration: 66 months

23 partners

Almost €29 million

astellas

Universities, public bodies and research organisations

Erasmus MC Universitair Medisch Centrum Rotterdam — **Academic coordinator**

Uppsala Monitoring Centre

universidade de aveiro

NICE National Institute for Health and Care Excellence

Tartu Ülikool 1632 Universitas Tartuensis

University of Oxford

SME & Mid-sized companies

SYNAPSE Research Management Partners

ODYSSEUS Data Services Inc

The Hyve

Non-profit organisations

EPF European Patients Forum

ICHOM

EFPIA & Associated partners

Janssen Pharmaceutical Companies of Johnson & Johnson

Bayer

Celgene

Pfizer

Lilly

abbvie

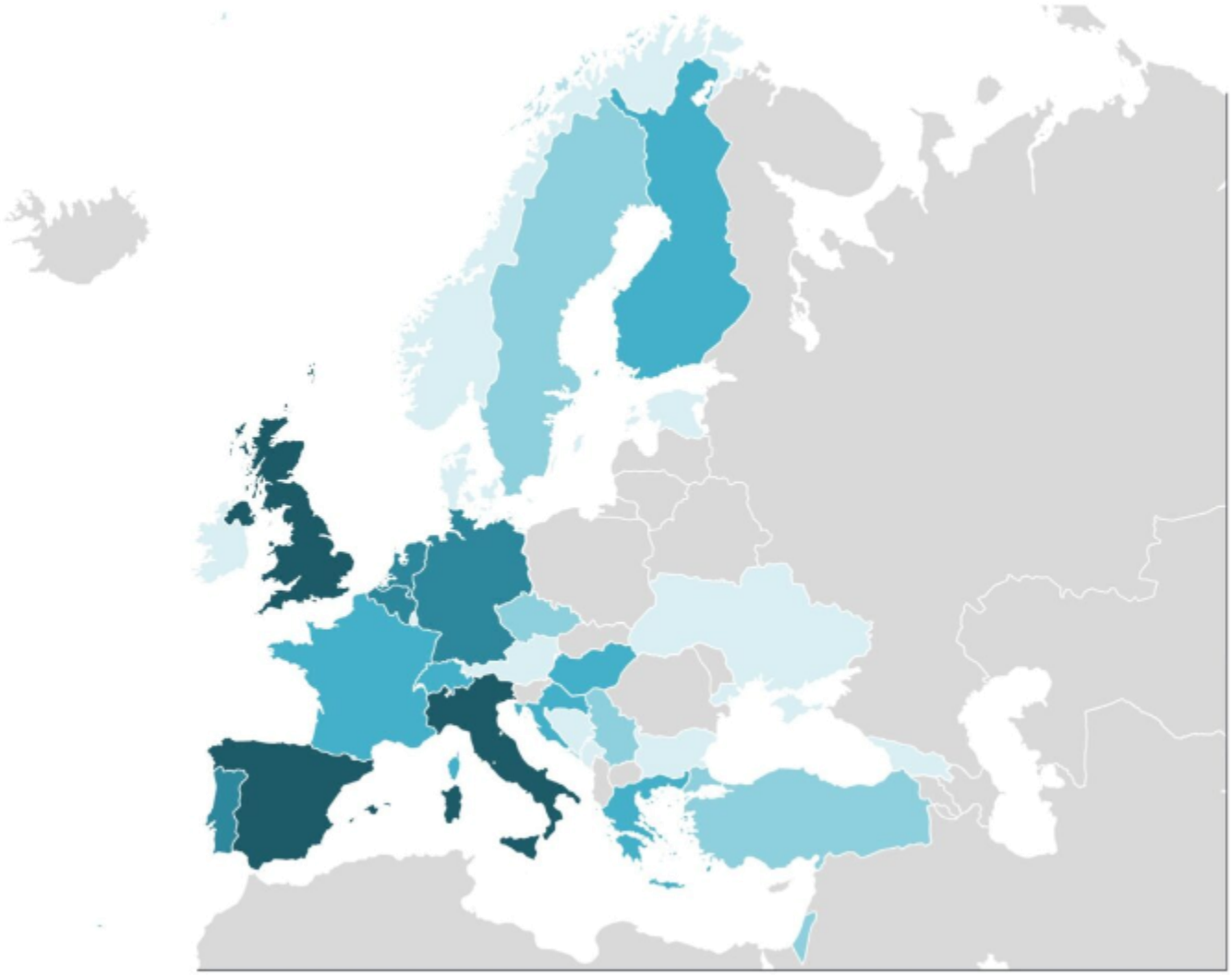
AstraZeneca

Sanofi

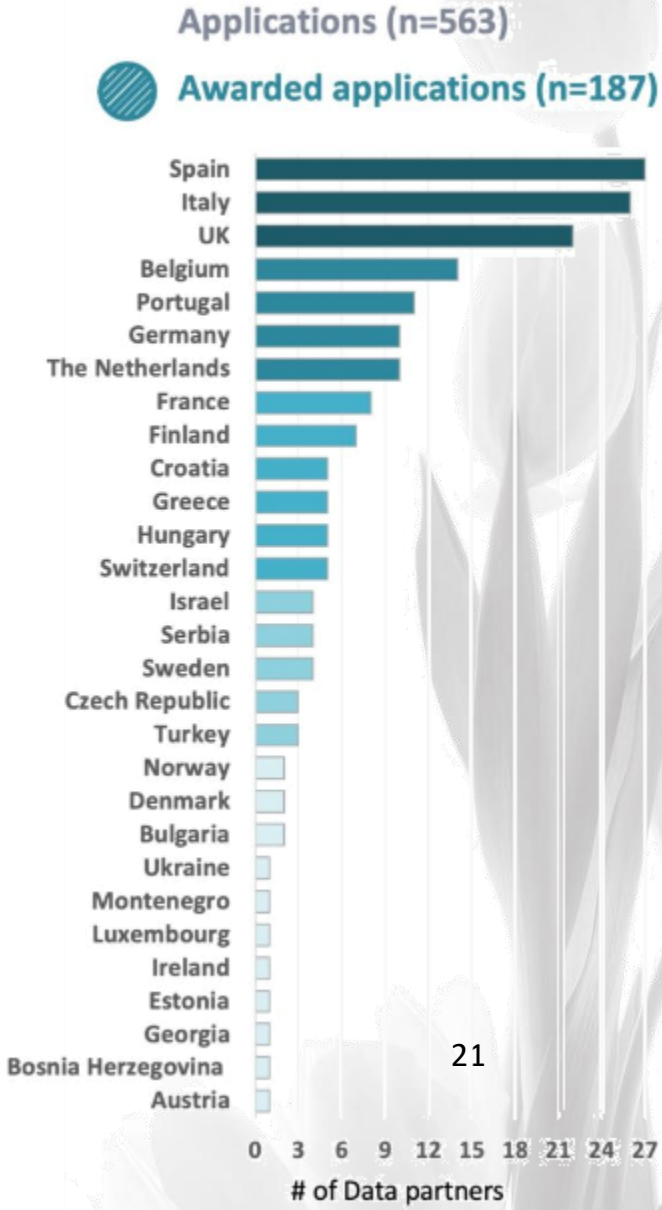
UCB

Novartis

Servier

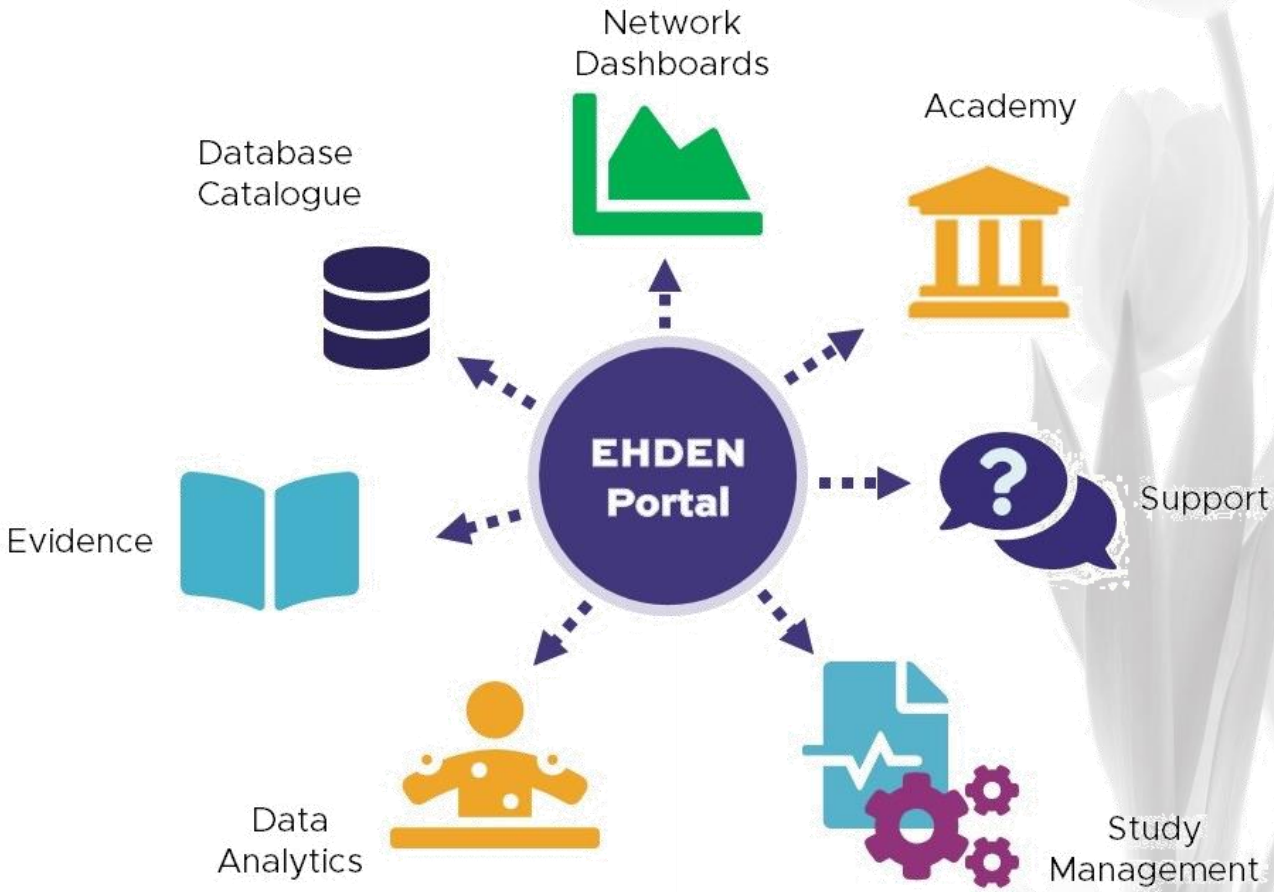


Geographic spread of data partners. The shade of blue indicates the # of data partners in that country (darker = more)



astellas

EHDEN is building the sociotechnical architecture to facilitate its European federated network and the research workflow from discovery to analysis within the Findable, Accessible, Interoperable and Reusable (FAIR) principles. Called the EHDEN portal.



PIONEER and the BD4BO mission

PIONEER is part of the Innovative Medicine Initiative's (IMI's) "Big Data for Better Outcomes" (BD4BO) programme

4 Disease-specific projects:

- **Roadmap**: Alzheimer's, ***HARMONY**: haematologic malignancies, ***BigData@heart**: Cardiovascular, ***PIONEER**: Pca

The **BD4BO mission** is to improve health outcomes and healthcare systems in Europe by maximising the potential of Big Data

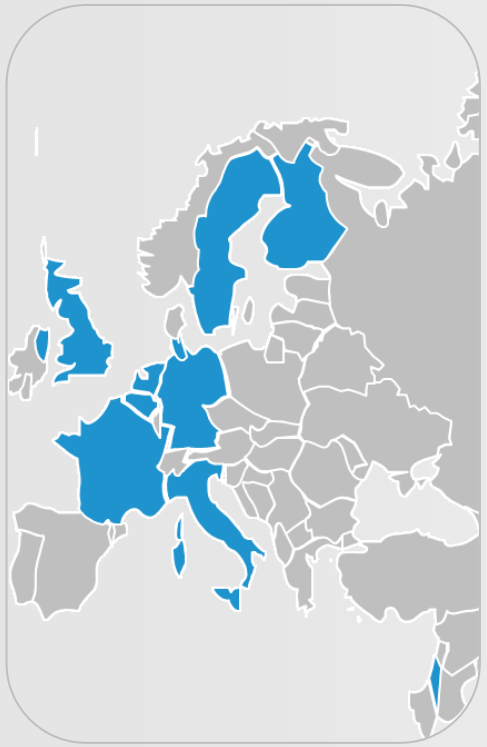
PIONEER aims to transform the field of prostate cancer care with particular focus on:

improving prostate-cancer related outcomes

health system efficiency

the quality of health and social care across Europe

PIONEER Consortium



36 private and public stakeholders in prostate cancer research and clinical care from across 9 countries

Technische Universität Dresden
Bayer
Sanofi
European Cancer Patient Coalition
Pinsent Masons
eisbm European Institute for Systems Biology & Medicine
ttopstart a PNO life sciences & health company
Lund University
Göteborgs Universitet
Astellas
ecancer
thehyve
janssen
European Alliance for Personalised Medicine
IHE The Swedish Institute for Health Economics
Imperial College London
IQVIA
European Association of Urology
Radboudumc university medical center
Orion
מכון ויצמן למדע Weizmann Institute of Science
King's College London
Erasmus MC
Covance Solutions Made Real
ICHOM
University of Tampere
sas
EORTC
University of Aberdeen
AstraZeneca
Fraunhofer IZI
Universitätsklinikum Hamburg-Eppendorf

PIONEER's primary objective

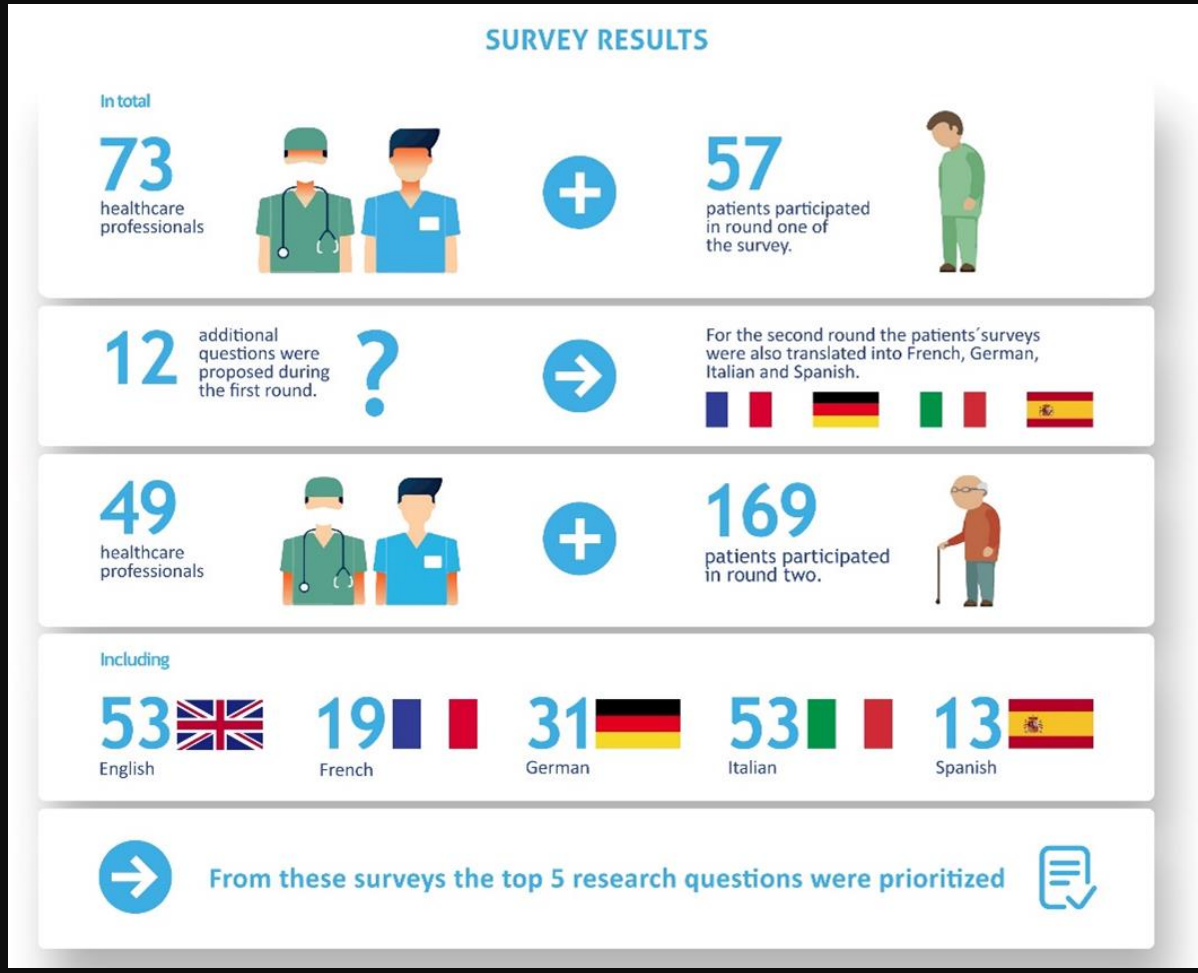
- By applying advanced data analytics, and developing a data-driven platform of unparalleled scale, quality and diversity,

  - PIONEER will work towards **meaningful improvement in clinical practice, PCa disease-related outcomes, and health-economic outcomes** across the European healthcare landscape

  - PIONEER will assemble, standardise, harmonise and analyse high-quality big data from diverse populations of PCa patients across different stages of the disease to provide **evidence-based data for improving decision-making** by key stakeholders

Planned outcomes

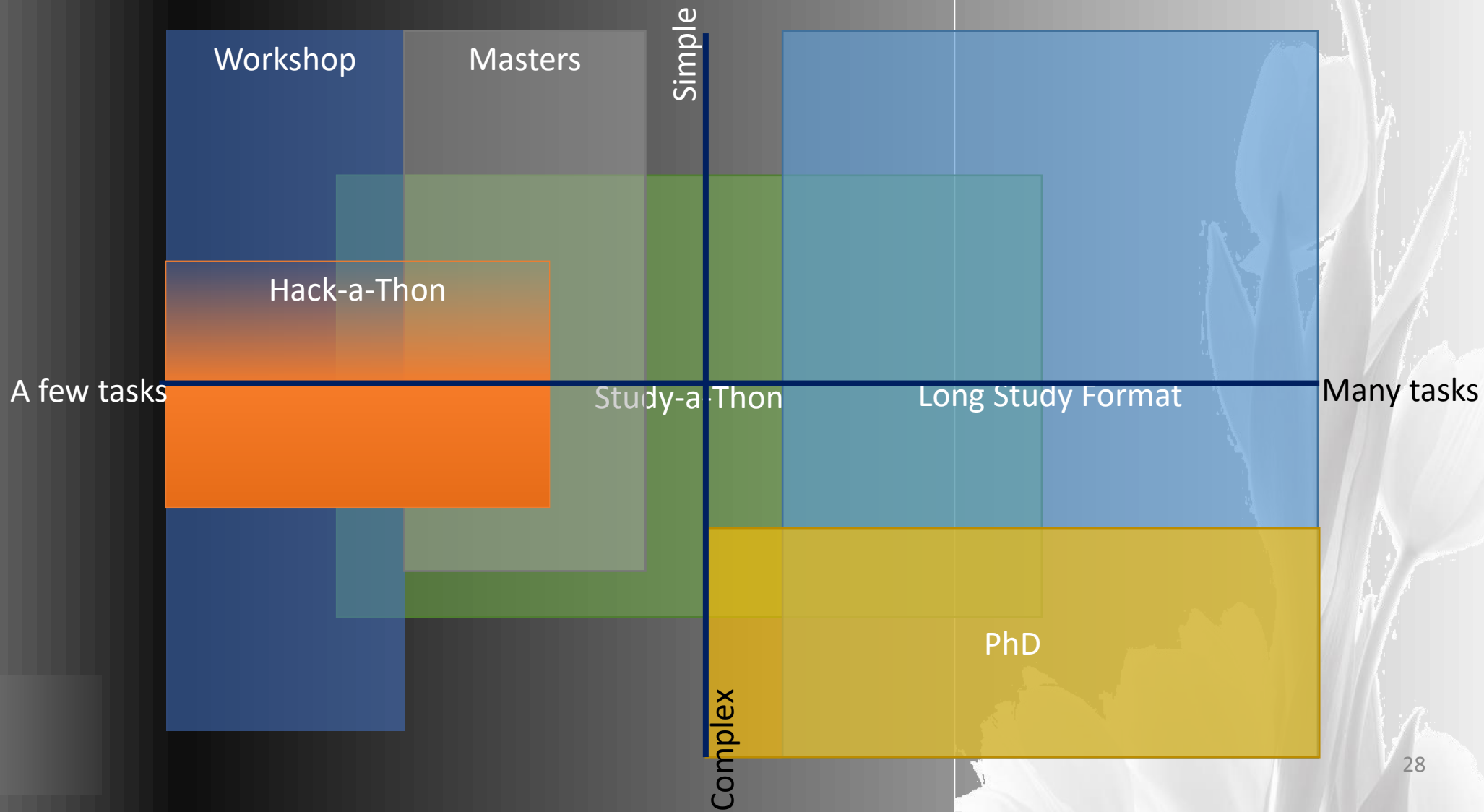
- Consensus on the most important **prostate cancer outcomes**
- Identification of **critical evidence gaps** in prostate cancer
- Standardisation of **definition and outcome measures**
- New insights into improved **stratification**
- Improved **standardised care pathways** with better predictable outcomes



PIONEER Research Questions

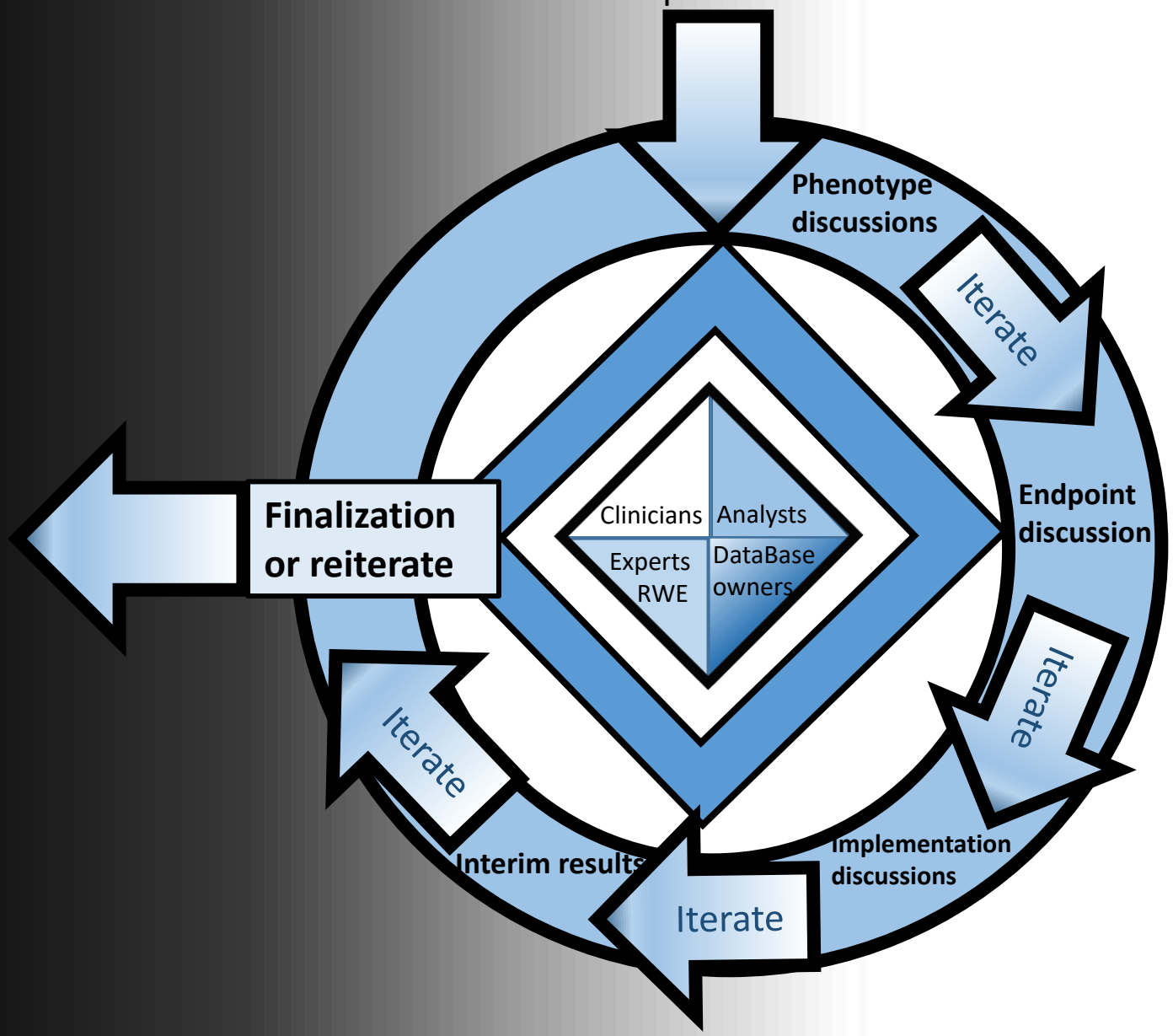
RQ #	Title	Lead
In preparation		
2	**Which are the long-term outcomes of prostate cancer patients undergoing non-interventional management (i.e., watchful waiting) and what is the impact of comorbidities and life expectancy?**	Giorgio Gandaglia 1st Study-A-Thon
4	When should we treat patients who experience prostate cancer recurrence after primary treatment and which are the most effective therapeutic approaches?	EAU (Nicolas Mottet & Vasileios), Florence (Mauro Gacci & team), EFPIA (IQVIA)
5	**Which specific patient groups benefit most of upfront chemotherapy? What are the side effects and what is impact on quality of life in real-life practice of chemotherapy in this setting? The benefit of potentially toxic upfront chemotherapy appears to be highly individual. Other factors to predict who would benefit most are needed. The benefit of chemotherapy in the subgroup patients who have recurrence after primary treatment is not known.**	Robert, Bertrand, Juan 2nd Study-A-Thon
8	Which specific patient population benefits from the different available treatment options for nm CRPC	Bayer and SAS
	Imaging Team	Michael Bussman & Henkjan Huisman
30	Development of a reference model for economic evaluation. What are the most important outcomes across different parts of the prostate cancer care pathway? The outcome domains can be subdivised into the following groups: a. Oncological; b: Functional; c: Process of recovery; d: Complications and/or adverse events; e: Quality of life: f: Health economics and cost effectiveness"	linked to task 6.2 - Thomas Hofmarcher

Study formats and study questions



What is a studyathon?

- « A **short**, concentrated face-to-face gathering of **multidisciplinary** group of scientists aimed at answering an important, clinically relevant **research question** using the OMOP data model and the OHDSI tools. » (Kees van Bochove, The Book of OHDSI, section 3.2).



PIONEER Study-A-Thon - Workflow

Preparation

Protocol development
Phenotyping
Analysis plan
System build
Literature review

Monday

Introduction day
Split the team in small groups
Start drafting literature review
Create cohort diagnostics

Tuesday

Review cohort diagnostics
Run clinical characterization
Draft patient-level prediction design

Wednesday

Review characterization results
Execute prediction package

Thursday

Continue clinical characterization
Review patient-level prediction results
Externally validate prediction model

Friday

Review of results
Feedback on studyathon
Plan for completing publications

Wrap up

Package updates
Final result review
Manuscript
Deep dives

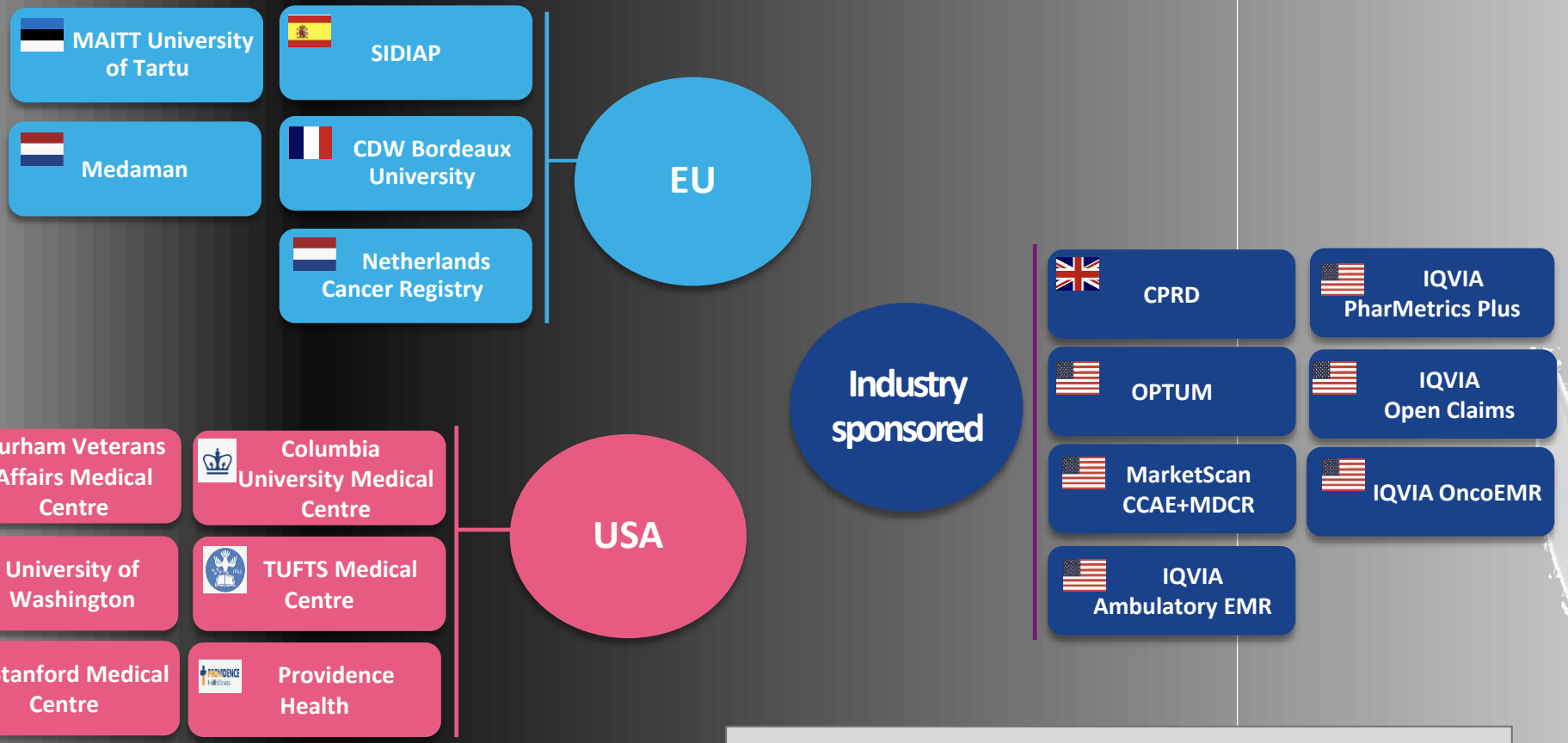
- 245 participants
- 4 patients
- 875 team engagements

- 20 countries
- 5 time zones
- 5 days

PIONEER 1st Study-A-Thon – Sub-teams

Sub-teams	Objectives
Phenotyping	Define the study phenotypes clearly, unambiguously and accurately to generate meaningfully evidence considering differences/nuances of the databases
Data network	Identify & recruit appropriate databases to the study; execute the R packages to conduct clinical characterisation and prediction analysis
Clinical characterisation	Describe the demographic and clinical characteristics of patients with prostate cancer under conservative management & estimated clinical outcomes of these patients including those who initiated treatment.
Prediction	Develop a prediction model that predicts an outcome (symptomatic progression, death, death without symptoms) at a specific moment in time (6, 12, 24 months) based on a combination of patient characteristics.

PIONEER 1st Study-A-Thon Data Network



18 data partners

- Academic EHR (EU & US)
- Community-based EHR (EU & US)
- Nationwide claims (US)
- Cancer registry (EU)

Publications

ABSTRACT | VOLUME 81, SUPPLEMENT 1, S1546-S1547, FEBRUARY 2022

Clinical characterization and outcomes of prostate cancer patients undergoing immediate vs. conservative management: A PIONEER study

Gandaglia G. • Omar M.I. • Maresca G. • ... Smith E.J. • N'Dow J. • PIONEER Consortium • Show all authors

DOI: https://doi.org/10.1016/S0302-2838(22)01127-7

- Another paper titled “Baseline Characteristics and Clinical Outcomes of Prostate Cancer Patients on Delayed Palliative Management“ is in the final stages of drafting

Lessons learned and outlook

Challenges	Lessons learned

• Not all information available in all data sources:
  • E.g. Gleason Score / PSA level available only in some data sources
  • Unavailability/under-reporting of death and clinical progression (e.g. claims)
	➢ Data fitness to answer a specific RQ to be ascertained upfront

• Differentiation between WW and AS not possible in RWD: • Phenotype based on lack of event (i.e. health care encounter in the data)	➢ This objective was not met ➢ Focus on RQ questions that can be answered via RWD

Opportunities

- Potential for advanced analytics and large-scale studies in multinational networks
- Generation of evidence across a network of data from US and EU - geographical representation
- Study-a-thon is just one modality of conducting network studies in Pioneer, ad-hoc studies are another
- Patient representative participation: patients' voice can be included

Studyathon 2

16 datasets

Study package successfully completed in 10

- Hybrid event
- Leiden November 2022
- 35 participants including patients
- 10 countries
- 5 days hybrid event

Research question:

- Which specific patient will benefit the most according to the different treatment schemes in metastatic hormone-sensitive prostate cancer?
- Two parts – characterisation & prediction

The crew

Studyathon 3

20 datasets identified

2 test R packages successfully run
on 5 datasets

- Hybrid event
- Stockholm May 2023
- 53 participants including patients
- 9 countries
- 4 days hybrid event

Research goal:

- An observational health data analysis on the adverse events of systemic treatment in patients with metastatic hormone-sensitive prostate cancer (mHSPC)
- Systemic treatment: Androgen deprivation therapy (ADT); Taxane chemotherapy; and Androgen receptor Signalling Inhibitor (ARSI)

Learnings

The following allow for faster protocol development and results

Multi-day F2F meetings benefit the process

Empowered study team consisting of decision makers

Iterative flexible approaches and rapid decision making

Bring together diverse stakeholders with different expertise

Breakout teams with specialized focus and core study team

Other learnings

Be science focused and build credibility

Speak up and do not be afraid to challenge assumptions