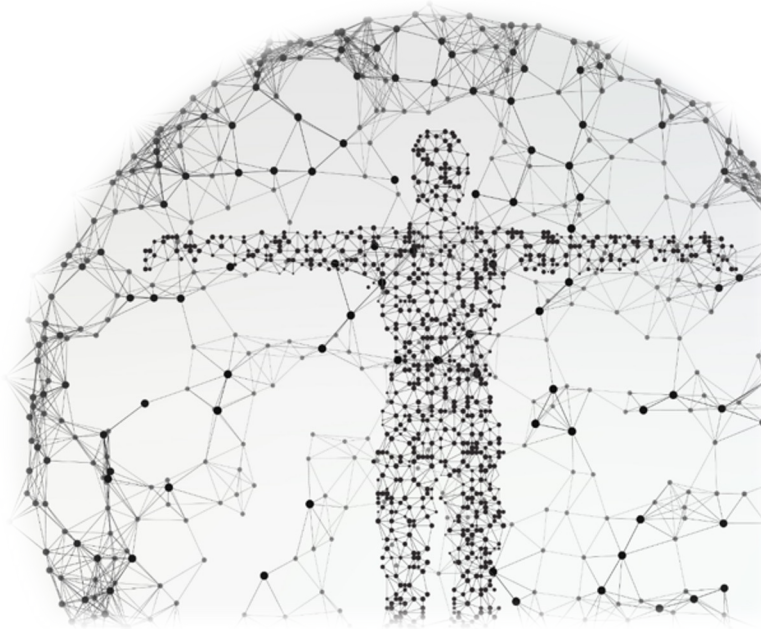
Coordination Centre

Introduction of the DARWIN EU® Coordination Centre

Prof. Peter R. Rijnbeek
Executive Director

Chair Department of Medical Informatics Erasmus MC



Disclaimer

This presentation represents the views of the DARWIN EU® Coordination Centre only and cannot be interpreted as reflecting those of the European Medicines Agency or the European Medicines Regulatory Network.

By 2025 the use of Real-World Evidence will have been enabled and the value will have been established across the spectrum of regulatory use cases

- European Medicines Regulatory Network (EMRN) strategy to 2025 -

Enabling use & establishing the value of RWE

Clin Pharmacol Ther. 2022 Jan;111(1):21-23. doi: 10.1002/cpt.2479.

- Facilitating access
- Build business processes
- Set standards
- Validate methods
- Train/share knowledge
- Establish value across use cases
- International collaboration:
  - build on ICMRA → RWE statement: 4 collaboration areas
  - ICH RWE reflection paper 'International harmonisation of real-world evidence (RWE) terminology, and convergence of general principles regarding planning and reporting of studies using real-world data, with a focus on effectiveness of medicines' →public consultation

PERSPECTIVES

PERSPECTIVE

Real-World Evidence in EU Medicines Regulation: Enabling Use and Establishing Value

Peter Arlett[1]*, Jesper Kjær[2], Karl Broich[3] and Emer Cooke[1]

We outline our vision that by 2025 the use of real-world evidence will have been enabled and the value will have been established across the spectrum of regulatory use cases. We are working to deliver this vision through collaboration where we leverage the best that different stakeholders can bring. This vision will support the development and use of better medicines for patients.

Real-world data (RWD) and real-world evidence (RWE) are already used in the regulation of the development, authorization, and supervision of medicines in the European Union. Their place in safety monitoring and disease epidemiology are well-established while their evidentiary value for additional use cases, notably for demonstrating efficacy, requires further evaluation.[1] During the coronavirus disease 2019 (COVID-19) pandemic, RWE rapidly provided impactful evidence on drug safety, vaccine safety, and effectiveness and we were reminded of the importance of robust study methods and transparency.[2] Our vision, anchored in the European Medicines Regulatory Network (EMRN) strategy to 2025, is that by 2025 the use of RWE will have been enabled and the value will have been established across the spectrum of regulatory use cases.[3] Delivering this vision will support the development and use of better medicines for patients.

In December 2018, the US Food and Drug Administration (FDA) published its framework for RWE underpinned by three pillars: whether RWD are fit for use, whether the study design can provide adequate evidence, and whether the study conduct meets regulatory requirements.[4] In 2019 in the European Union, we published the OPTIMAL framework for RWE also consisting of three pillars: operational, technical, and methodological.[5] More recently, the EU approach places RWE in the wider context of big data and is guided by the policy recommendations of the Big Data Task Force. These recommendations are being implemented through the Big Data Steering Group and the second multi-annual work plan was published in August 2021.[6] Figure 1 represents the workplan with its 11 workstreams which will deliver our vision for RWE by 2025. The workplan places emphasis on collaboration across stakeholders and with international regulatory partners. This work also needs to be seen in the wider EU policy context, most notably the European Commission's plans for a European Health Data Space.[7]

Acknowledging different frameworks to conceptualize the challenges and opportunities of RWE, we believe the two main priorities for the European Union are to enable its use and establish its value for regulatory decision making. The EMRN is working to deliver on both priorities through a collaborative approach where we leverage the best that different stakeholders can bring, and where those stakeholders can complement the central role of industry in generating evidence.

ENABLING USE

To enable use, we are working on multiple fronts with our stakeholders, including patients, healthcare professionals, industry, regulatory and public health agencies, health technology assessment bodies, payers, and academia. We are initiating work to establish a data quality framework, not just for RWD but for all data used in regulatory decision making. We are striving to improve the discoverability (findability) of RWD through agreement of metadata for RWD and through a public catalogue of RWD sources[8] that builds on the early work of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). The ENCePP Guide on Methodological Standards in Pharmacoepidemiology,[9] extensively updated in 2021, is the core of our efforts to drive up the standards of study methods for RWE, and this is complemented by recently published guidance on conducting studies based on patient registries.[10]

The European Medicines Agency (EMA) and some national medicines agencies

[1]European Medicines Agency, Amsterdam, Netherlands; [2]Danish Medicines Agency, Copenhagen, Denmark; [3]BfArM, Bonn, Germany. *Correspondence: Peter Arlett (Peter.Arlett@ema.europa.eu)

Received March 1, 2021; accepted November 1, 2021. doi:10.1002/cpt.2479

Towards delivering the 2025 RWE vision

Countdown to 2025

Enabling use

EMA studies using in-house databases

- **Primary care** health records from the **France**, **Germany, UK, Italy, Spain** and **Romania.** Some data sources include data on specialist.

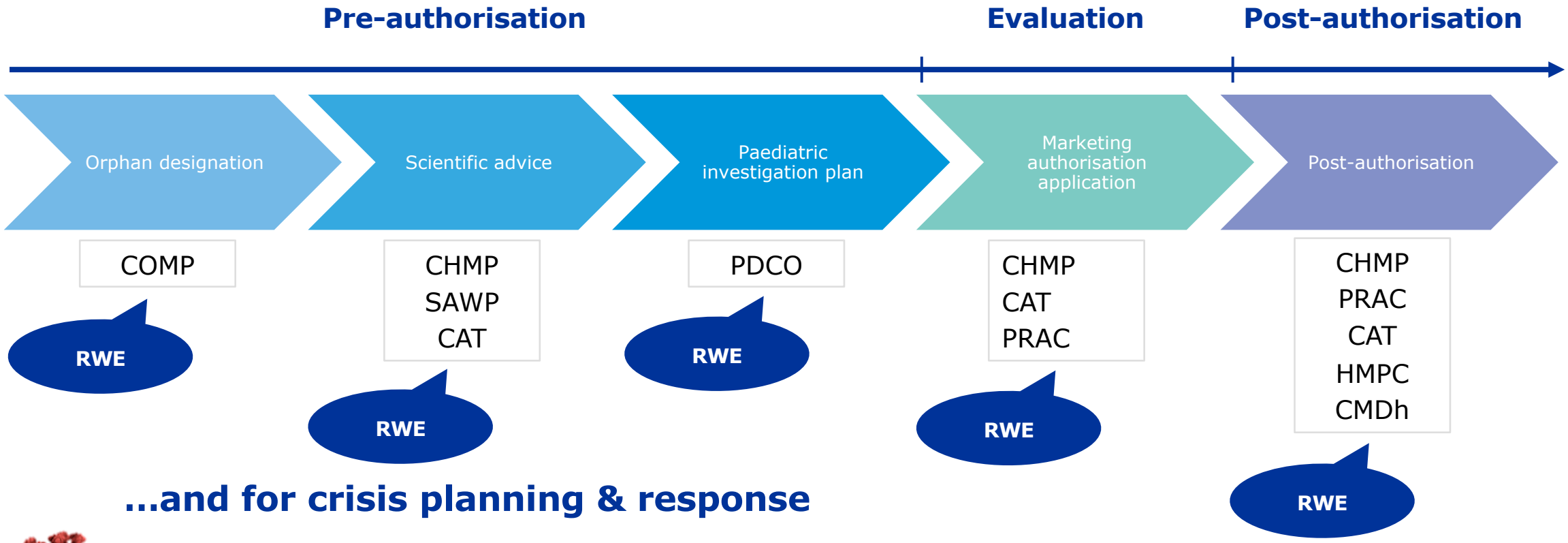
Studies procured through EMA FWCs

- New framework contract (FWC) since September 2021: services of **8 research organisations** and academic institutes
- Access to **wide network of data sources**: 59 data sources from 21 EU countries
- Ability to leverage external **scientific expertise**

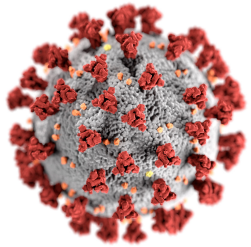
DARWIN EU®

- Coordination Centre launched February 2022
- Onboarded first **10 data partners**
- **First studies** finalised
- Additional 10 data partners are foreseen to **be added each year** for 2023-2025

Demand: RWE use across the medicinal product lifecycle

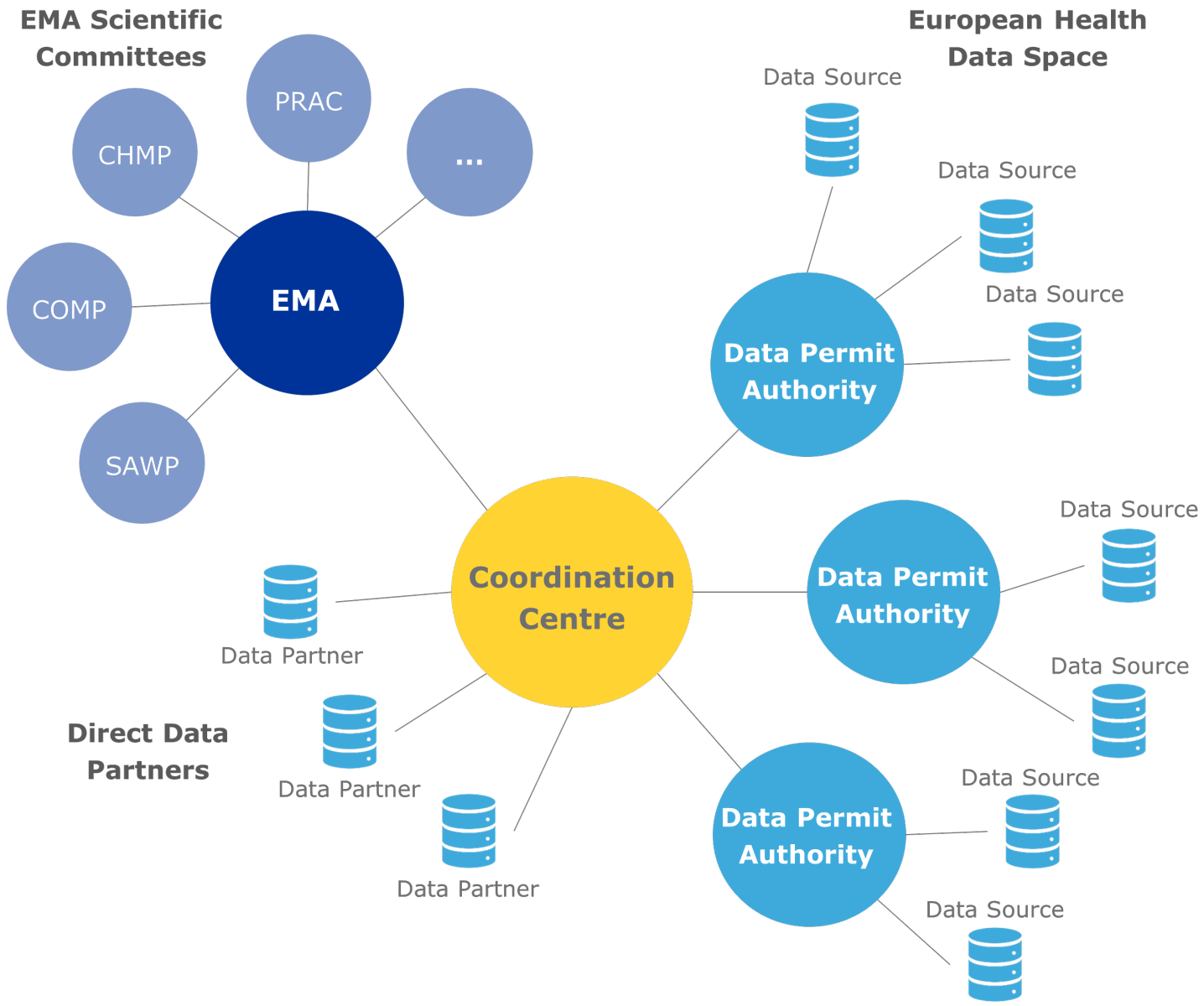


...and for crisis planning & response



- Monitoring the use of medicines to predict demand and shortages
- Understanding the disease natural history → development of vaccines and therapeutics
- Provide evidence for repurposing existing medicines
- Monitor the safety and effectiveness of vaccines and therapeutics post-authorisation

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**



FEDERATED NETWORK PRINCIPLES

- Data stays **local**
- **Use of OMOP Common Data Model** (where applicable) to perform studies in a timely manner and increase consistency of results

DARWIN EU® Coordination Centre

Executive Director
Prof. Peter Rijnbeek
Head of the Department of Medical Informatics
Erasmus MC

Deputy Director
Prof. Daniel Prieto Alhambra
Erasmus MC, Oxford University

Deputy Director
Associate Prof. Katia Verhamme
Erasmus MC

Contractor

Erasmus MC
Universitair Medisch Centrum Rotterdam

Sub-contractors

University of Oxford

SYNAPSE
RESEARCH MANAGEMENT PARTNERS

IQVIA

The Hyve

ODYSSEUS
DATA SERVICES INC

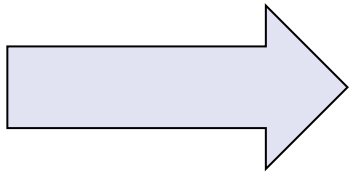
DARWIN EU® establishment in 2022 and 2023

- ✓ 2nd year of establishment in progress, delivery on target and according to plan
- ✓ Focus on selection of further Data Partners and study conduct (various use cases)
- ✓ Establishment of standard analytical pipelines and codes

		Phase I	Phase II	Phase III	Option I	Option II
Studies	Off the shelf	2	6	30	60	60
	Routine repeated	1	6	30	60	60
	Complex study	1	4	12	24	24
	Very complex	0	0	0	1	1
Data Partners (total)		10	20	30	40	40

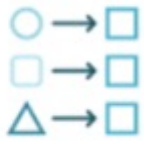
DARWIN EU® Implementing a paradigm shift

- A highly needed paradigm shift for the fast delivery of reliable evidence for regulatory decision-making on the utilisation, safety and effectiveness of medicinal products throughout their lifecycle
- A long-term investment needed to significantly scale up the number of studies on more databases and improve public health.

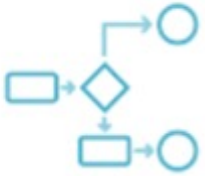


Not possible by simply scaling up the traditional approaches.

What is needed to facilitate observational studies at scale?



Data interoperability



Standardised analytics

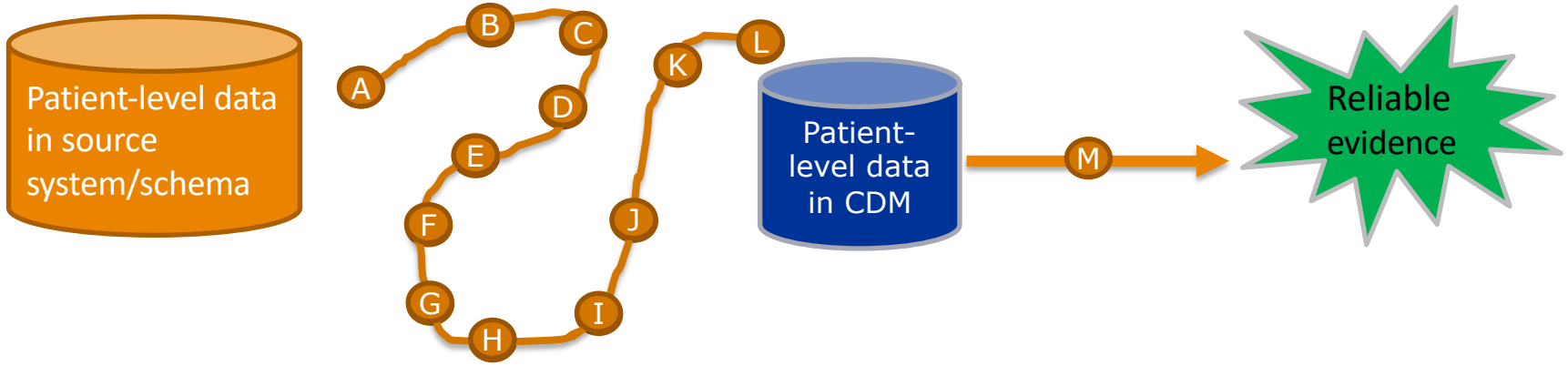
Technical Infrastructure

Data network

Research Memory.

Generating Reliable Evidence using the OMOP Common Data Model

We need to make studies repeatable, reproducible, replicable, generalisable, and robust



A Common Data Model enables standardised analytics to generate reliable evidence.

Operating a high-quality Data Network

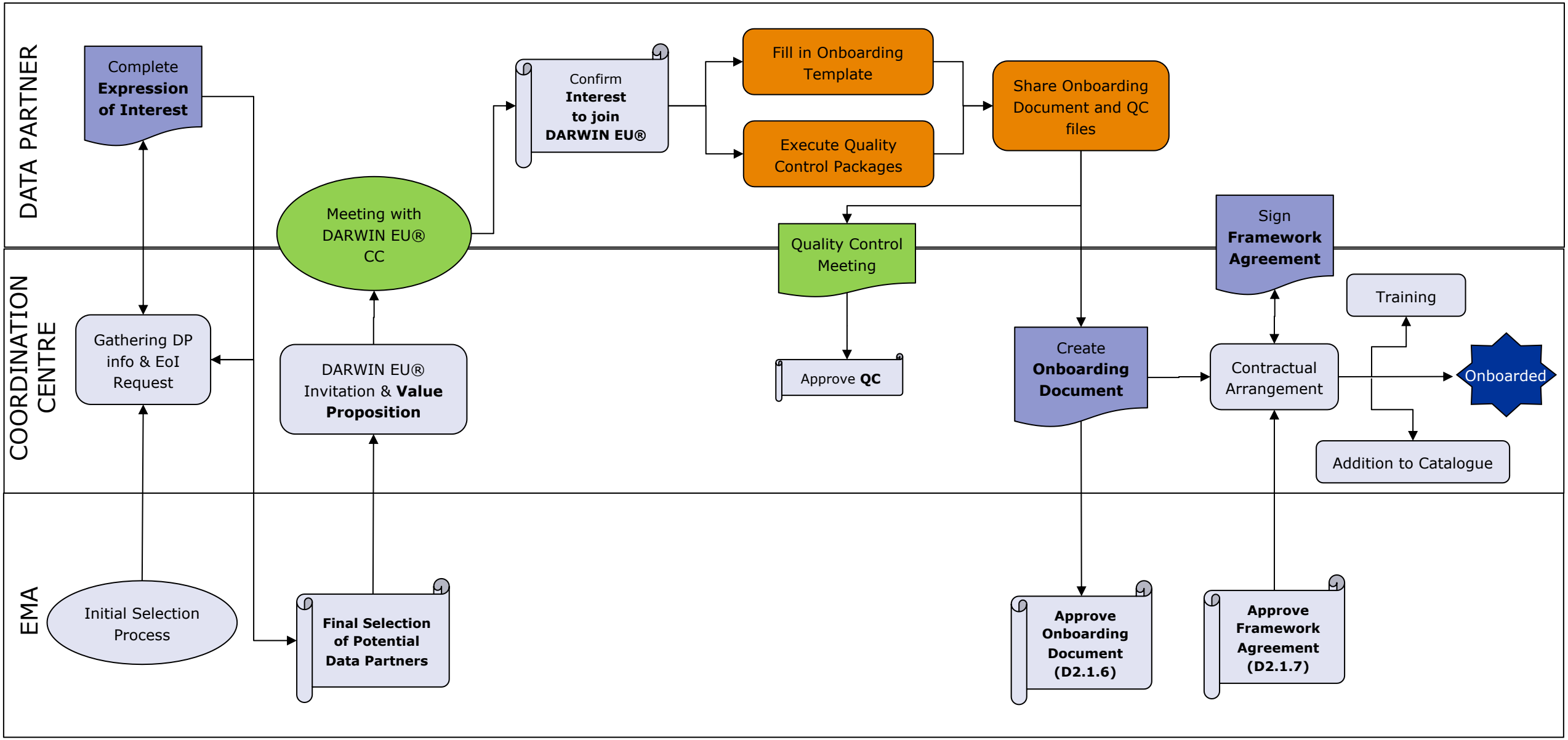
- Selection of data partners
  1) Prioritisation of already converted data sources
  2) Potentially mapping highly valued data sources

- All data sources will go through an onboarding process approved by EMA including quality control steps

See https://darwin-eu.org/index.php/data/how-to-join-the-network for more information.

Deadline Open Call for Expression of Interest: 31st October.

ONBOARDING PROCESS (already on OMOP CDM)



Data Partners – Phase I

UK

1. Clinical Practice Research Datalink (CPRD GOLD)

Belgium

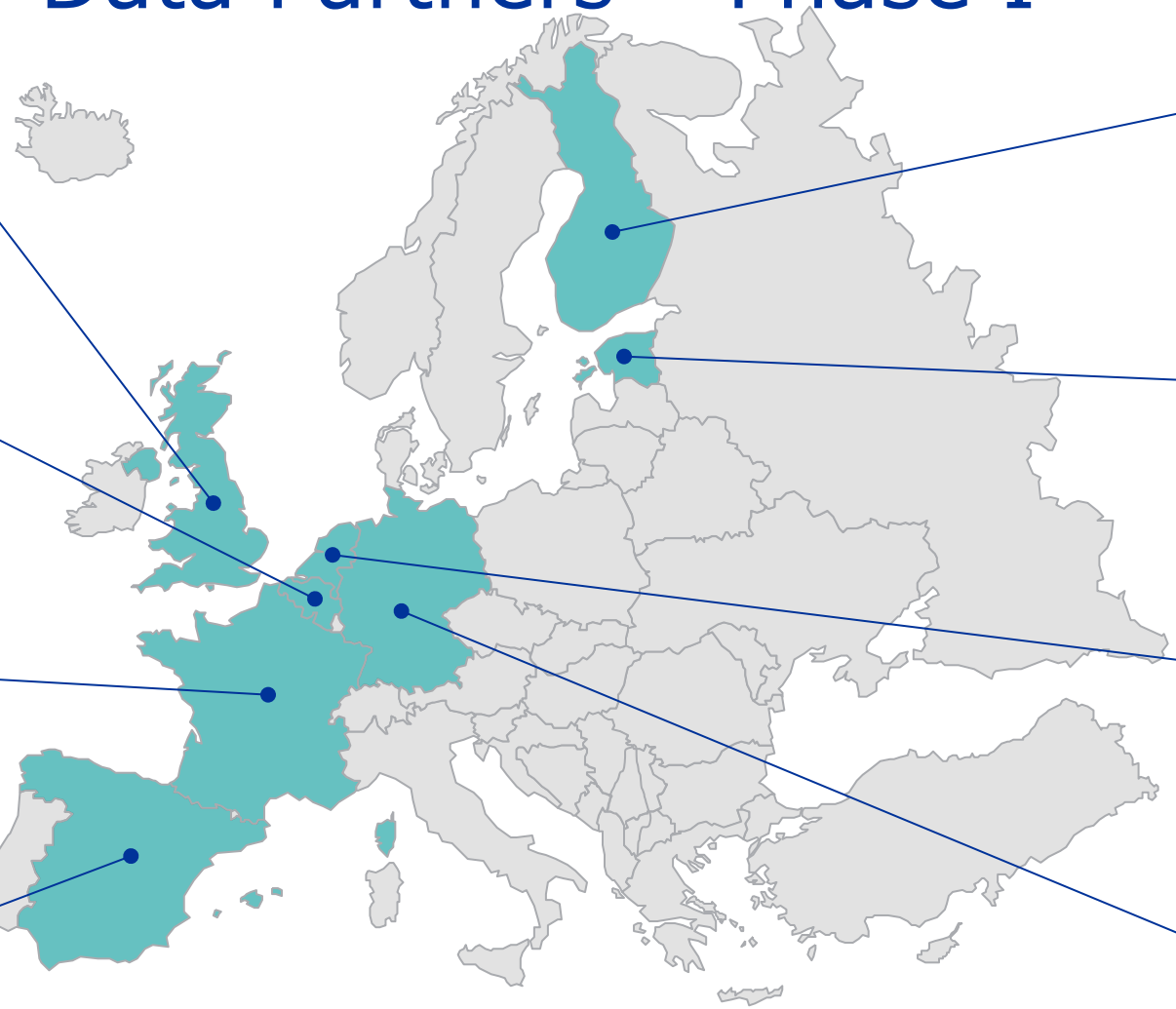
2. IQVIA Belgium Longitudinal Patient Data

France

3. Bordeaux University Hospital

Spain

4. IDIAPJGol
5. Parc Salut Mar Barcelona, Hospital del Mar (IMIM)



Finland

6. Auria Clinical Informatics at Hospital District of Southwest Finland (HDSF)

Estonia

7. University of Tartu (Biobank)

Netherlands

8. Integrated Primary Care Information
9. Netherlands Comprehensive Cancer Organisation

Germany

10. IQVIA Germany Disease Analyser

~26 million active patients

Currently **selecting Phase II DPs** via **open call for expression of interest**, then Phase III to follow

Standardising the analytics

A catalogue of open source standardised analytics is needed to support ”all” regulatory decision-making on the utilisation, safety and effectiveness of medicinal products

Will require alignment on the priority and choice of the analytical methods, and the standardised output!

https://www.darwin-eu.org/index.php/methods/standardised-analytics

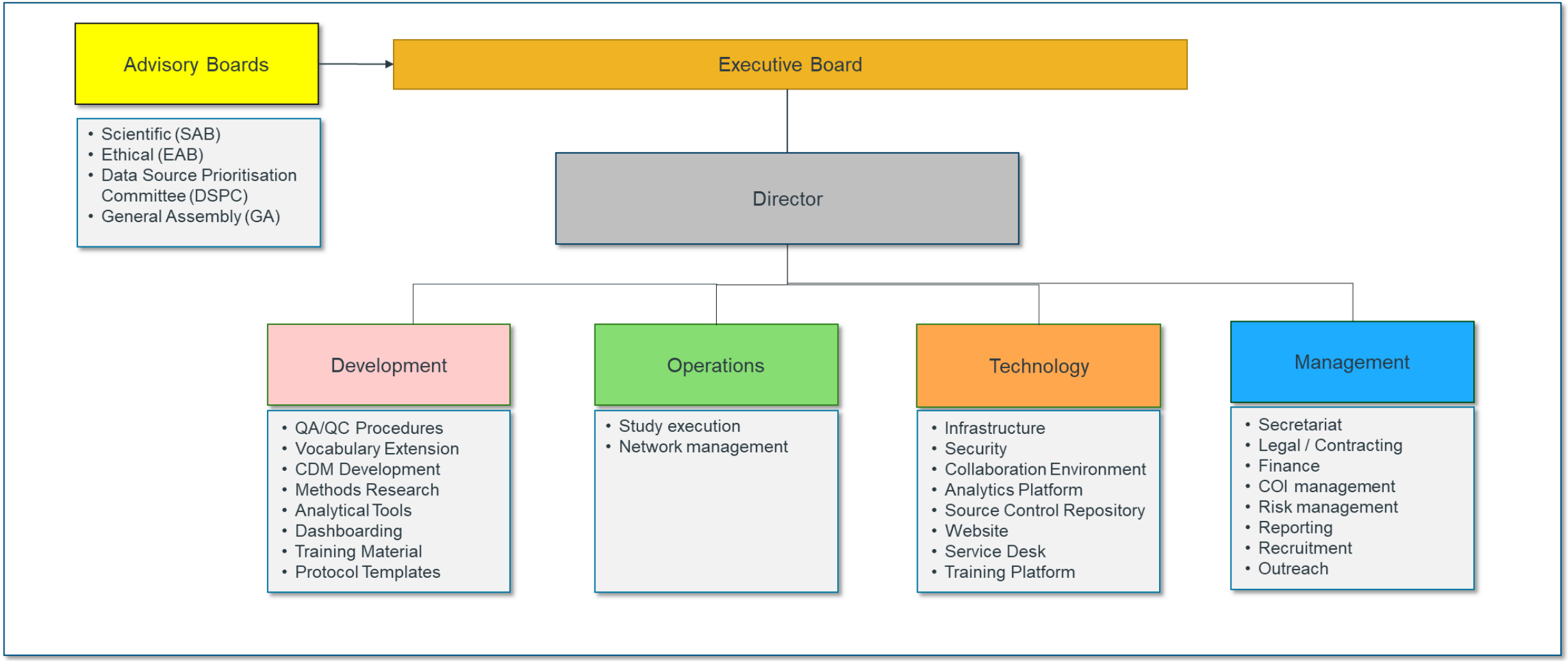
Creating a strong technical infrastructure

Required components:

- Collaboration Space for CC and Study Teams
- Analytics Platform
- Study Execution Platform
- Training Platform
- Service Desk
- Source Control Repository
- Phenotype Library
- DARWIN EU Website:
- …

https://www.darwin-eu.org/

Establishment and Evolution of the Coordination Centre



DARWIN EU

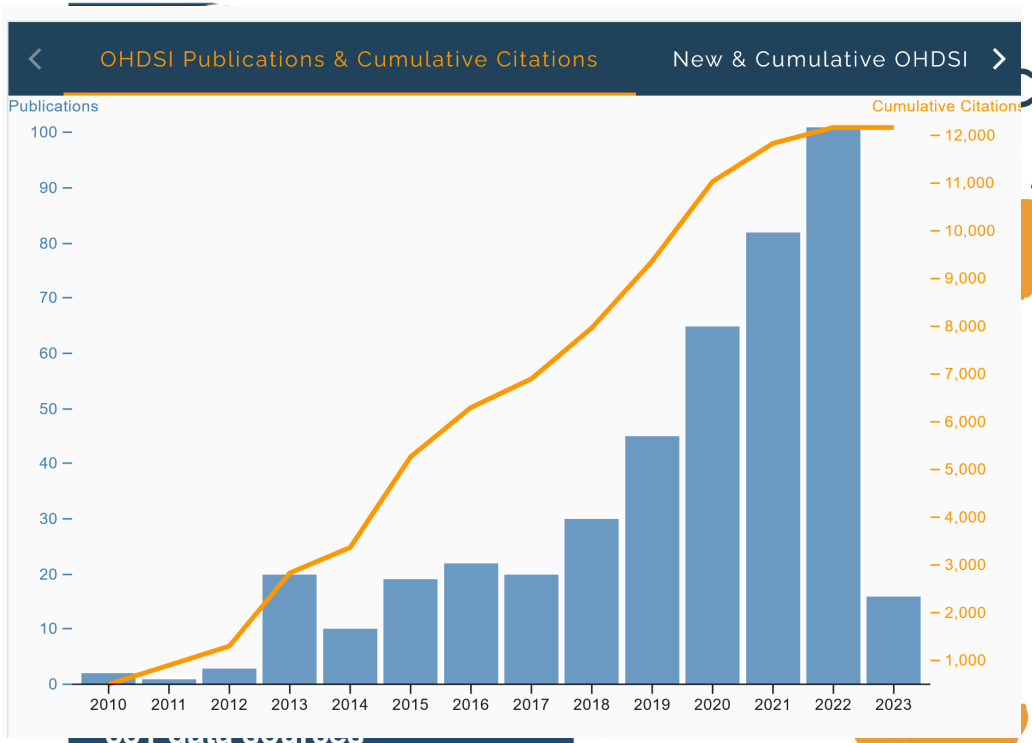
Coordination Centre

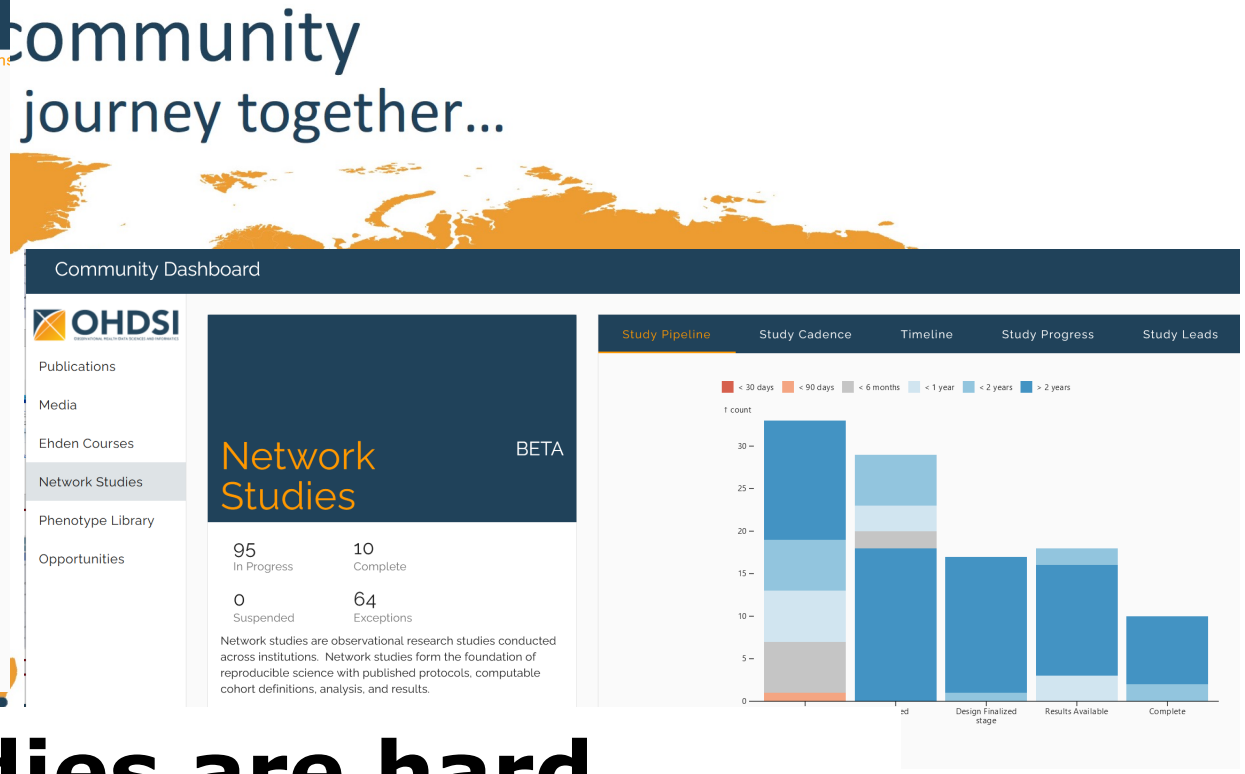
Development

Ed Burn

145 studies in 2025

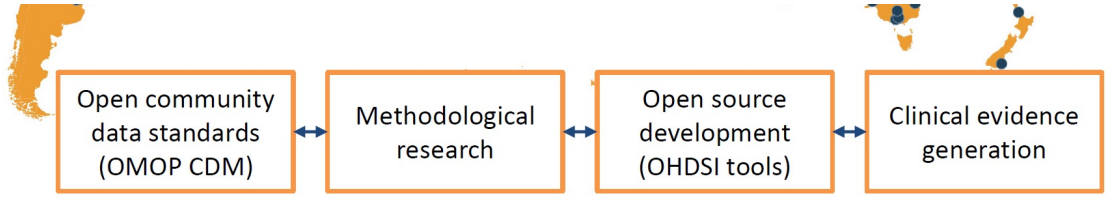
		Phase I	Phase II	Phase III	Option I	Option II
Studies	Off the shelf	2	6	30	60	60
	Routine repeated	1	6	30	60	60
	Complex study	1	4	12	24	24
	Very complex	0	0	0	1	1
Data Partners (total)		10	20	30	40	





- 284 EHRs
- 28 administr
- 34 countries
- 810 million unique patient records

Network studies are hard .....



Catalogue of Standard Data Analyses

Off-the-shelf studies

These are mainly characterisation questions that can be executed with a generic protocol. This includes disease epidemiology, for example the estimation of the prevalence, incidence of health outcomes in defined time periods and population groups, or drug utilization studies at the population or patient level.

- Patient-level characterisation
- Patient-level DUS analyses
- Population-level DUS analyses
- Population-level descriptive epidemiology

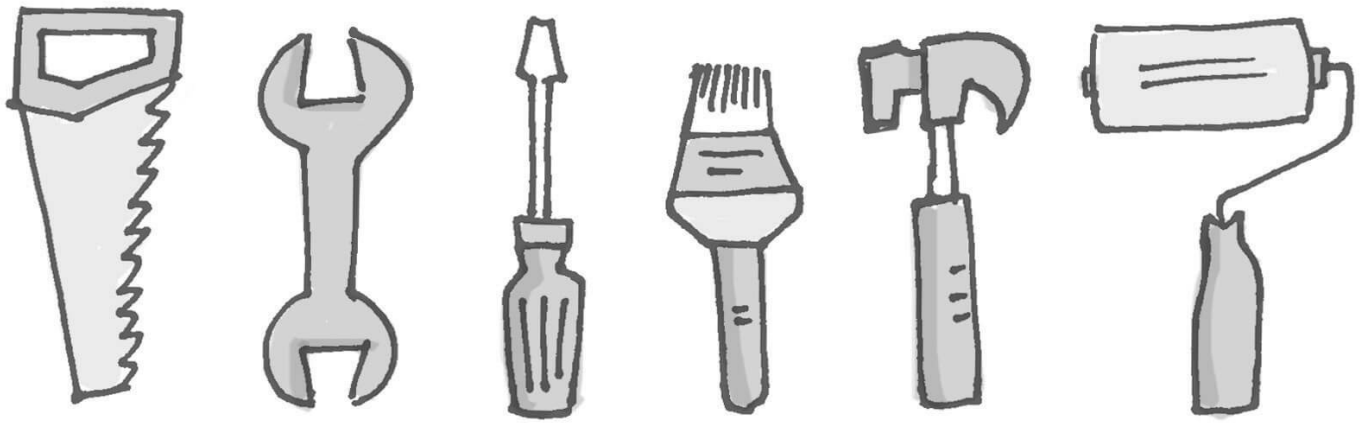
Complex

These are studies requiring development or customisation of specific study designs, protocols, analytics, phenotypes. This includes studies on the safety and effectiveness of medicines and vaccines.

- Prevalent user active comparator cohort studies
- New user active comparator cohort
- Self-controlled case risk interval
- Self-controlled case series
- Time series analyses and Difference-in-difference studies
- RMM effectiveness

Building tools

Primary focus of the development pillar is providing tools (mostly R packages) to help users to perform standard data analyses

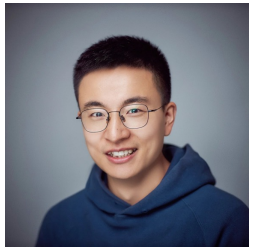


User profiles

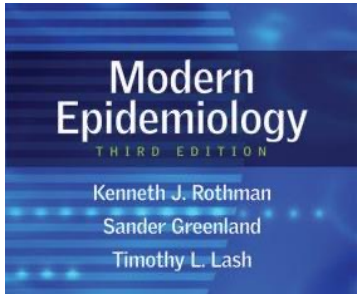


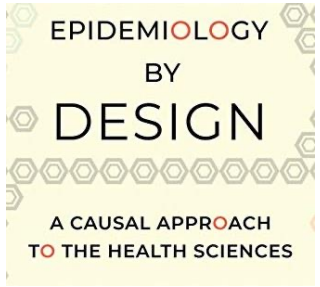


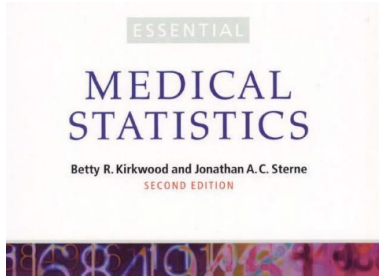


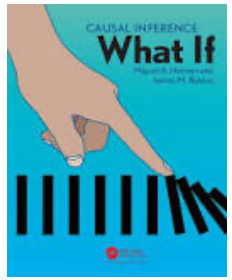


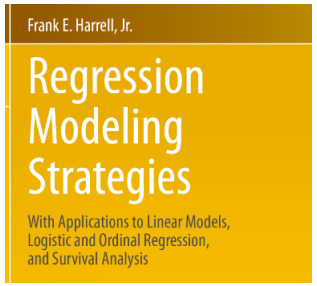
- Epidemiologists and data scientists
- Interact with our R packages directly, preparing analysis scripts that use them

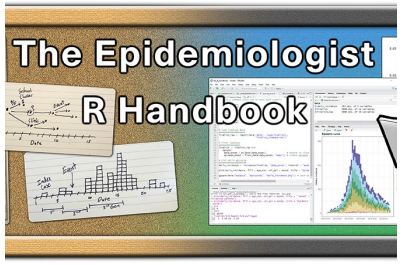










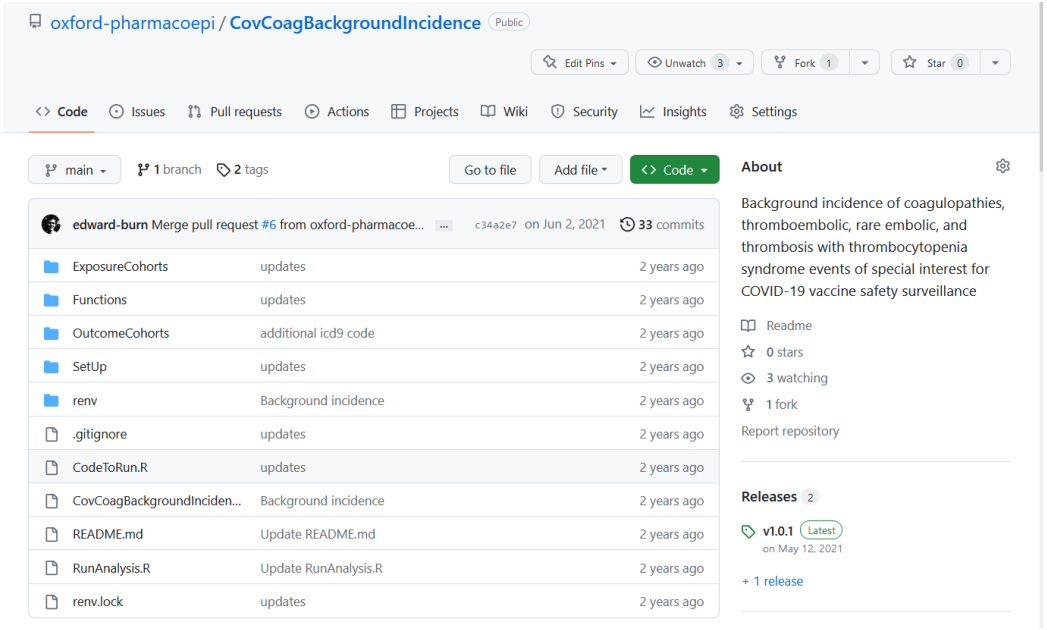


PDS Pharmacoepidemiology & Drug Safety
ispe Official Journal of the International Society for Pharmacoepidemiology

ORIGINAL ARTICLE | Full Access

Background rates of five thrombosis with thrombocytopenia syndromes of special interest for COVID-19 vaccine safety surveillance: Incidence between 2017 and 2019 and patient profiles from 38.6 million people in six European countries

Edward Burn, Xintong Li, Kristin Kostka, Henry Morgan Stewart, Christian Reich, Sarah Seager, Talita Duarte-Salles, Sergio Fernandez-Bertolin, María Aragón, Carlen Reyes ... See all authors



main CovCoagBackgroundIncidence / RunAnalysis.R

Code Blame 1378 lines (1206 loc) · 53.2 KB Raw

```
}
}
}

IR.summary<-bind_rows(IR.summary, .id = NULL)
IR.summary$db<-db.name

# save ----
save(IR.summary, file = paste0(output.folder, "/IR.summary_", db.name, ".RData"))
save(Patient.characteristcis, file = paste0(output.folder, "/Patient.characteristcis_", db.name, ".RData"))
save(Patient.characteristcis.for.plotting, file = paste0(output.folder, "/Patient.characteristcis.for.plotting_", db.name, ".RData"))

# # zip results
print("Zipping results to output folder")
unlink(paste0(output.folder, "/OutputToShare_", db.name, ".zip"))
zipName <- paste0(output.folder, "/OutputToShare_", db.name, ".zip")

files<-c(paste0(output.folder, "/IR.summary_", db.name, ".RData"),
         paste0(output.folder, "/Patient.characteristcis_", db.name, ".RData"),
         paste0(output.folder, "/Patient.characteristcis.for.plotting_", db.name, ".RData"))
files <- files[file.exists(files)==TRUE]

createZipFile(zipFile = zipName,
              rootFolder=output.folder,
              files = files)

print("Done!")
print("-- If all has worked, there should now be a zip folder with your results in the output folder to share")
print("-- Thank you for running the study!")
```

```
3     # link to db tables -----
```

CDMConnector 0.6.0 Reference Articles Changelog

CDMConnector

Are you using the tidyverse with an OMOP Common Data Model?

Interact with your CDM in a pipe-friendly way with CDMConnector.

- Quickly connect to your CDM and start exploring.
- Build data analysis pipelines using familiar dplyr verbs.
- Easily extract subsets of CDM data from a database.

```
20                    cdm_database_schema,
21                    ".concept_ancestor")))
```

PatientProfiles 0.1.0 Reference Articles

Function reference

Add individual patient characteristics

Add patient characteristics to a table in the OMOP Common Data Model

`addAge()`
Compute the age of the individuals at a certain date

`addFutureObservation()`
Compute the number of days till the end of the observation period at a certain date

`addInObservation()`
Indicate if a certain record is within the observation period

`addPriorHistory()`
Compute the number of days of prior history in the current observation period at a certain date

`addSex()`
Compute the sex of the individuals

54",

IncidencePrevalence 0.3.0 Reference Articles

Collect population incidence estimates

Collect population incidence estimates

Usage

```
estimateIncidence(
  cdm,
  denominatorTable,
  outcomeTable,
  denominatorCohortId = NULL,
  outcomeCohortId = NULL,
  interval = "years",
  completeDatabaseIntervals = TRUE,
  outcomeWashout = Inf,
  repeatedEvents = FALSE,
  minCellCount = 5,
  temporary = TRUE,
  returnParticipants = FALSE
)
```

```
804     # see end.date defined above
```

main C1-001-RareBloodCancersPrevalence / RunStudy.R

Code Blame 120 lines (117 loc) · 3.84 KB

```
81      outcomeCohortName = outcome_cohorts$cohortName,
82      interval = c("years"),
83      verbose = TRUE
84    )
85
86    info(logger, paste0("- getting period prevalence for ", denominators[[i]]))
87    prevalence_estimates[[paste0(
88      "period_prevalence_",
89      denominators[[i]]
90    )]] <- estimatePeriodPrevalence(
91      cdm = cdm,
92      tablePrefix = table_period_prev,
93      denominatorTable = denominators[[i]],
94      outcomeTable = table_outcome,
95      outcomeCohortId = outcome_cohorts$cohortId,
96      outcomeCohortName = outcome_cohorts$cohortName,
97      completeDatabaseIntervals = TRUE,
98      fullContribution = c(TRUE, FALSE),
99      interval = c("years"),
100     verbose = TRUE
101   )
102 }
103 # gather results and export -----
104 info(logger, "GATHERING RESULTS")
105 study_resuls <- gatherIncidencePrevalenceResults(cdm,
106                                                  resultList = prevalence_estimates,
107                                                  databaseName = db_name
108 )
109 info(logger, "ZIPPING RESULTS")
110 exportIncidencePrevalenceResults(
111   result = study_resuls,
112   zipName = paste0(c(db_name,
113                      "C1_001_Results",
114                      format(Sys.Date(), format="%Y%m%d")),
115                    collapse = "_"),
116   outputFolder = output_folder
117 )
118
119 print("-- Thank you for running the study!")
120 print("-- If all has worked, there should now be a zip folder with your results in the output folder to share")
```

IncidencePrevalence

- Population-level DUS analyses
- Population-level descriptive epidemiology
- Time series analyses and Difference-in-difference studies
- RMM effectiveness

IncidencePrevalence 0.4.0 Reference Articles

Estimate period prevalence

Estimate period prevalence

Usage

```
estimatePeriodPrevalence(
  cdm,
  denominatorTable,
  outcomeTable,
  denominatorCohortId = NULL,
  outcomeCohortId = NULL,
  outcomeLookbackDays = 0,
  interval = "years",
  completeDatabaseIntervals = TRUE,
  fullContribution = FALSE,
  minCellCount = 5,
  temporary = TRUE,
  returnParticipants = FALSE
)
```

IncidencePrevalence 0.4.0 Reference Articles

Collect population incidence estimates

Collect population incidence estimates

Usage

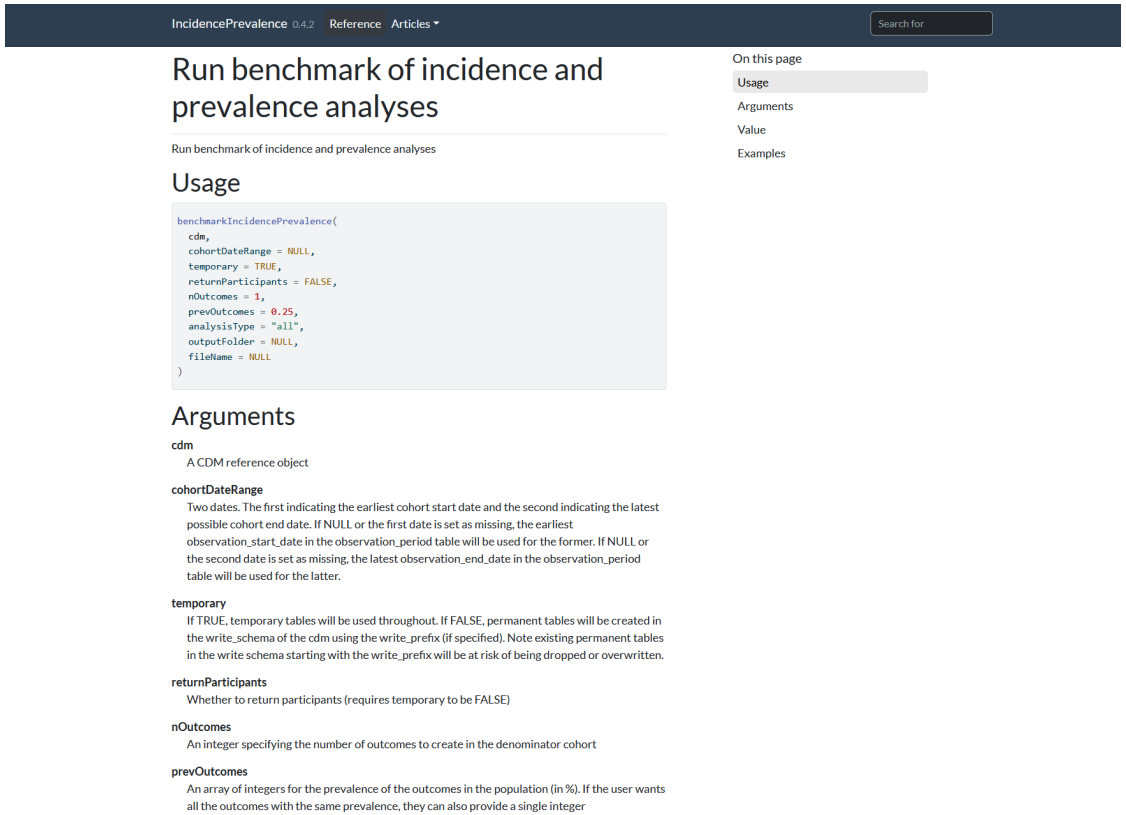
```
estimateIncidence(
  cdm,
  denominatorTable,
  outcomeTable,
  denominatorCohortId = NULL,
  outcomeCohortId = NULL,
  interval = "years",
  completeDatabaseIntervals = TRUE,
  outcomeWashout = Inf,
  repeatedEvents = FALSE,
  minCellCount = 5,
  temporary = TRUE,
  returnParticipants = FALSE
)
```

Software validation: unit testing

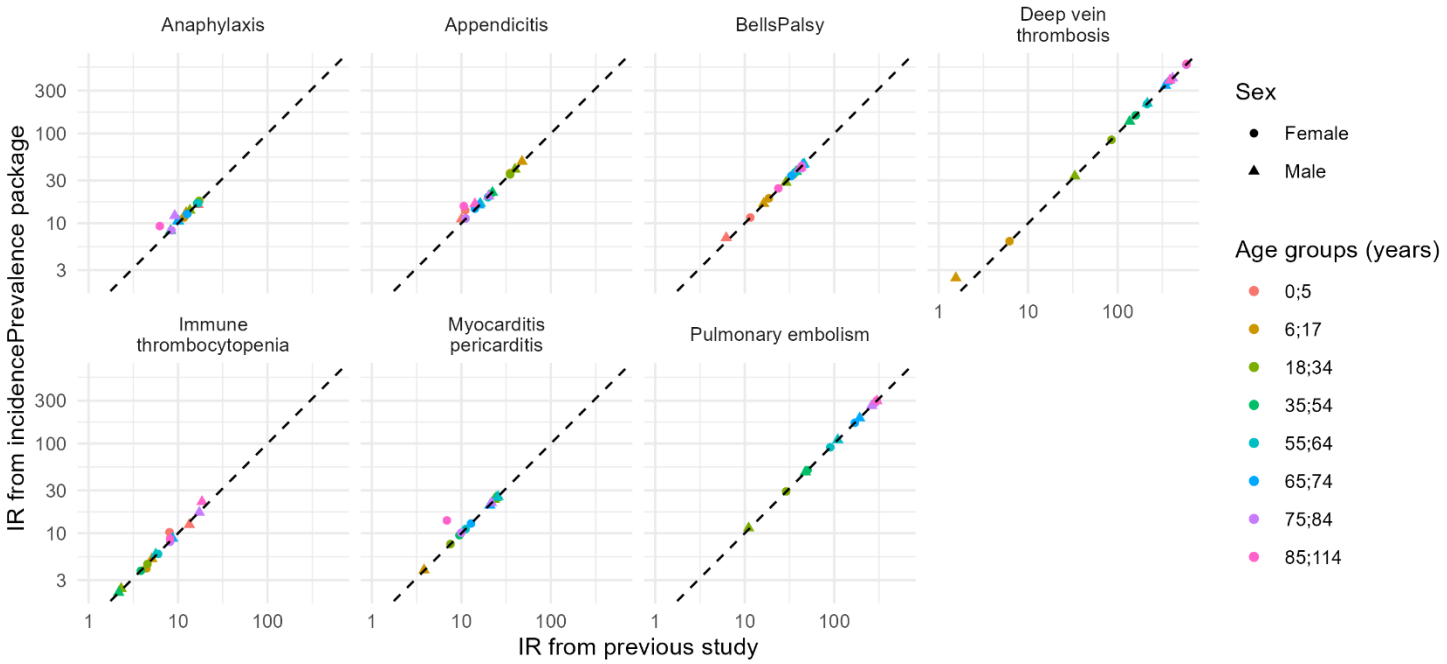
IncidencePrevalence coverage - 98.10%

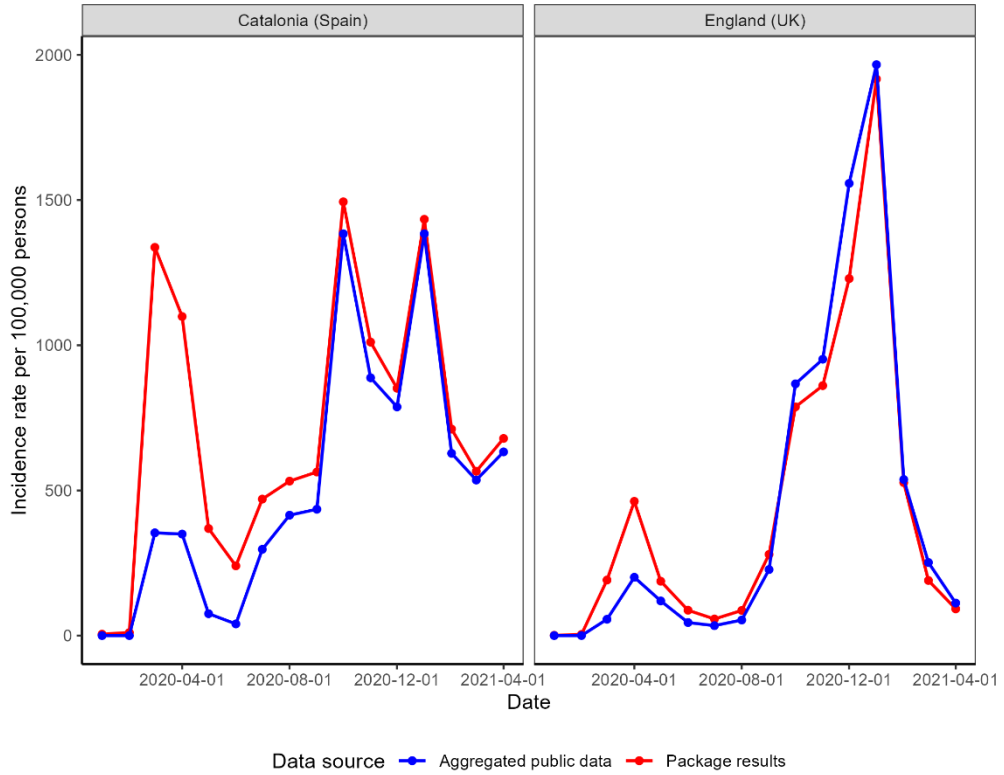
Files R/utils.R

File	Lines	Relevant	Covered	Missed	Hits / Line	Coverage
R/utils.R	207	47	41	6	10	87.23%
R/plotting.R	248	108	100	8	5	92.59%
R/getDenominatorCohorts.R	501	304	288	16	111	94.74%
R/exportIncidencePrevalenceResults.R	109	41	39	2	3	95.12%
R/getIncidence.R	374	229	221	8	235	96.51%
R/inputValidation.R	386	230	222	8	102	96.52%
R/generateDenominatorCohortSet.R	568	313	313	0	160	100.00%
R/estimateIncidence.R	538	294	294	0	82	100.00%
R/mockIncidencePrevalenceRef.R	492	279	279	0	76	100.00%
R/estimatePrevalence.R	556	259	259	0	63	100.00%
R/getPrevalence.R	363	222	222	0	498	100.00%
R/benchmarkIncidencePrevalence.R	321	185	185	0	4	100.00%
R/getStudyDays.R	141	84	84	0	207	100.00%
R/obscureCounts.R	73	39	39	0	110	100.00%
R/bindEstimates.R	135	38	38	0	7	100.00%
R/recordAttrition.R	65	33	33	0	2671	100.00%



Software validation: face validity

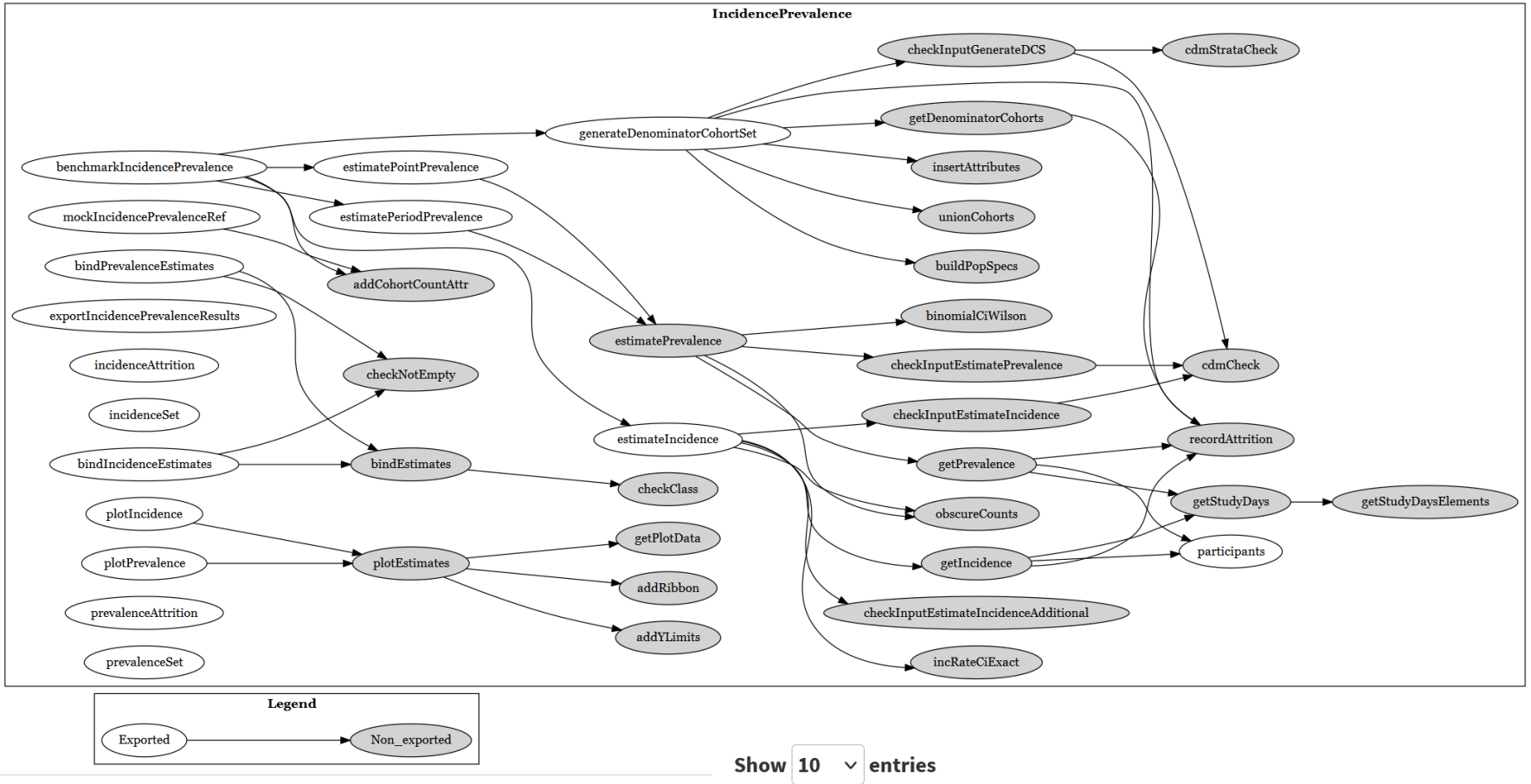




Software performance

Task	CPRD AURUM (n=39,999,011)	CPRD GOLD (n=15,662,217)	SIDIAP (n=8,265,343)	IPCI (n=2,612,850)
Generating denominator (8 cohorts)	19 mins	8 mins	3 mins	1 min
Yearly period prevalence	11 mins	5 mins	5 mins	1 min
Monthly period prevalence	17 mins	6 mins	8 mins	2 mins
Yearly incidence	8 mins	3 mins	4 mins	1 min
Monthly incidence	13 mins	5 mins	7 mins	1 min

Software review



Show 10 entries

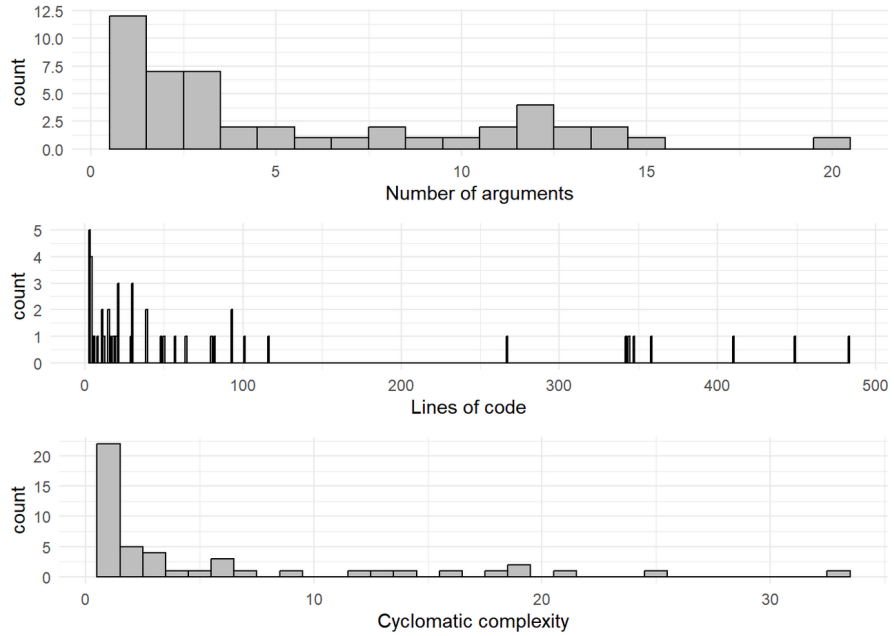
Search:

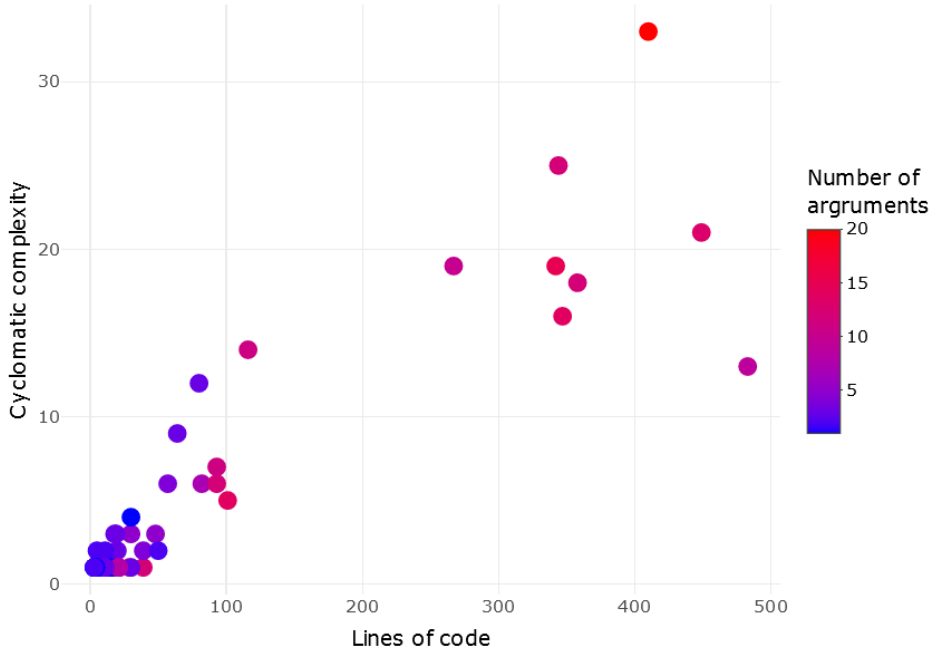
	R	cpp	o	h	java	sql
# lines of code	5099	0	0	0	0	0

Software review

Summary of package functions

	min	median	max
Number of arguments	1	3	20
Lines of code	3	25	483
Cyclomatic complexity	1	2	33





Software dissemination

IncidencePrevalence: Estimate Incidence and Prevalence using the OMOP Common Data Model

Calculate incidence and prevalence using data mapped to the Observational Medical Outcomes Partnership (OMOP) common data model. Incidence and prevalence can be estimated for the total population in a database or for a stratification cohort.

Version:	0.4.1
Depends:	R (≥ 4.0)
Imports:	CDMConnector (≥ 1.0.0), checkmate (≥ 2.0.0), cli (≥ 3.0.0), DBI (≥ 1.0.0), dbplyr (≥ 2.0.0), dplyr (≥ 1.1.0), glue (≥ 1.5.0), ggplot2 (≥ 3.4.0), scales (≥ 1.1.0), lubridate (≥ 1.0.0), magrittr (≥ 2.0.0), purrr (≥ 0.3.5), rlang (≥ 1.0.0), stringr (≥ 1.5.0), tidyr (≥ 1.2.0), tidyselect (≥ 1.2.0), zip (≥ 2.2.0)
Suggests:	knitr, rmarkdown, RPostgres, tibble, duckdb, odbc, here, Hmisc, epitools, tictoc, testthat (≥ 0.3.1), spelling, PaRe
Published:	2023-07-11
Author:	Edward Burn [aut, cre], Berta Raventos [aut], Marti Catala [aut], Mike Du [ctb], Yuchen Guo [ctb], Adam Black [ctb], Ger Inberg [ctb], Kim Lopez [ctb]
Maintainer:	Edward Burn <edward.burn at ndorms.ox.ac.uk>
License:	Apache License (≥ 2)
URL:	https://darwin-eu.github.io/IncidencePrevalence/
NeedsCompilation:	no
Language:	en-US
Materials:	README
CRAN checks:	IncidencePrevalence results

Documentation:

Reference manual:	IncidencePrevalence.pdf
Vignettes:	a01_Introduction_to_IncidencePrevalence a02_Creating_denominator_cohorts a03_Creating_outcome_cohorts a05_Calculating_prevalence a06_Calculating_incidence

Downloads:

Package source:	IncidencePrevalence_0.4.1.tar.gz
Windows binaries:	r-devel: IncidencePrevalence_0.4.1.zip, r-release: IncidencePrevalence_0.4.1.zip, r-oldrel: IncidencePrevalence_0.4.1.zip
macOS binaries:	r-release (arm64): IncidencePrevalence_0.4.1.tgz, r-oldrel (arm64): IncidencePrevalence_0.4.1.tgz, r-release (x86_64): IncidencePrevalence_0.4.1.tgz, r-oldrel (x86_64): IncidencePrevalence_0.4.1.tgz
Old sources:	IncidencePrevalence archive

IncidencePrevalence: An R package to calculate population-level incidence and prevalence rates using the OMOP Common Data Model.

Berta Raventós[1,2*], Martí Català[3*], Mike Du[3], Yuchen Guo[3], Adam Black[4], Ger Inberg[5], Xintong Li[3], Kim López-Güell[3], Danielle Newby[3], Maria de Ridder[5], Cesar Barboza[5], Talita Duarte-Salles[3], Egill Fridgeirsson[5], Sofia Verhamme[5], Peter Rijnbeek[5], Daniel Prieto Alhambra[3,5], Edward Burn[3]

Affiliations

1. Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain
2. Universitat Autònoma de Barcelona, Can Ruti Campus (Cerdanyola del Vallès), Barcelona, Spain
3. Centre for Statistics in Medicine (CSM), Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, UK
4. Odysseus Data Services, Cambridge, MA, USA.
5. Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

*joint first-authors

Corresponding author: Edward Burn, Botnar Research Centre, Windmill Road, OX37LD, Oxford, UK, edward.burn@ndorms.ox.ac.uk

Summary

Software development in DARWIN EU® to support high throughput network studies

- Designed to deliver catalogue of standard data analyses
- Testing to ensure validity and usability
- Flexibility in design to account for expected updates in the catalogue over time
- Extensibility to allow complex analyses to build on top of off-the-shelf tools
- R packages published on CRAN with thorough documentation

DARWIN EU

Coordination Centre

Study operations

Katia Verhamme, Department of Medical Informatics, EMC Rotterdam, The Netherlands

What analyses and studies will DARWIN EU® deliver?

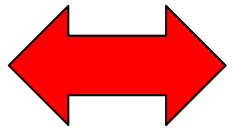
Category of observational analyses and studies	Description
Routine repeated analyses	**Routine analyses** based on a **generic study protocol** • Periodical estimation of drug utilisation • Safety monitoring of a medicinal product • Estimation of the incidence of a series of adverse events
Off-the-shelf studies	Studies for which a **generic protocol** is adapted to a research question • Estimate the prevalence, incidence or characteristics of exposures • Health outcomes • Describe population characteristics
Complex Studies	Studies **requiring development or customisation** of specific study designs, protocols and Statistical Analysis Plans (SAPs), with extensive collection or extraction of data • Etiological study measuring the strength and determinants of an association between an exposure and the occurrence of a health outcome considering sources of bias, potential confounding factors and effect modifiers
Very Complex Studies	Studies which **cannot rely only on electronic health care databases**, or which would require **complex methodological work** • Studies where it may be necessary to combine a diagnosis code with other data such as results of laboratory investigations, or studies requiring additional data collection

Expected number of studies

PHASE I Establishment – 1st year → PHASE II Establishment – 2nd year → PHASE III Operation – 1st year → Operation 2nd year → Operation 3rd year

	Year 1	Year 2	Year 3	Year 4	Year 5
Phases	Phase I	Phase II	Phase III	Option 1	Option 2
Routine Repeated analysis	At least 1 study	-	30	60	60
Off the shelf studies	At least 2 studies	6 + 8	30	60	60
Complex Studies	1	4	12	24	24
Very Complex Studies	0	0	0	1	1

1. Study Exploration	<u>Study Feasibility- Study request by EMA:</u> - Do we have the data? – Darwin Portal - Do we have the analytical pipelines? (OTS)
2. Study Initiation	- Work Order Form Data Partners - Creation of Study Team: PI/data analyst - Declaration of Interest
3. Study Implementation	- Study outline/Protocol – Upload to EUPAS register - IRB approval - Kick-off meeting - Phenotyping – Cohort Diagnostics - Study Package – Test Run
4. Study Execution	- Data Partners run Study Package - Data QC by Data Partners - Results uploaded to DRE - Results reviewed by PI
5. Study Dissemination	- Generation of Study Report (ENCePP template) - Upload to EUPAS register - Manuscript generation - Study archiving



EMA

Database Partners

Darwin CC:
- Network Pillar
- Development Pillar
- Technology Pillar
- Management Pillar

STUDIES YEAR 1

Study Title	Committees	Study Type	Type of analysis	Data bases	Status
DARWIN EU® - **Prevalence of rare blood cancers in Europe (ICPE P1180)**	COMP	OTS	Disease Epidemiology	IPCI (Nl) SIDIAP (Spain) CPRD Gold (UK) IQVIA LPD (Be) IQVIA DA (Ge)	Completed
DARWIN EU® - **Drug utilisation of valproate**-containing medicinal products in women of childbearing potential (**ICPE P42**)	Following safety referral	OTS	Drug Utilisation Study	IPCI (Nl) SIDIAP (Spain) CPRD Gold (UK) IQVIA LPD (Be) IQVIA DA (Ge) HDSF (Fi)	Completed
DARWIN EU® - **DUS of Antibiotics** in the 'Watch' category of the WHO AWaRe classification of antibiotics for evaluation and monitoring of use (**ICPE P331**)	PRAC/CHMP	OTS	Drug Utilisation Study	IPCI (Nl) CHUBX (France) SIDIAP (Spain) IMASIS (Spain) IQVIA DA (Ge) CPRD Gold (UK)	Completed
DARWIN EU® - **Background rates** of serious adverse events to contextualise safety assessments in clinical trials and non-interventional studies in adolescent and adult patients with **severe asthma**.	CHMP	Complex (complex phenotype)	Disease Epidemiology	IPCI (Nl) SIDIAP (Spain) IMASIS (Spain) CPRD Gold (UK) Estonian Biobank	Ongoing

STUDIES YEAR 2 (1)

Study Title	Committees	Study Type	Type of analysis	Data bases	Status
DARWIN EU® - **Multiple myeloma**: patient characterisation, treatments and survival in the period 2012-2022	HTA/Payers	OTS	Disease Epidemiology and Treatment Pattern analysis	IQVIA DA (Ge) SIDIAP (Spain) IMASIS (Spain) Estonian Biobank ACI Varha (Fi) CHUBX (France) IKNL (Nl)	Ongoing
DARWIN EU® **Drug Utilisation Study** of **prescription opioids**.	PRAC	OTS	Drug Utilisation Study	Estonian Biobank IPCI (NL)	Ongoing
D **N** **C** **v** c					
DARWIN EU® - Co-prescribing of endothelin receptor antagonists (ERAs) and phosphodiesterate-5 inhibitors (PDE-5is) in **pulmonary arterial hypertension** (PAH)	CHMP	OTS	Disease Epidemiology and Treatment Pattern analysis	CHUBX (France) CPRD GOLD (UK) Estonian Biobank IQVIA DA (Ge)	Ongoing

Study Feasibility:

As of to date: **22 feasibility requests** in year 2:

- 10 studies received green light
- 3 studies suggested to put on hold ➔ different reasons: lack of data or need of more recent data
- 7 Feasibility assessments either just received or under review by EMA

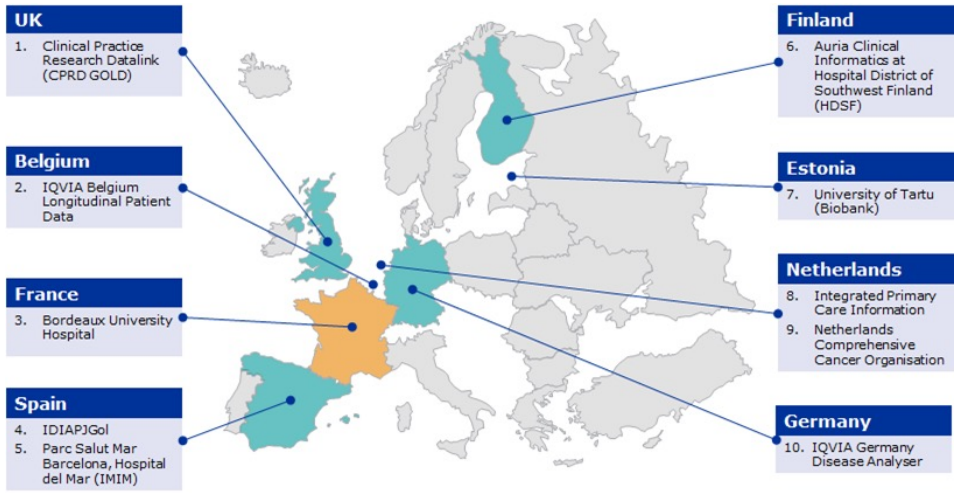
STUDIES YEAR 2 (2)

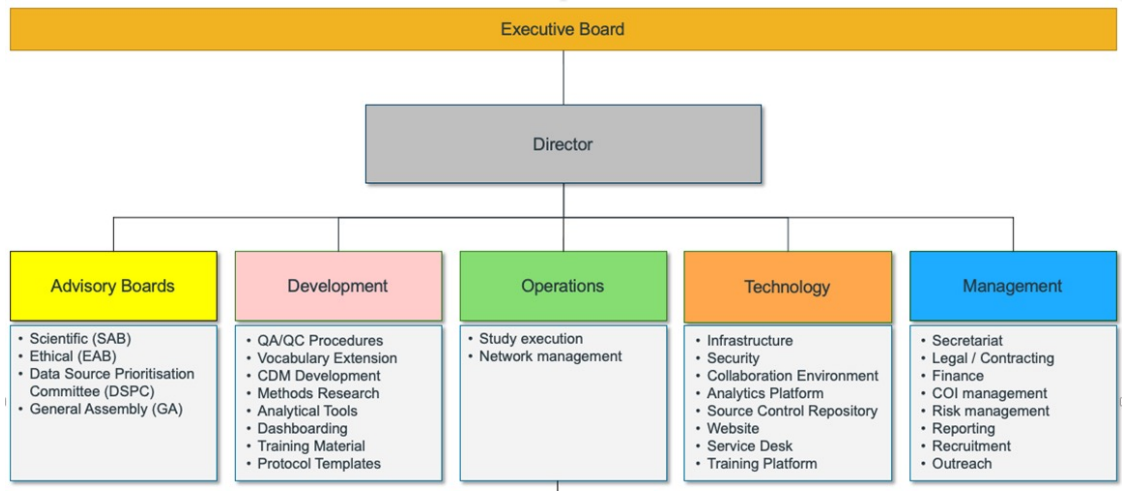
Study Title	Committees	Study Type	Type of analysis	Data bases	Status
DARWIN EU® - **Use of take-home naloxone** for opioid overdose treatment	CHMP	OTS	Drug Utilisation Study	IQVIA DA (Ge) IQVIA DA (Be) CPRD Gold (UK) SIDIAP	Ongoing
DARWIN EU® **DUS of medicines with prokinetic properties** in children and adults diagnosed with gastroparesis	NCA	OTS	Drug Utilisation Study	IPCI (Nl) CHUBX (France) SIDIAP (Spain) IMASIS (Spain) IQVIA DA (Ge) IQVIA LPD (Be) CPRD Gold (UK)	Ongoing
DARWIN EU® - **Effectiveness of COVID-19 vaccines** against severe COVID-19 and post-acute outcomes of SARS-CoV-2 infection	ECDC/ Vaccine monitoring platform	Complex	Comparative Effectiveness	CPRD Gold IPCI SIDIAP	Ongoing

Expected number of studies

PHASE I Establishment – 1st year → PHASE II Establishment – 2nd year → PHASE III Operation – 1st year → Operation 2nd year → Operation 3rd year

	Year 1	Year 2	Year 3	Year 4	Year 5
Phases	Phase I	Phase II	Phase III	Option 1	Option 2
Routine Repeated analysis	At least 1 study	-	30	60	60
Off the shelf studies	At least 2 studies	6 + 8	30	60	60
Complex Studies	1	4	12	24	24
Very Complex Studies	0	0	0	1	1





EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

EMA Colleagues:
RWE group

More Information

Data Analysis and Real World Interrogation Network (DARWIN EU) | European Medicines Agency (europa.eu)

www.darwin-eu.org

For questions to the Coordination Centre, please contact: enquiries@darwin-eu.org

For regular updates on DARWIN EU® Subscribe to the Big Data Highlights newsletter by sending an email to: bigdata@ema.europa.eu

01 Issue 1 February 2022

HMA Heads of Medicines Agencies

EMA EUROPEAN MEDICINES AGENCY

BIG DATA

HIGHLIGHTS

Quarterly update on implementation activities of the HMA-EMA Big Data Steering Group workplan

An agency of the European Union

Editorial

Peter Arlett

Co-chair of Big Data Steering Group/ Head of Data Analytics and Methods Task Force, EMA

Jesper Kjær

Co-chair of Big Data Steering Group/ Director of Data Analytics Centre, DKMA

INSIDE THIS ISSUE

Welcome to the first edition of the newsletter on Big Data which reports on the implementation of the HMA-EMA Big Data Steering Group workplan 2021-2023 and the data and digital pillar of the Network Strategy 2025.

The BDSG was established in May 2020 with the mandate to take forward and advise on implementation of the priority recommendations set out in the Big Data Task Force final report (phase two). The vision, set out by the Big Data Taskforce that guides delivery of the workplan is that:

> *by delivering a regulatory system able to integrate Big Data into its assessment and decision making, we can support the development of innovative treatments more quickly and optimise the safe and effective use of medicines.*
>
> **Big Data Task Force Final Report, Dec 2019**

Already in 2021, we have seen an impressive array of activities and outputs that support transformation to data-driven medicines regulation. These include the finalisation of the network's data standardisation strategy and a series of technical workshops covering standardisation, real-world data, artificial intelligence, veterinary data and real-world evidence. More details on the 2021 deliverables are included later in this newsletter.

Moving forward, we are excited to be able to share with you more regular updates on big data and the implementation of the BDSG workplan via this newsletter. We look forward to collaborating and leveraging the work of stakeholders to increase the utility of big data in medicines regulation.

The Big Data Steering Group welcomes your feedback and questions you may have by email at bigdata@ema.europa.eu .

We hope you find the Big Data Highlights informative!