EHDEN

EUROPEAN HEALTH DATA & EVIDENCE NETWORK

Scaling up the big data ecosystem across Europe

Prof Dani Prieto-Alhambra

Prof of Pharmaco- and Device Epidemiology
CSM, NDORMS, University of Oxford

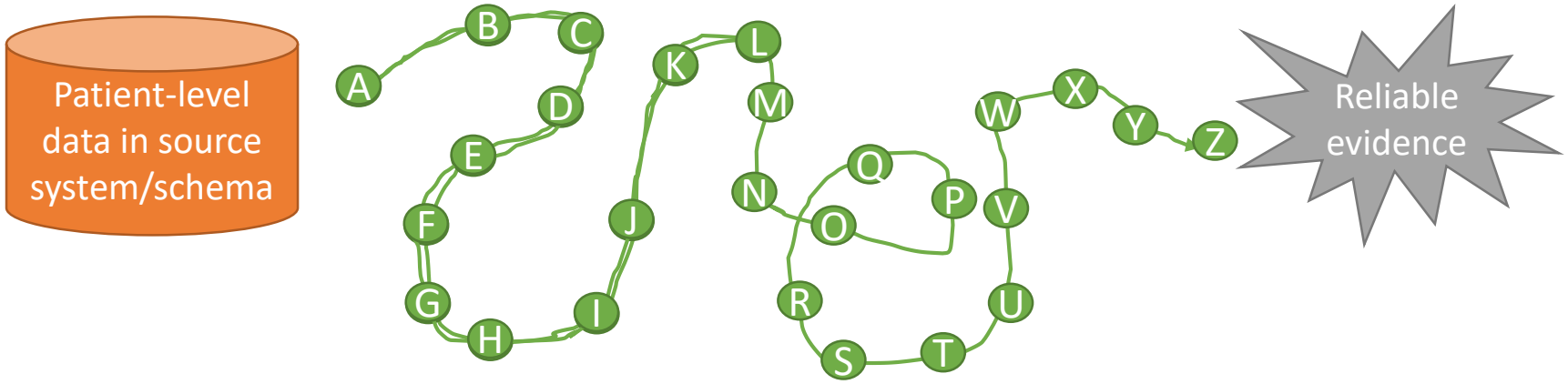
imi efpia

AGENDA

- EHDEN – An Introduction
- The Oxford Study-A-Thon
- The Barcelona Study-A-Thon (if time allows…)

ALL TOO OFTEN REAL WORLD RESEARCH IS A CHALLENGING JOURNEY….

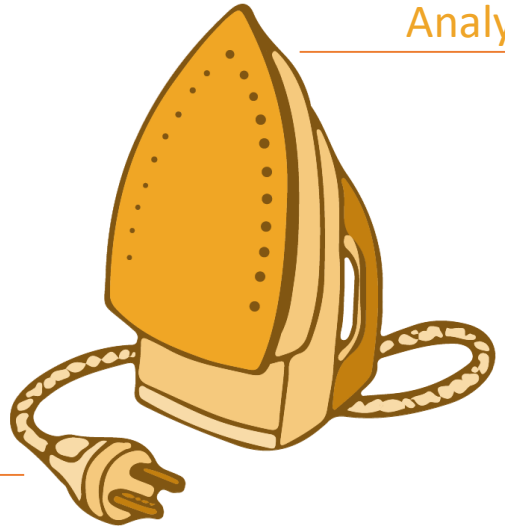
Generating Reproducible and Reliable Evidence



We need fully transparent and reproducible pipelines that enable large-scale federated analyses across Europe.

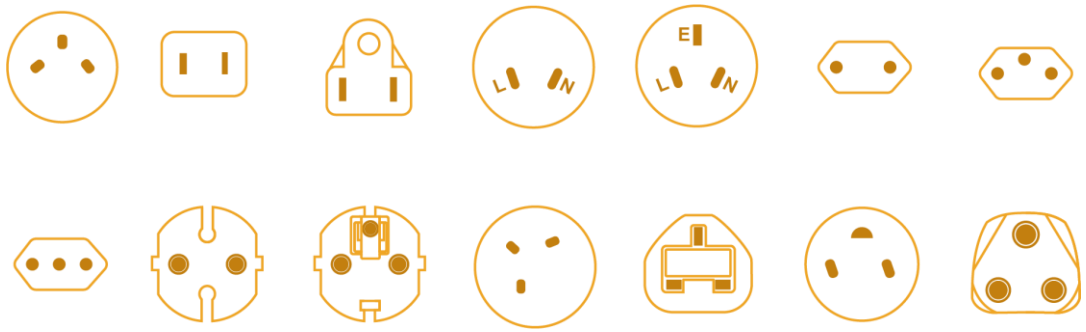
WHY IS THIS NOT CURRENT PRACTICE?

Analytical method



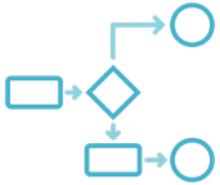
Link to data

The data…



What will it require?

Data interoperability

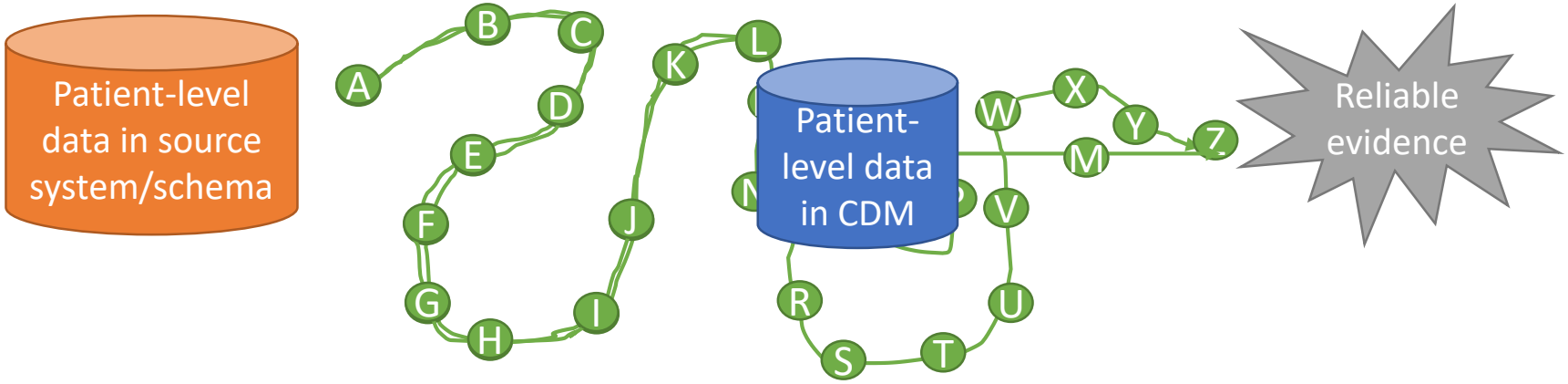


Standardised analytics

Data network

Strong community

GENERATING RELIABLE EVIDENCE USING THE OMOP CDM

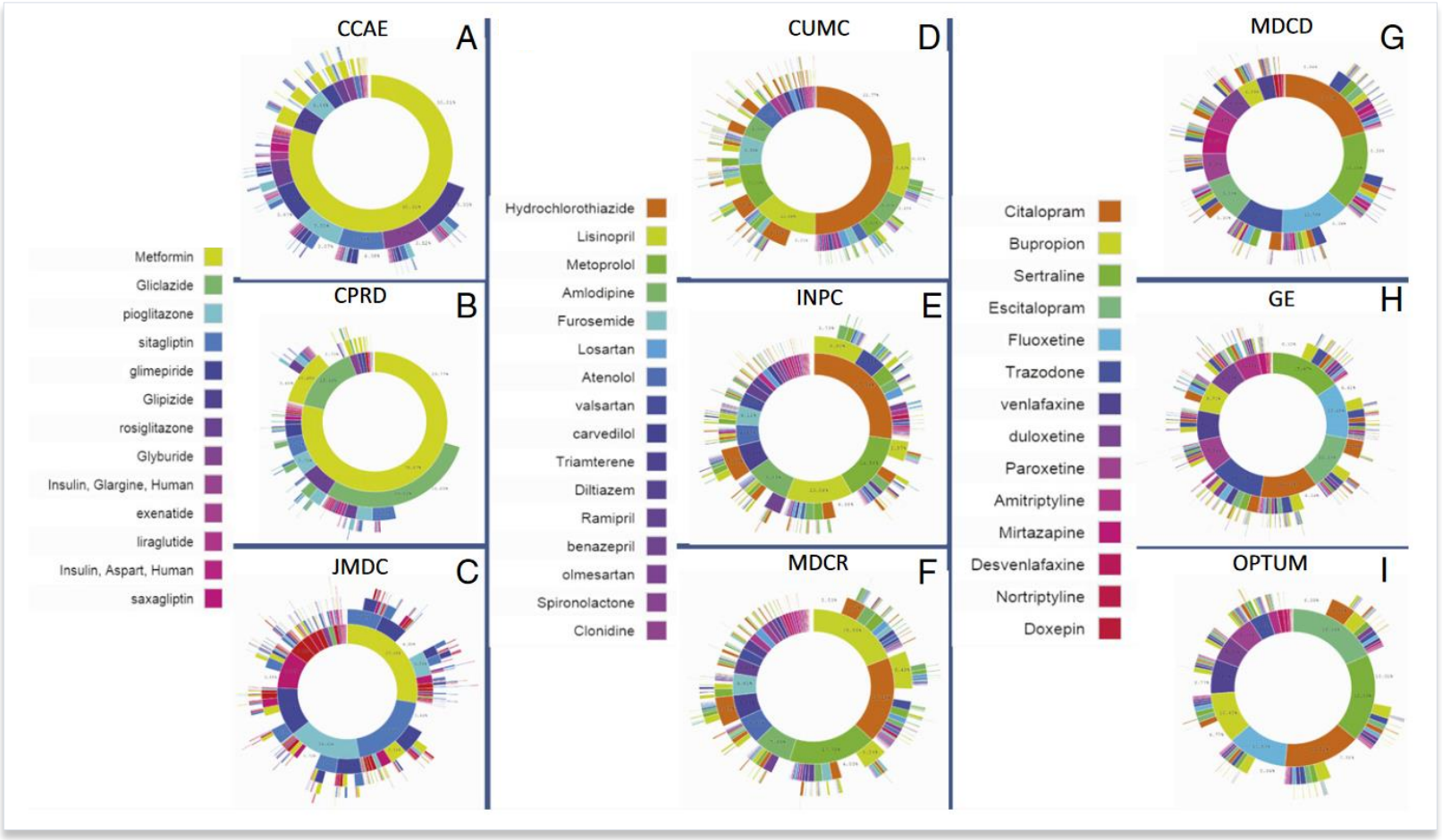


EHDEN will build on expertise and tools from prior IMI projects, such as EMIF, and will collaborate intensively with the global OHDSI community.

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

LARGE-SCALE OBSERVATIONAL RESEARCH IS FEASIBLE

T2 Diabetes Mellitus | Hypertension | Depression



11 Data sources

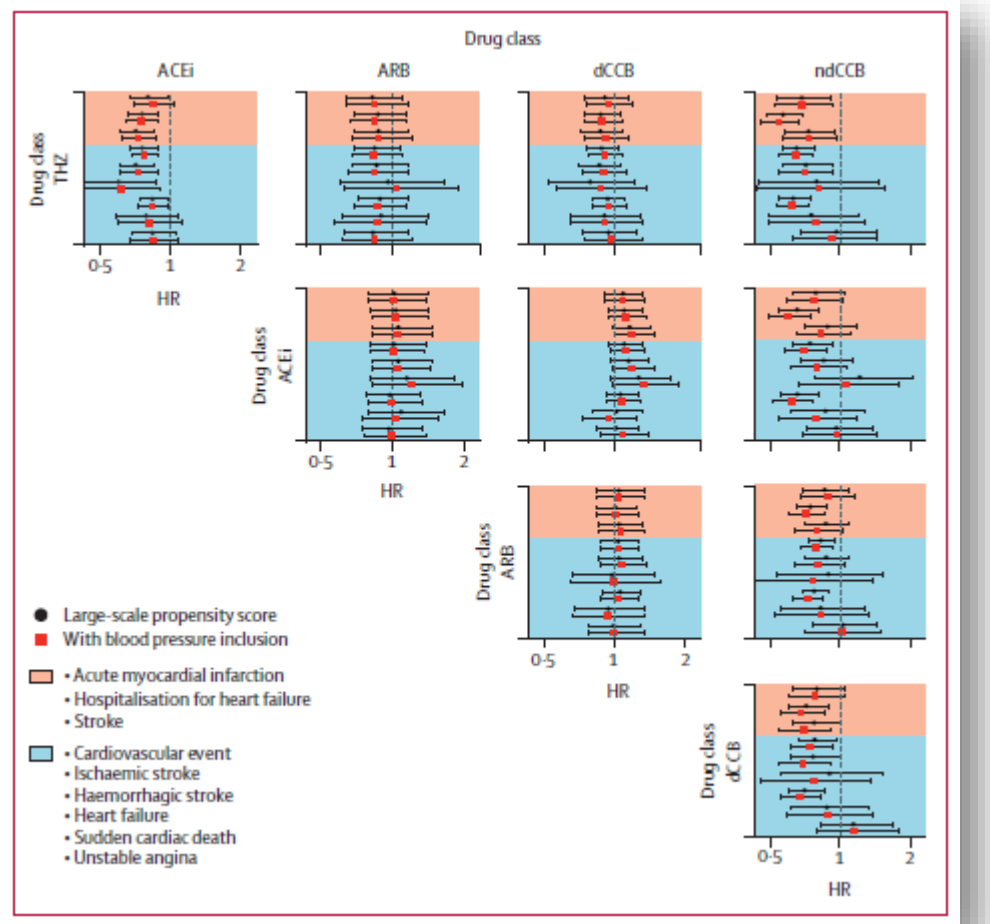
4 Countries (incl Japan)

> 250 million patients

"Characterizing treatment pathways at scale using the OHDSI network."
George Hripcsak et al. - PNAS (2016)27:7329–7336

imi efpia

LARGE-SCALE OBSERVATIONAL RESEARCH TO INFORM GUIDELINES



LEGEND-HTN

- Systematic, large-scale study
- New-user cohort
- 3 primary outcomes
- 6 secondary
- 46 safety outcomes

9 Data sources
6 Claims
3 EHR

4 Countries

4.9 million patients

Equivalent to 22,000 observational studies

" Comprehensive comparative effectiveness and safety of 1st line antihypertensive drug classes."
Suchard MA, et al. – Lancet (2019): online 24 Oct

EHDEN: Vision and Mission

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

Our 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

EHDEN IS ABOUT ...

FEDERATION

Creation of an EU-wide architecture for federated analyses of real world data

HARMONISATION

Harmonise more than 100 million anonymised health records to the OMOP common data model

COMMUNITY

Establish a self-sustaining open science collaboration in Europe, supporting academia, industry, regulators, payers, government, NGOs and others

OUTCOMES

Enabling outcomes driven healthcare at a European level

EDUCATION

The establishment of an EHDEN Academy, webinars and face-to-face training sessions to train all stakeholders

FOUNDATIONAL PILLARS

Infrastructure

Creation of an **EU-wide** federated network architecture

Privacy by design

Data harmonisation to the OMOP common data model

Training & certification of **SMEs**

Research & Outcomes

Use cases to evaluate the EHDEN federated network

Collaboration on consistent **methodologies**

Collaboration with the global **OHDSI** research network

Incorporation of the **ICHOM** health outcome standards

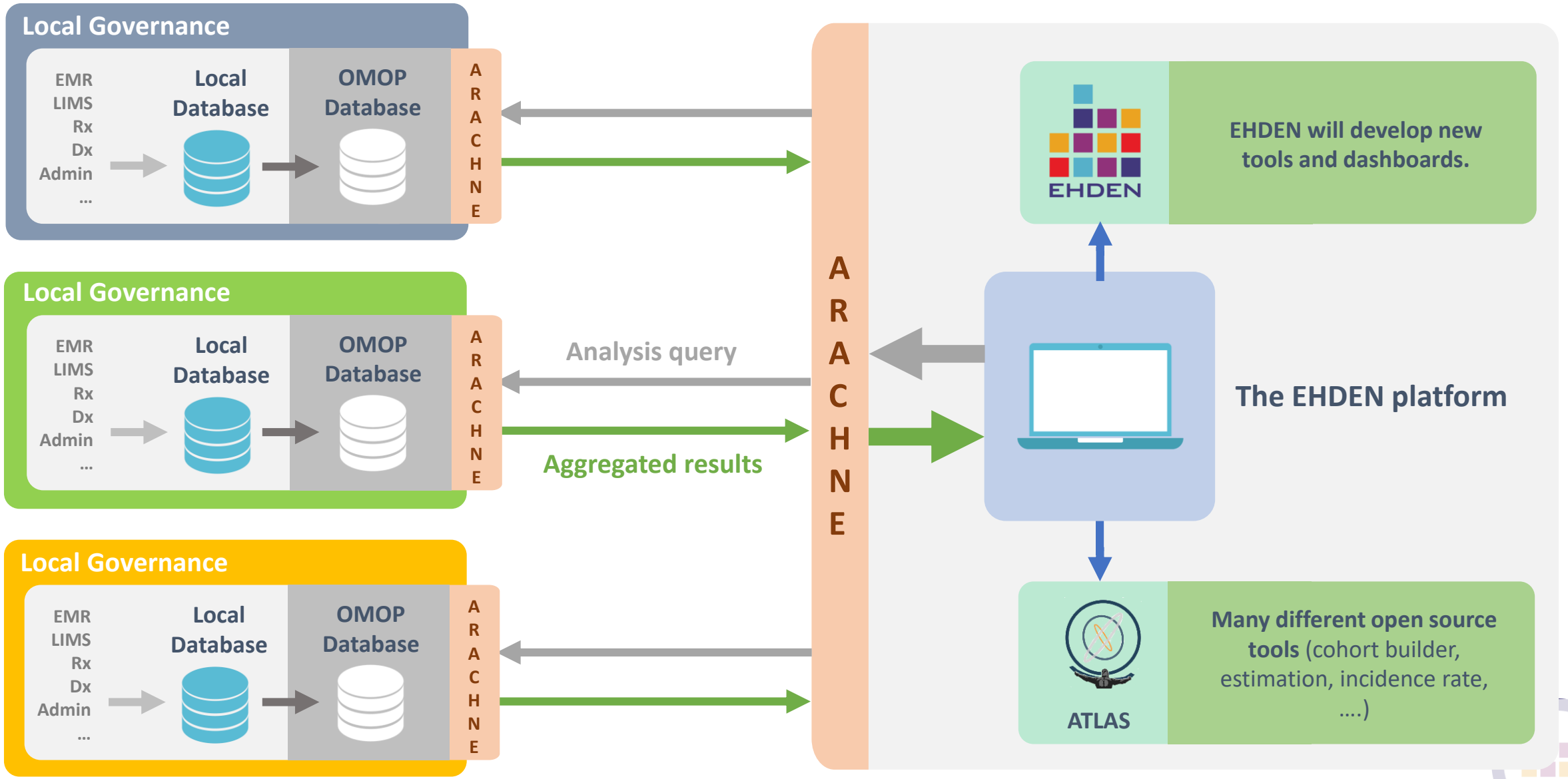
Education & Community

Establishment of an **EHDEN Academy**

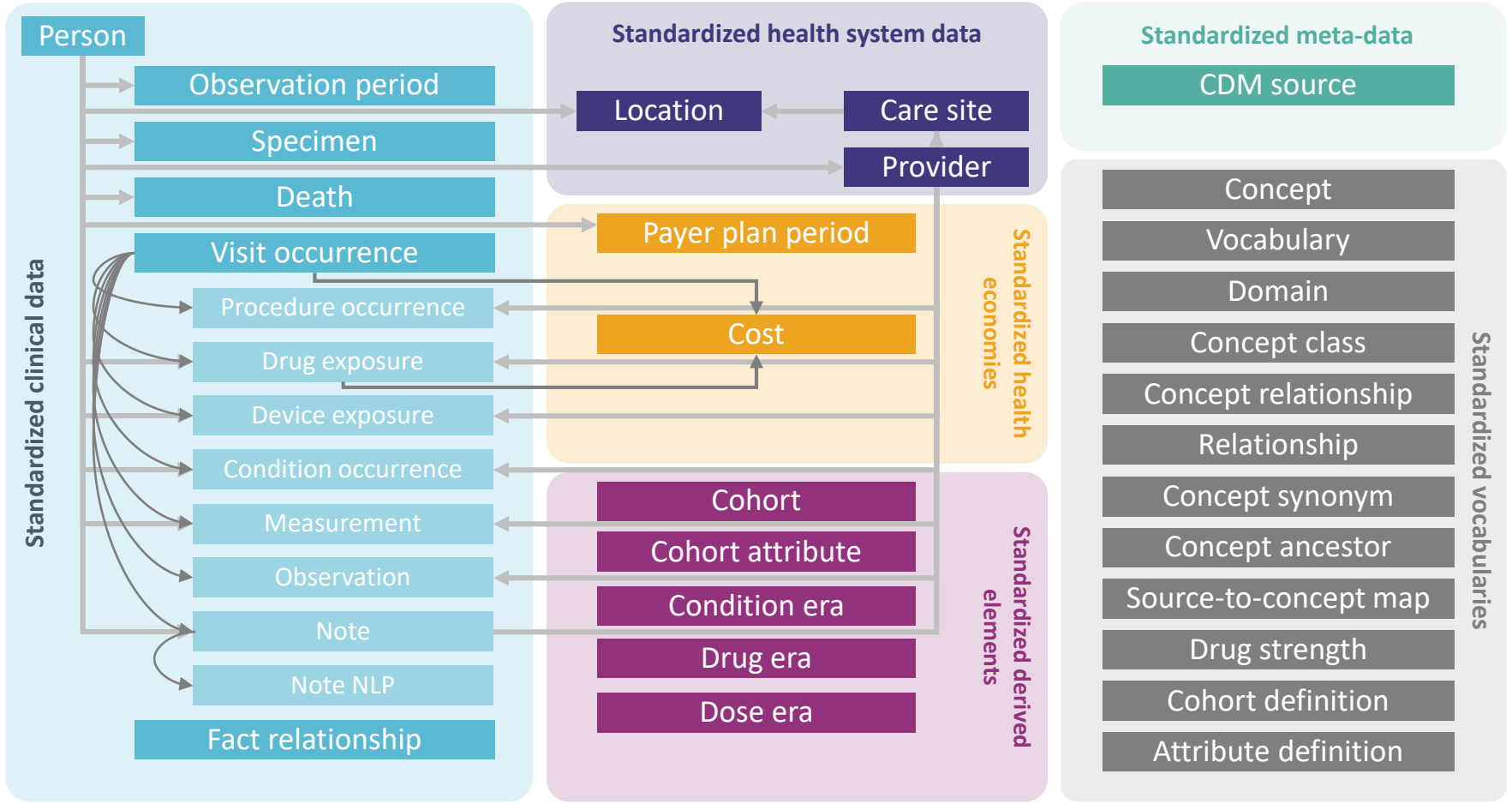
Expansion of the **OHDSI network** in Europe

Collaboration on **collective memory** for research use cases

The EHDEN Federated Data Network



THE OMOP COMMON DATA MODEL



Patient-centric

Tabular

Extendable

Built for analytics

Relational design

v 5.0.1

CALL PROCESS OVERVIEW

Harmonisation fund

Data sources

Open calls

Tailored for project objectives and sustainability

Grant awarding

Evaluated via a pre-defined set of criteria by the Data source prioritisation committee

Data sources can choose the SME from the pool of EHDEN certified SMEs

SMEs are paid via grants from the harmonisation fund

Payments are milestone based

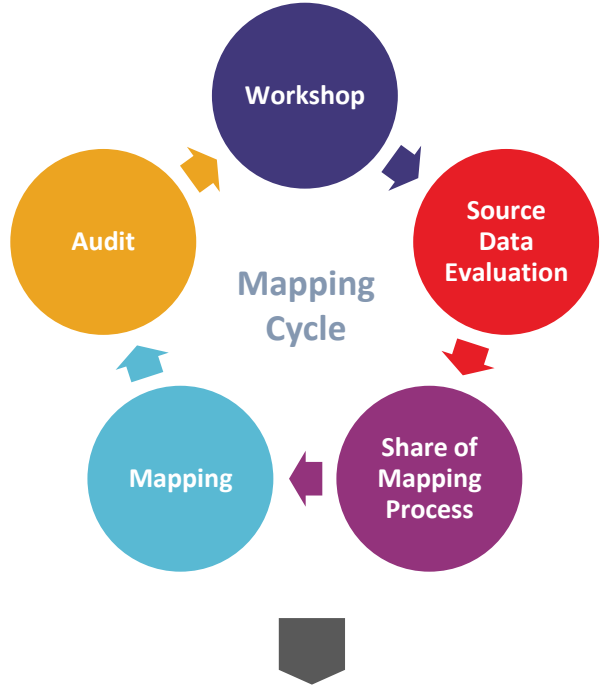
Supporting SMEs

Open calls

Focusing on SMEs able to support mapping and sustainability

Training & Certification

SME certification committee prioritizes SMEs for training and certification



Mapped data sources are encouraged to be active members of the **EHDEN community**, participating in **research studies**.

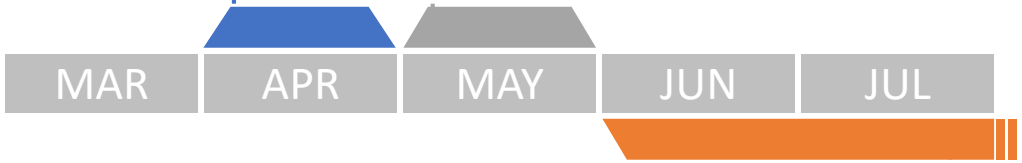
SME Pilot Call

Open Call

Submission of applications for the open call for SMEs from the 1st of April until the 1st of May (17h00) via the EHDEN website.

Evaluation

Following an eligibility check, applications were evaluated by the SME certification committee.



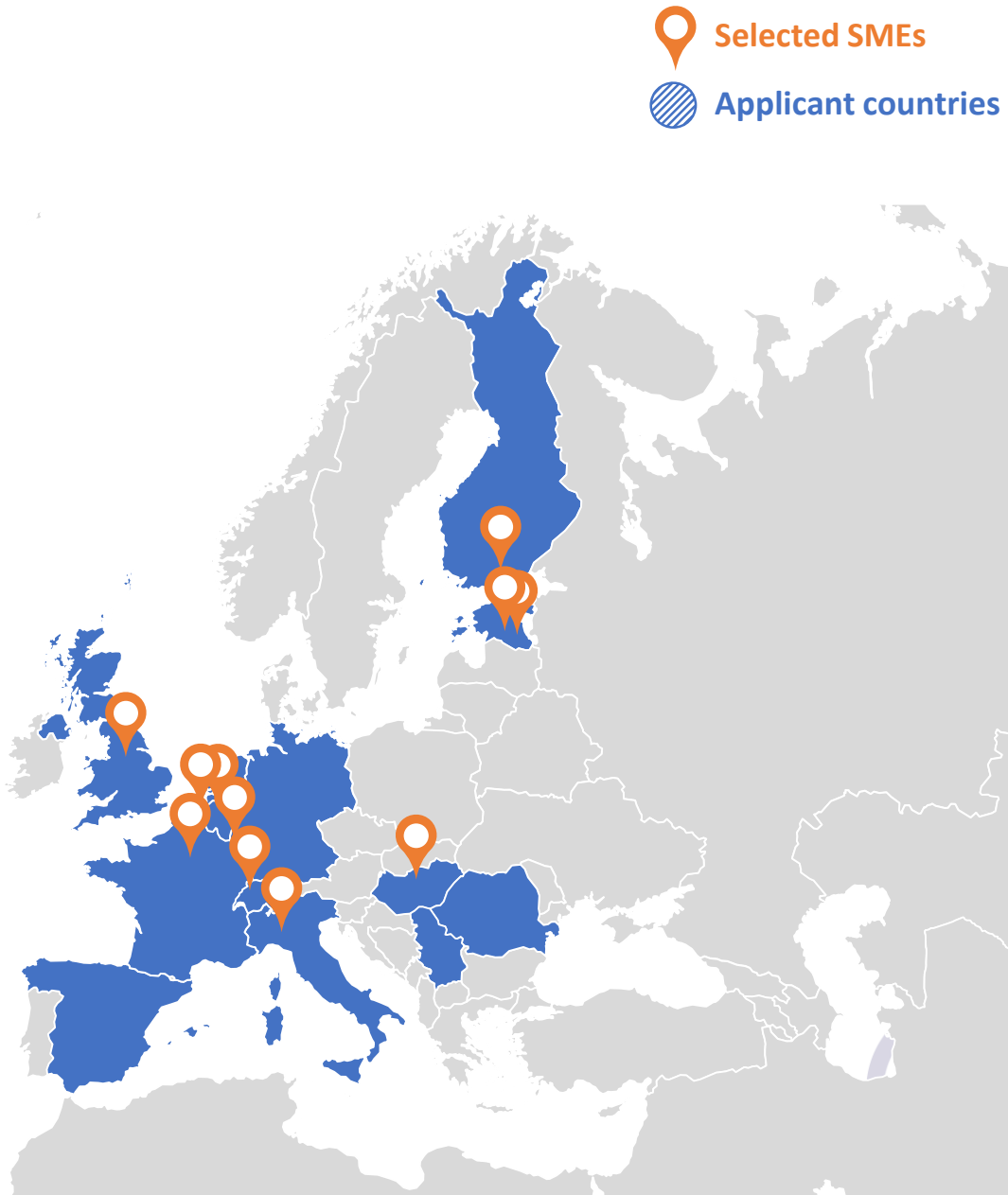
Training & Certification

Certification and training of selected SMEs in all necessary competencies.

34 SME profiles made

28 Eligible applications

11 SMEs initially selected



Data Partner Pilot Call

(Preliminary results)

Public review of the call
Initial phase during which all interested data partners can review the call description, ask questions and comment on the call description (July 15 – August 15).

Grant application portal open
Open call for data partners (September 1 – September 15)

Evaluation
Evaluation of all applications by our committee of both internal and external experts.

Agreement
Grant awarding and signing of grant agreement.

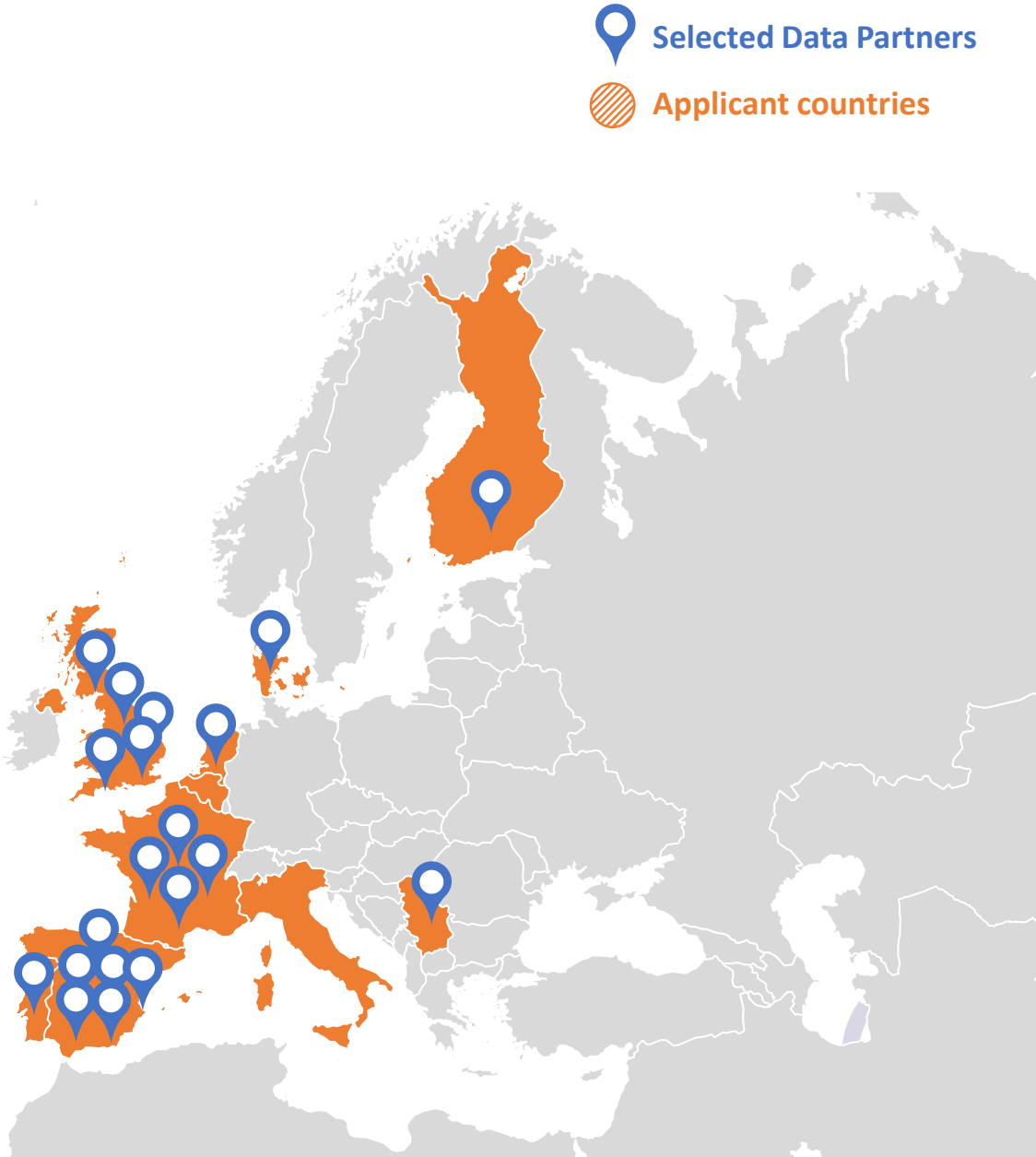
SME linking
Identification and linking up with the SME of choice.

Harmonisation
Initiation of the data harmonisation.

(Timelines are provisional)

20 Data Partners

>120 million patient records



imi efpia

2020 CALL TIMELINES

SME
The second open call for SMEs.

SME
The third open call for SMEs.

2020

JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC

Data partner
The second open call for Data partners.

Data partner
The third open call for Data partners.

imi efpia

USE-CASE DRIVEN

Use Case 1 – Drug utilisation study

Validate the OMOP CDM in European data sources comparing results from source versus OMOP data

Use Case 2 – Drug and device safety study

Test and further develop existing (OHDSI) population-level estimation and population-level prediction packages in EU Data sources

Use Case 3 – HTA study

Assess whether data commonly used for HTA purposes can easily be measured and analysed using the OMOP CDM

FAST OBSERVATIONAL RESEARCH IS FEASIBLE (STUDY-A-THONS…)

"To compare the **risk** of post-operative **complications** and **mortality** between unicompartmental *vs* total knee replacement."

Monday

Group consensus on the **problem**
Draft cohort definitions

Tuesday

Review clinical characterisation
Draft patient-level prediction design

Wednesday

Review patient-level prediction results
Externally validate prediction model

Thursday

Draft population-level effect estimation design
Review population-level effect estimation diagnostics

Friday

Review of results
Plan for completing **publications**

(EHDEN 1st Study-a-thon, Oxford, December 2018)

ACADEMY

Aim

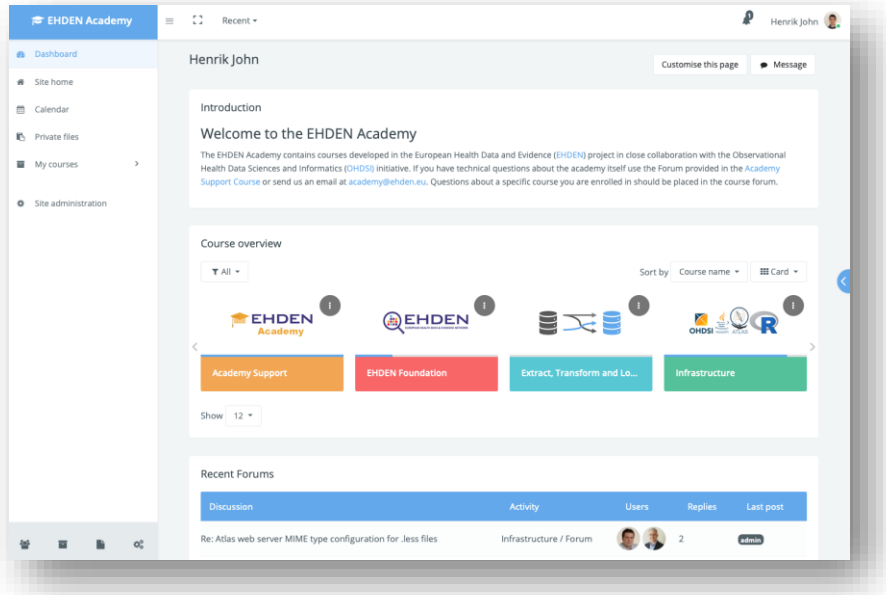
To develop an **e-learning** environment **to train all stakeholders** on the use of the tools and processes that are being adopted within EHDEN and OHDSI.

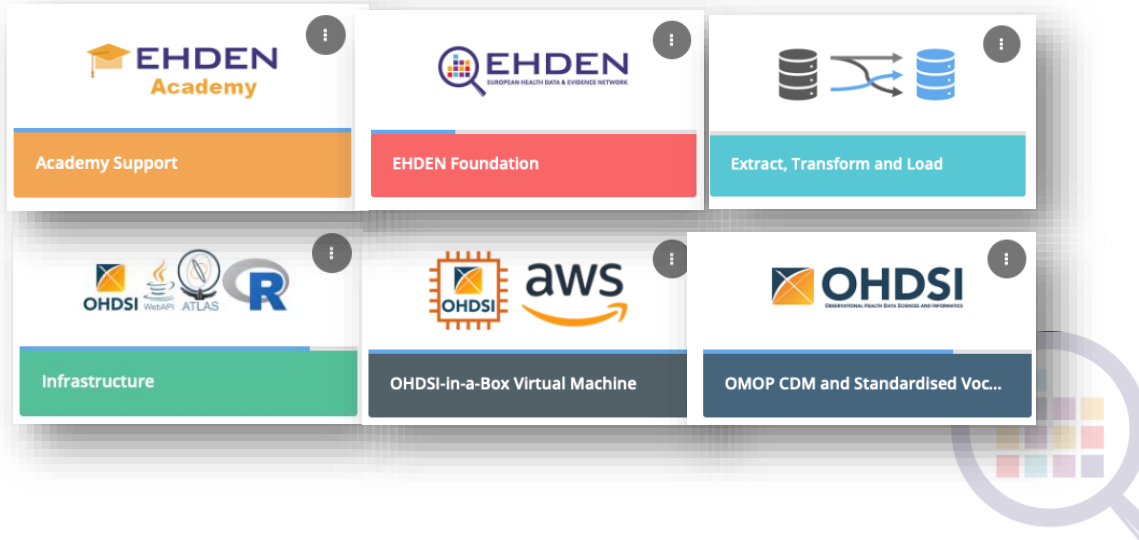
Collaboration

Course development on the OMOP Common Data Model and the rich set of OHDSI tools that are being developed in collaboration with the **OHDSI community.**

Infrastructure

The EHDEN Academy is developed in **Moodle** and is hosted in the **Amazon AWS cloud.** We use virtual machines for assignments.





QUESTIONS?

EHDEN
EUROPEAN HEALTH DATA & EVIDENCE NETWORK

https://forum.ehden.eu

enquiries@ehden.eu

www.ehden.eu

@IMI_EHDEN

IMI_EHDEN

github.com/EHDEN

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 806968. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

The Oxford Study-Athon

OUR JOURNEY WITH OHDSI AND EHDEN TO REAL-WORLD USEFUL EVIDENCE

Important clinical question

Patient-level data in CDM

Data partners standardized to OMOP CDM:
Iqvia
Janssen

ATLAS

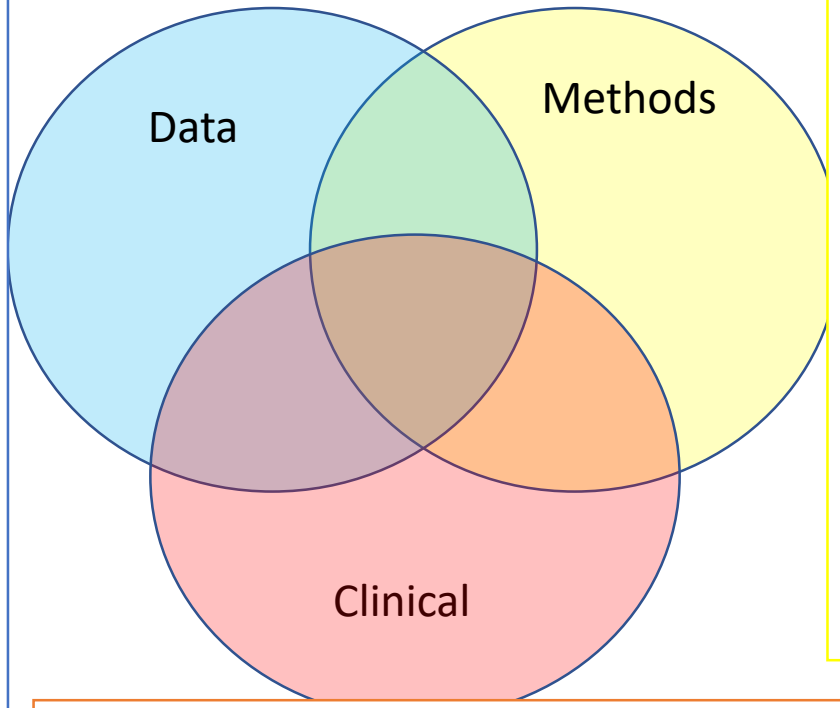
Standardized analysis tools from OHDSI

Reliable evidence

Valuable clinical answers disseminated to medical decision-makers

WHAT WE NEEDED IN THE MIX …

Hands-on knowledge of 1+ data source/s, including its structure and content, the provenance of the underlying population and data capture process, data quality issues and temporal variability, so that you can responsibly use the data to generate reliable evidence and recognize its limitations



Hands-on knowledge in designing studies and executing statistical methods to generate aggregate summary statistics from patient-level data. Different expertise required for clinical characterization, patient-level prediction, and population-level effect estimation

Direct knowledge of the diagnosis, treatment, and management of severe knee OA, including healthcare delivery, natural history and patient prognosis

Expertise Required

- Clinical knowledge in knee oa/arthroplasty?
- UK electronic health records (THIN)?
- US claims (MarketScan, Optum, PharMetrics)?
- OHDSI tools?
- R programming?
- Literature review?
- Publication writing?

Who has all of these prerequisites?

THE VOTED RESEARCH QUESTION…

"To compare the risk of **post-operative complications (infection, venous thrombo-embolism, mortality) and long-term implant revision** between **unicompartmental** *vs* **total knee replacement**."

WHAT WE KNEW BEFORE WE STARTED

- N = 60 quality studies
- From 1998 to 2018 (20 years of research!!)
- Reduced risk of VTE w UKR

Study	UKA No of events	UKA Total no of procedures	TKA No of events	TKA Total no of procedures	Risk ratio M-H Random (95% CI)	Weight (%)	Risk ratio M-H Random (95% CI)
Group 1							
Beard 2017	0	264	1	248		30.4	0.31 (0.01 to 7.65)
Newman 1998	1	50	5	52		69.6	0.21 (0.03 to 1.72)
Subtotal		314		300		100.0	0.24 (0.04 to 1.37)
Total (95% CI)	0		6				
Test for heterogeneity: τ^2=0.00; χ^2=0.04, df=1, P=0.83; I^2=0%							
Test for overall effect: Z=1.61, P=0.11							
Group 2							
Bolognesi 2013	0	3098	438	65 505		1.8	0.02 (0.00 to 0.39)
Courtney 2018	10	1351	823	49 136		21.2	0.44 (0.24 to 0.82)
Drager 2016	7	1340	580	36 274		17.0	0.33 (0.16 to 0.69)
Duchman 2014	10	1588	31	1588		18.1	0.32 (0.16 to 0.66)
Liddle 2015	86	25 334	525	75 996		41.9	0.49 (0.39 to 0.62)
Subtotal		32 711		228 499		100.0	0.39 (0.27 to 0.57)
Total (95% CI)	113		2397				
Test for heterogeneity: τ^2=0.08; χ^2=7.16, df=4, P=0.13; I^2=44%							
Test for overall effect: Z=4.82, P<0.001							
Group 3							
Berger 2009	0	25	1	86		7.7	1.12 (0.05 to 26.57)
Fabre Aubrespry 2018	3	101	12	101		51.0	0.25 (0.07 to 0.86)
Weal 2001	2	31	7	130		33.6	1.20 (0.26 to 5.49)
Yang 2003	0	50	1	50		7.7	0.33 (0.01 to 7.99)
Subtotal		207		367		100.0	0.49 (0.20 to 1.17)
Total (95% CI)	5		21				
Test for heterogeneity: τ^2=0.00; χ^2=2.82, df=3, P=0.42; I^2=0%							
Test for overall effect: Z=1.61, P=0.11							

Fig 3 | Forest plot comparing risk of venous thromboembolism after unicompartmental (UKA) versus total knee replacement (TKA). Also appears in the supplementary material as supplementary figure 5. M-H=Mantel-Haenszel test

WHAT WE KNEW BEFORE WE STARTED (2)

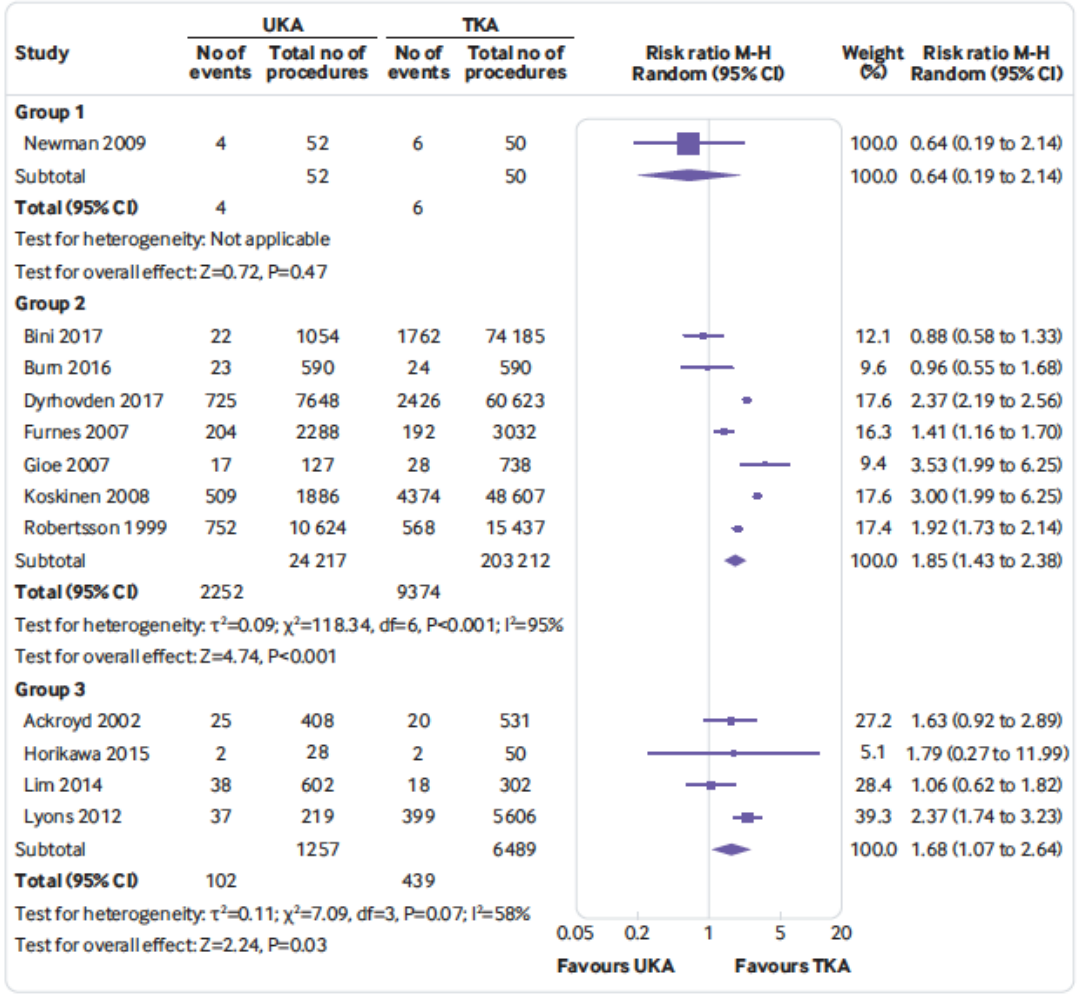
- Possibly a reduction in post-operative mortality …
- [although little data available on this]

Study	UKA No of events	UKA Total no of procedures	TKA No of events	TKA Total no of procedures	Risk ratio M-H Random (95% CI)	Weight (%)	Risk ratio M-H Random (95% CI)
Group 1							
Beard 2017	0	260	0	248			Not estimable
Costa 2011	0	34	0	34			Not estimable
Newman 1998	0	50	0	52			Not estimable
Sun 2010	0	28	0	38			Not estimable
Subtotal		314		300			
Total (95% CI)	0		6				
Test for heterogeneity: Not applicable							
Test for overall effect: Not applicable							
Group 2							
Bolognesi 2013	0	3098	455	65 505		3.3	0.02 (0.00 to 0.37)
Courtney 2018	0	1351	24	49 136		3.2	0.74 (0.05 to 12.9)
Drager 2016	0	1340	0	36 274			Not estimable
Duchman 2014	0	1588	3	1588		2.9	0.14 (0.01 to 2.76)
Hunt 2014	33	40,428	1,183	467 779		61.6	0.32 (0.23 to 0.46)
Liddle 2015	7	25358	90	76 074		29.1	0.23 (0.11 to 0.50)
Subtotal		75,163		696 356		100.0	0.27 (0.16 to 0.46)
Total (95% CI)	40		1755				
Test for heterogeneity: τ^2=0.08; χ^2=5.04, df=4, P=0.13; I^2=21%							
Test for overall effect: Z=5.02, P<0.001							
Group 3							
Yang 2003	0	50	0	50			Not estimable
Subtotal		50		50			
Total (95% CI)	0		0				
Test for heterogeneity: Not applicable							
Test for overall effect: Not applicable							

g 4 | Forest plot comparing risk of early mortality (at 45 days) after unicompartmental (UKA) versus total knee placement (TKA). Also appears in the supplementary material as supplementary figure 7. M-H=Mantel-Haenszel test

WHAT WE KNEW BEFORE WE STARTED (3)

- … BUT
- An increase (around double) in long-term revision risk



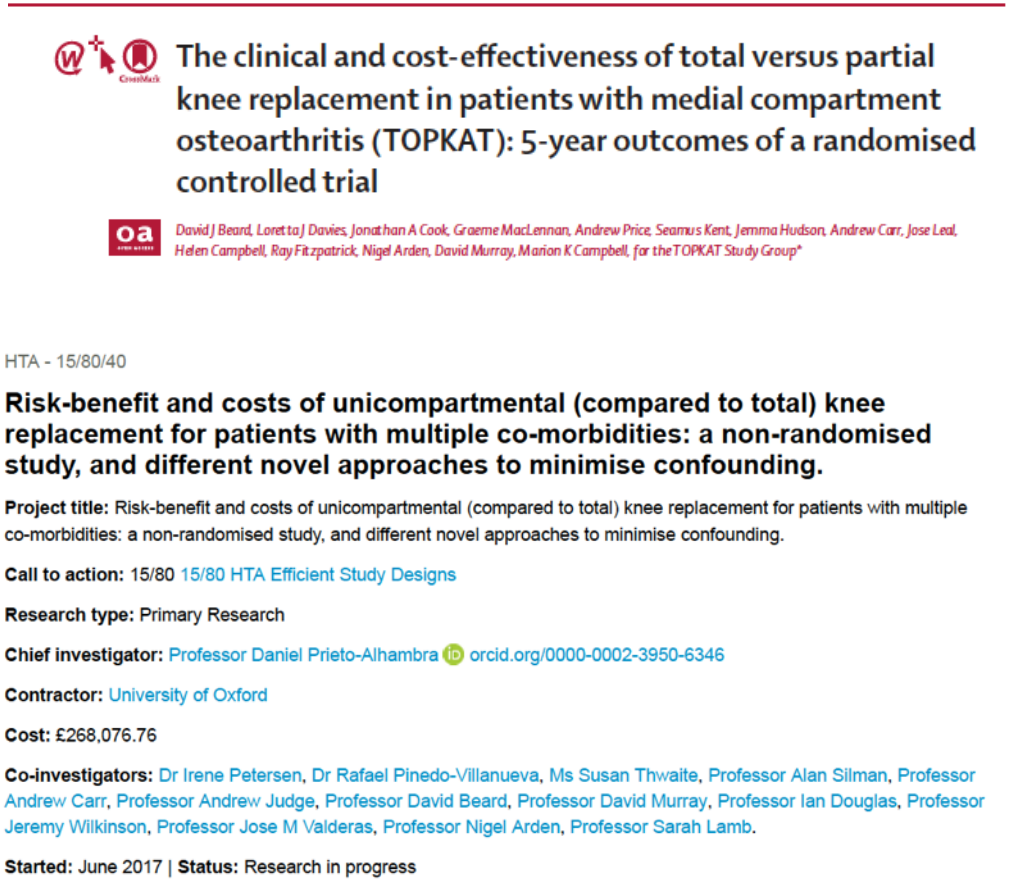
Study	UKA No of events	UKA Total no of procedures	TKA No of events	TKA Total no of procedures	Risk ratio M-H Random (95% CI)	Weight (%)	Risk ratio M-H Random (95% CI)
Group 1							
Newman 2009	4	52	6	50		100.0	0.64 (0.19 to 2.14)
Subtotal		52		50		100.0	0.64 (0.19 to 2.14)
Total (95% CI)	4		6				
Test for heterogeneity: Not applicable							
Test for overall effect: Z=0.72, P=0.47							
Group 2							
Bini 2017	22	1054	1762	74 185		12.1	0.88 (0.58 to 1.33)
Burn 2016	23	590	24	590		9.6	0.96 (0.55 to 1.68)
Dyrhovden 2017	725	7648	2426	60 623		17.6	2.37 (2.19 to 2.56)
Furnes 2007	204	2288	192	3032		16.3	1.41 (1.16 to 1.70)
Gioe 2007	17	127	28	738		9.4	3.53 (1.99 to 6.25)
Koskinen 2008	509	1886	4374	48 607		17.6	3.00 (1.99 to 6.25)
Robertsson 1999	752	10 624	568	15 437		17.4	1.92 (1.73 to 2.14)
Subtotal		24 217		203 212		100.0	1.85 (1.43 to 2.38)
Total (95% CI)	2252		9374				
Test for heterogeneity: τ^2=0.09; χ^2=118.34, df=6, P<0.001; I^2=95%							
Test for overall effect: Z=4.74, P<0.001							
Group 3							
Ackroyd 2002	25	408	20	531		27.2	1.63 (0.92 to 2.89)
Horikawa 2015	2	28	2	50		5.1	1.79 (0.27 to 11.99)
Lim 2014	38	602	18	302		28.4	1.06 (0.62 to 1.82)
Lyons 2012	37	219	399	5606		39.3	2.37 (1.74 to 3.23)
Subtotal		1257		6489		100.0	1.68 (1.07 to 2.64)
Total (95% CI)	102		439				
Test for heterogeneity: τ^2=0.11; χ^2=7.09, df=3, P=0.07; I^2=58%							
Test for overall effect: Z=2.24, P=0.03							

Fig 7 | Forest plot comparing incidence of revision at 10 years after unicompartmental (UKA) versus total knee replacement (TKA). Also appears in the supplementary material as supplementary figure 16. M-H=Mantel-Haenszel test

What we knew before we started (5)

- Caveats with quality of these 60 papers (and 20y) of data (mostly observational and from different sources)

- NIHR UK-funded
  - 1 surgical RCT (TOPKAT)
  - 1 observational study (UTMOST)



The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial

David J Beard, Loretta J Davies, Jonathan A Cook, Graeme MacLennan, Andrew Price, Seamus Kent, Jemma Hudson, Andrew Carr, Jose Leal, Helen Campbell, Ray Fitzpatrick, Nigel Arden, David Murray, Marion K Campbell, for the TOPKAT Study Group*

HTA - 15/80/40

Risk-benefit and costs of unicompartmental (compared to total) knee replacement for patients with multiple co-morbidities: a non-randomised study, and different novel approaches to minimise confounding.

Project title: Risk-benefit and costs of unicompartmental (compared to total) knee replacement for patients with multiple co-morbidities: a non-randomised study, and different novel approaches to minimise confounding.

Call to action: 15/80 15/80 HTA Efficient Study Designs

Research type: Primary Research

Chief investigator: Professor Daniel Prieto-Alhambra orcid.org/0000-0002-3950-6346

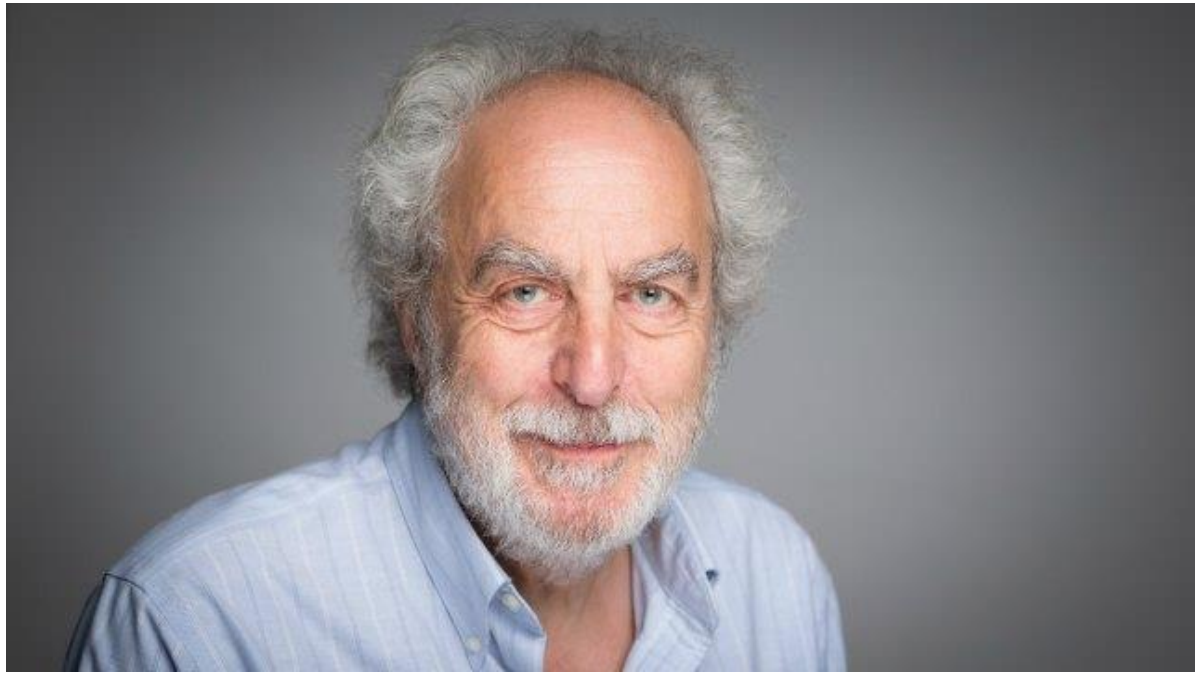
Contractor: University of Oxford

Cost: £268,076.76

Co-investigators: Dr Irene Petersen, Dr Rafael Pinedo-Villanueva, Ms Susan Thwaite, Professor Alan Silman, Professor Andrew Carr, Professor Andrew Judge, Professor David Beard, Professor David Murray, Professor Ian Douglas, Professor Jeremy Wilkinson, Professor Jose M Valderas, Professor Nigel Arden, Professor Sarah Lamb.

Started: June 2017 | **Status:** Research in progress

MOTIVATION



"We need less research, better research,
And research done for the right reasons"

i.e., can we do in a week a study what has
taken so far 20+ years, 60+ papers and loads of cash?

Intended for healthcare professionals

thebmj

Editorials

The scandal of poor medical research

BMJ 1994 ; 308 doi:
https://doi.org/10.1136/bmj.308.6924.283
(Published 29 January 1994)
Cite this as: *BMJ* 1994;308:283

Linked Opinion
Richard Smith: Medical research—still a scandal

Article | Related | Metrics | Responses

D G Altman

We need less research, better research, and research done for the right reasons

OUR AIMS!

- Can we ‘predict’ the TOPKAT results (on complications) before they publish?

- And more:
  - can we report on the results of UKR (vs TKR) in the older, more complex patients, excluded from TOPKAT?
  - Can we predict who is likely to have a post-operative complication following knee replacement surgery

WE CAN DO THIS IN ONE WEEK (STUDY-A-THON)??

"To compare the **risk** of post-operative **complications** and **mortality** between unicompartmental *vs* total knee replacement."

Monday

Group consensus on the **problem**
Draft cohort definitions

Tuesday

Review clinical characterisation
Draft patient-level prediction design

Wednesday

Review patient-level prediction results
Externally validate prediction model

Thursday

Draft population-level effect estimation design
Review population-level effect estimation diagnostics

Friday

Review of results
Plan for completing **publications**

(EHDEN 1st Study-a-thon, Oxford, December 2018)

First OHDSI-EHDEN Study-A-thon! Our final (refined) research question!!

"To compare the risk of **post-operative complications (infection, revision, and venous thrombo-embolism) and mortality** between **uni-compartmental** *vs* **total knee replacement**."

LET'S START COLLABORATING!

- Open the shared group notes: Link

- Ground rules:
  - During group exercises, take all your notes here together
  - During breakout exercises, assign one person in your team to make sure notes are recorded so other groups can learn from our experience

LET'S START WRITING OUR PAPERS!

- Patient-Level Prediction: Link
- Population-Level Effect Estimation: Link

LET'S START LEARNING ATLAS!

- Public version from OHDSI (v2.6, simulated data), go to: http://ohdsi.org/web/ATLAS

- Private version from IQVIA (v2.4, THIN data), go to: https://training.atlasplus.imshealth.com

SO WHAT DID WE LEARN (BY FRIDAY!!)

- Population-level effect estimation:

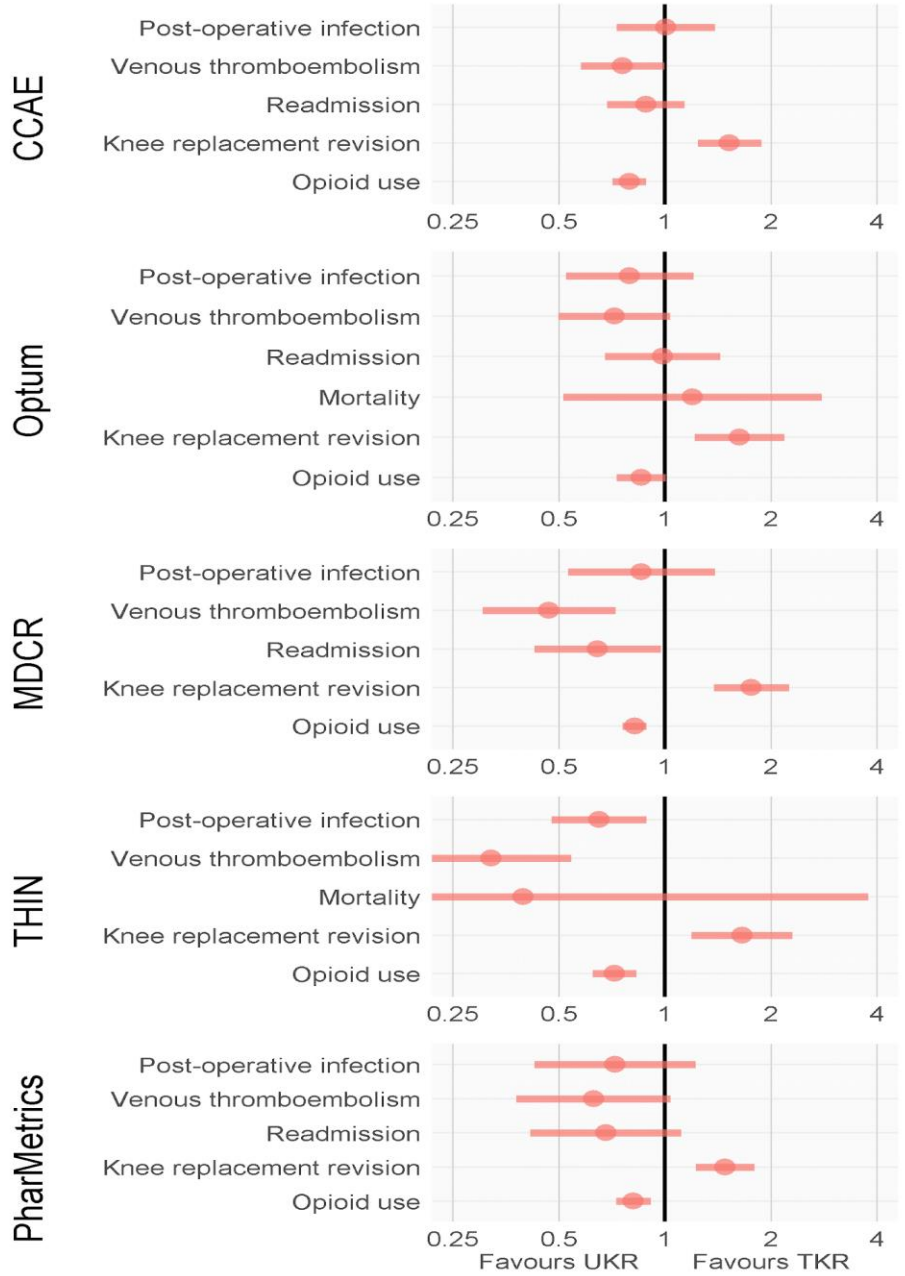
http://data.ohdsi.org/UkaTkaSafetyEffectiveness/

- Patient-level prediction

http://data.ohdsi.org/TKROutcomesExplorer/

OUR WEEK VS 20Y OF RESEARCH OX STUDY-A-THON VS BMJ SYST REV

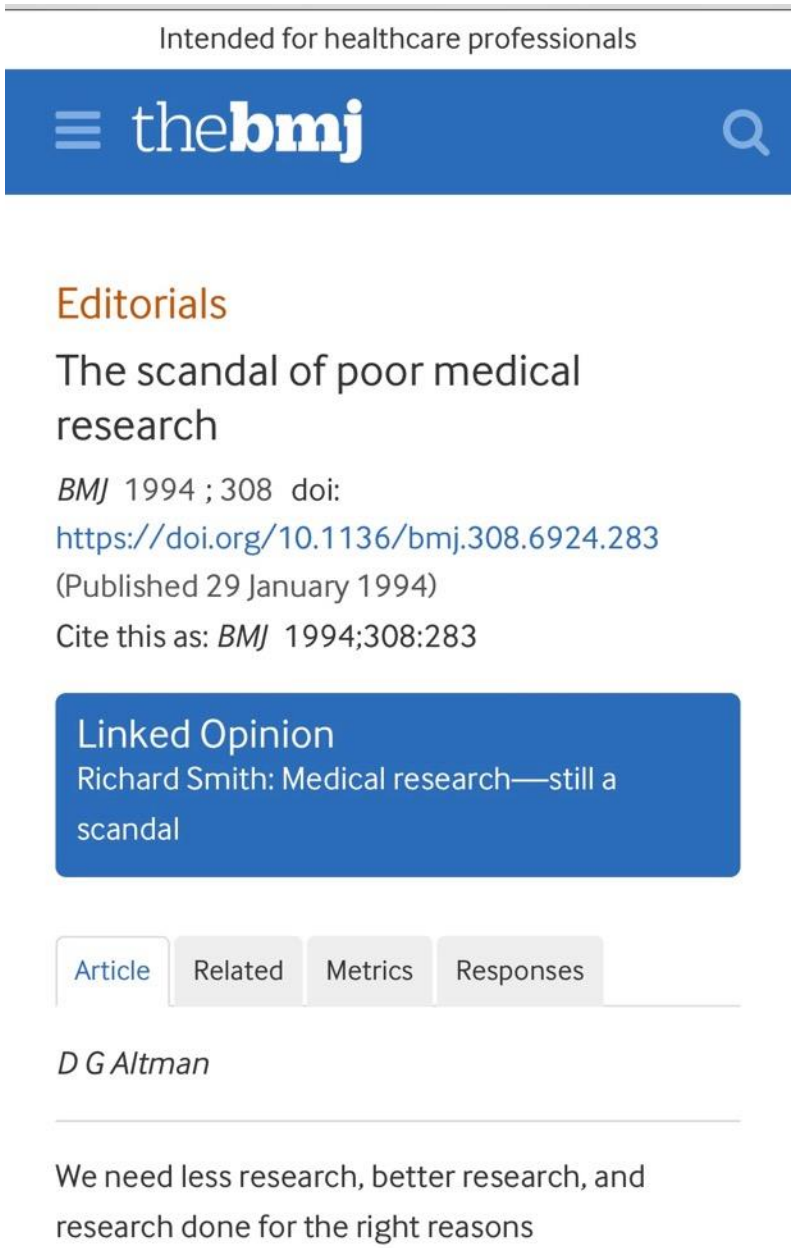
- VTE
  - RR 0.49 [0.20 to 1.17] (20 y)
  - vs HR 0.62 [0.36-0.96] (1 week)
- Long-term revision
  - RR 1.68 [1.07 to 2.64] (20y)
  - vs HR 1.51 to 2.16 (1 week)



OUR WEEK VS A £3M TRIAL [D BEARD ET AL. LANCET 2019]

- Small improvement in pain/function with UKR in TOPKAT
- = Small reduction in opioid/s use in Study-a-thon

- No power for safety in TOPKAT
- Findings compatible w 20y of data in Study-a-thon



Intended for healthcare professionals

thebmj

Editorials

The scandal of poor medical research

BMJ 1994 ; 308 doi: https://doi.org/10.1136/bmj.308.6924.283 (Published 29 January 1994)

Cite this as: *BMJ* 1994;308:283

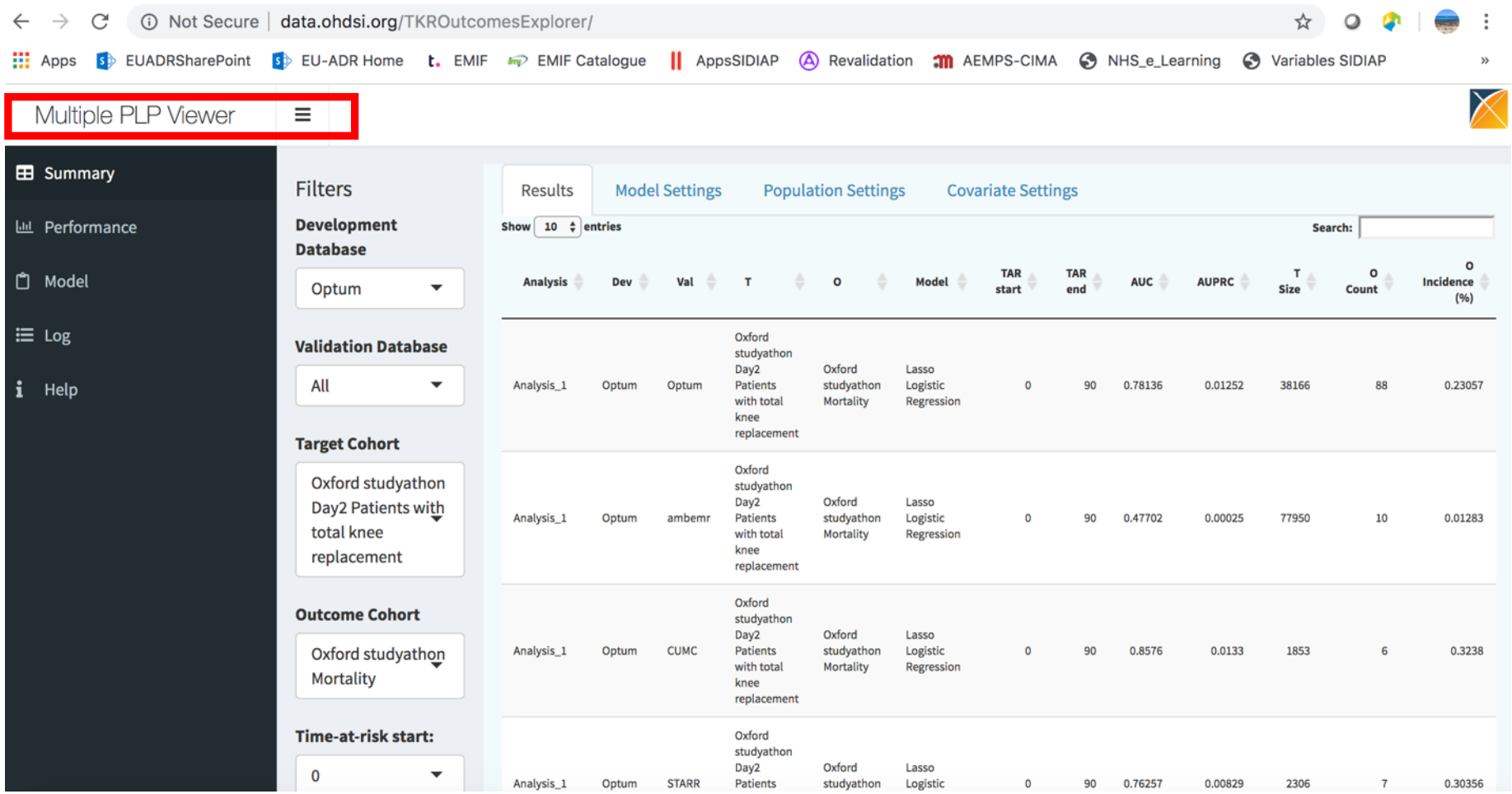
Linked Opinion
Richard Smith: Medical research—still a scandal

Article | Related | Metrics | Responses

D G Altman

We need less research, better research, and research done for the right reasons

PREDICTION… PREDICTING POST-OP MORTALITY



Analysis	Dev	Val	T	O	Model	TAR start	TAR end	AUC	AUPRC	T Size	O Count	O Incidence (%)
Analysis_1	Optum	Optum	Oxford studyathon Day2 Patients with total knee replacement	Oxford studyathon Mortality	Lasso Logistic Regression	0	90	0.78136	0.01252	38166	88	0.23057
Analysis_1	Optum	ambemr	Oxford studyathon Day2 Patients with total knee replacement	Oxford studyathon Mortality	Lasso Logistic Regression	0	90	0.47702	0.00025	77950	10	0.01283
Analysis_1	Optum	CUMC	Oxford studyathon Day2 Patients with total knee replacement	Oxford studyathon Mortality	Lasso Logistic Regression	0	90	0.8576	0.0133	1853	6	0.3238
Analysis_1	Optum	STARR	Oxford studyathon Day2 Patients	Oxford studyathon	Lasso Logistic	0	90	0.76257	0.00829	2306	7	0.30356

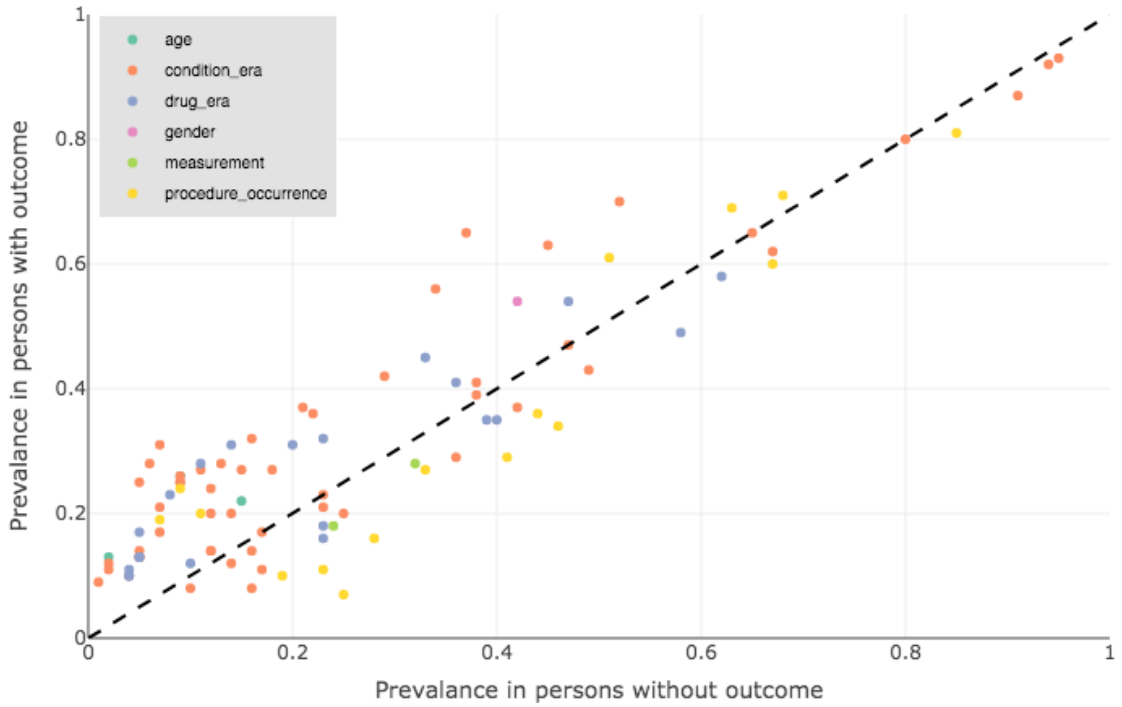
THE MODEL

Model Table

Download Model

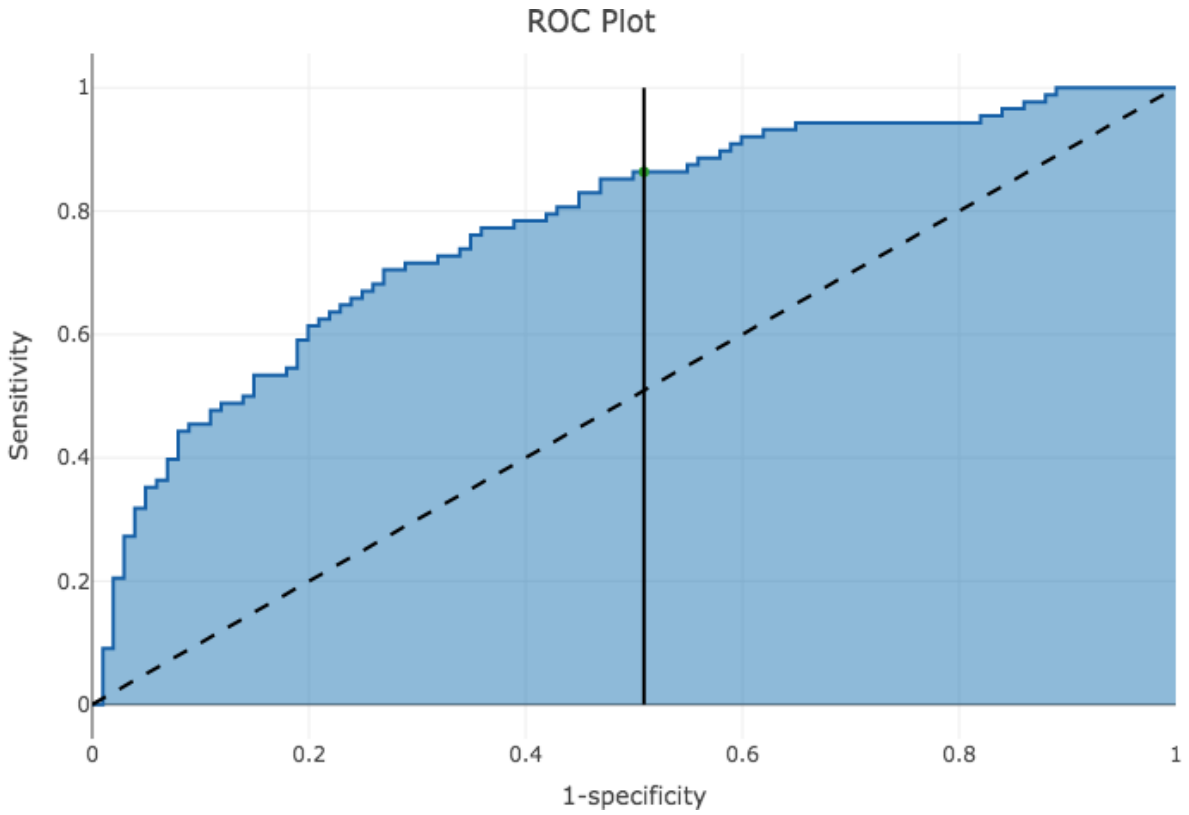
Show 10 entries Search:

	Covariate Name	Value	Outcome Mean	Non-outcome Mean
1	index month: 1	0	0.1	0.1
2	Charlson index - Romano adaptation	0	3.81	2.16
3	Diabetes Comorbidity Severity Index (DCSI)	0	3.71	1.96
4	CHADS2	0	2.54	1.57
5	CHADS2VASc	0	4.18	2.94
6	visit_occurrence concept count during day -1095 through -1 concept_count relative to index	0	71.99	50.13
7	visit_occurrence concept count during day -365 through -1 concept_count relative to index	0	32.46	24.26
8	index month: 2	0	0.1	0.08
9	index month: 3	0	0.1	0.09
10	index month: 4	0	0.08	0.08

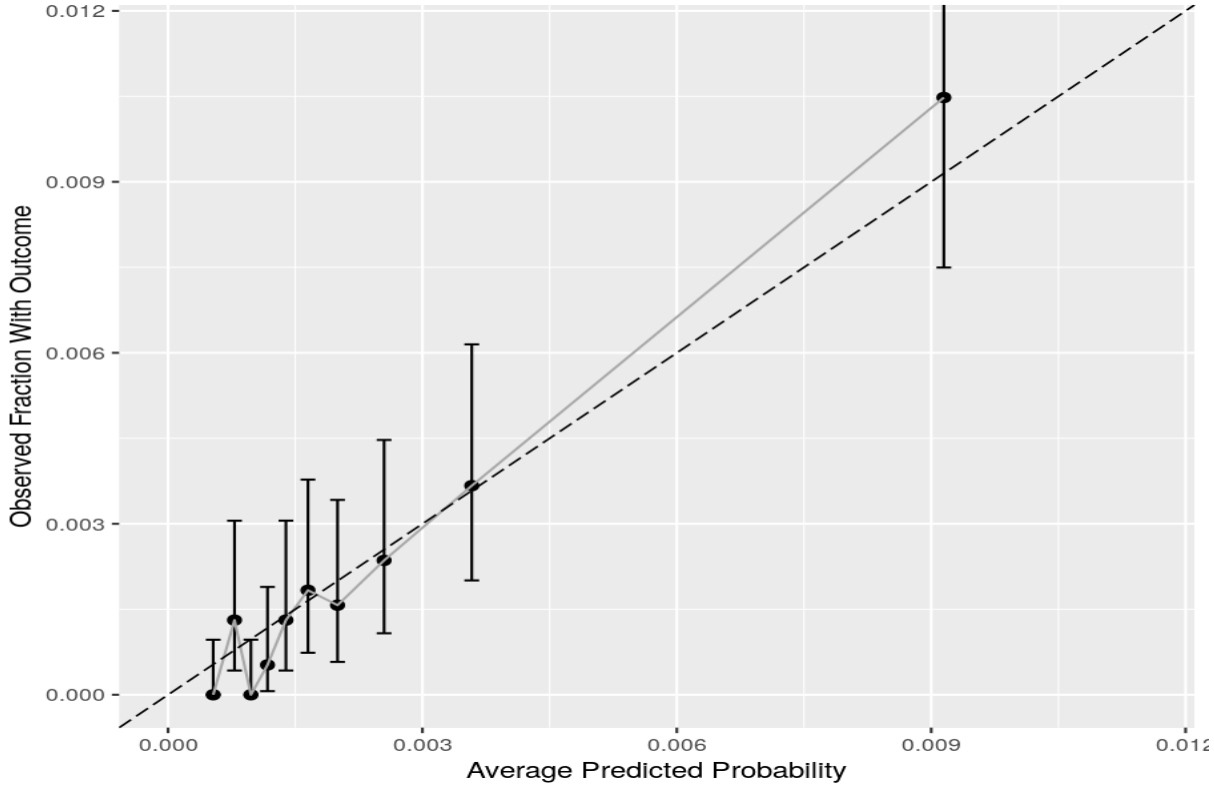


PREDICTION… PREDICTING POST-OP MORTALITY

DISCRIMINATION



CALIBRATION



OUTPUTS

- **Conference/s**
  - 1 Podium ppt at IOF/ESCEO Paris'19 (>3,000 researchers)
  - 1 Podium ppt and Press Release at EULAR Madrid '19 (>15,000 rheumatologists)
  - 1 Podium ppt and 1 Spotlight poster at ICPE'19 (>1,500 epidemiologists)

- **Scientific journal manuscripts**
  - PLE = Major review at Lancet Rheumatology
  - PLP = Submitted to JAMA Surgery

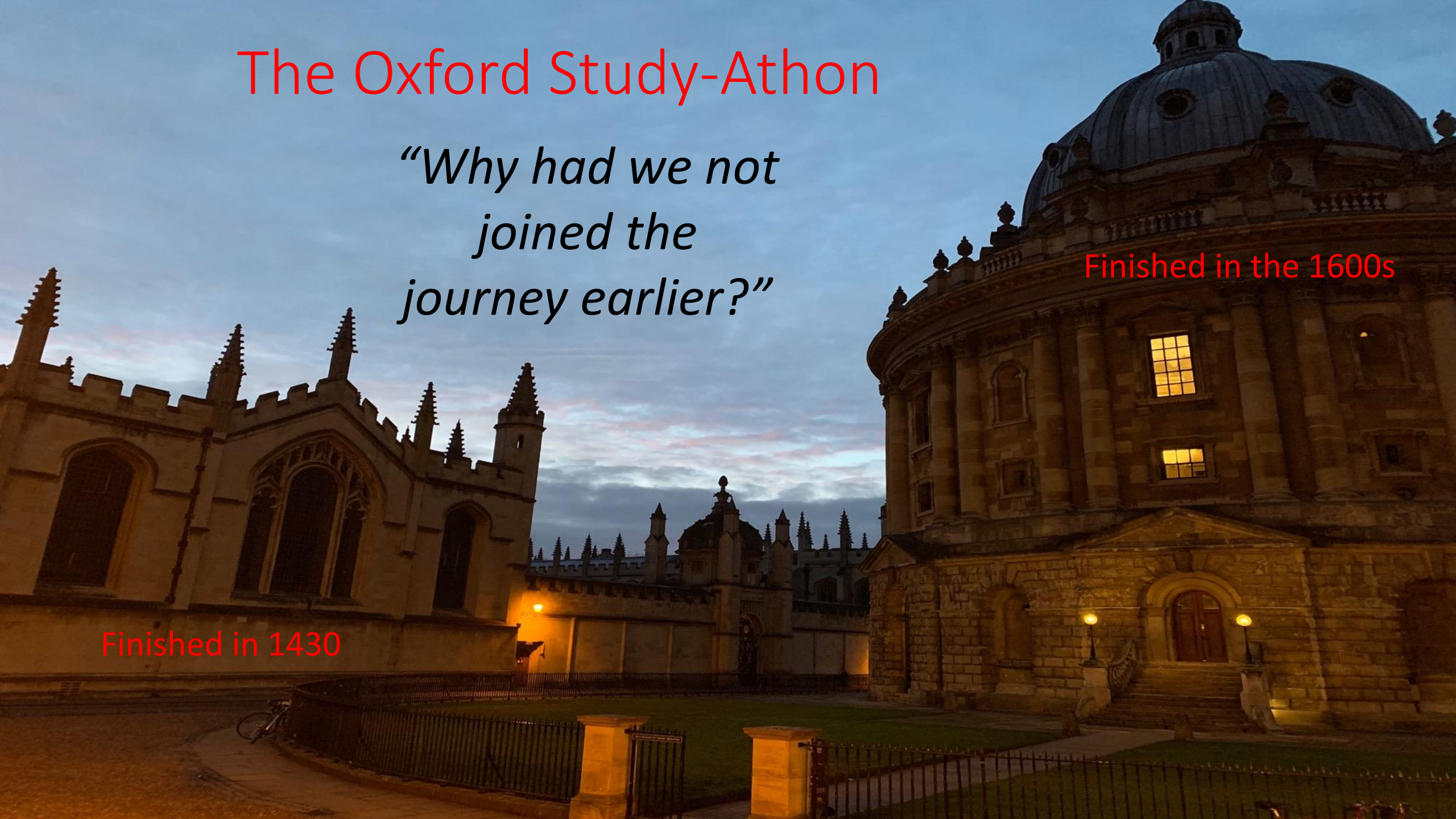
AND WHAT DID I LEARN (BY FRIDAY TOO!!)

The Oxford Study-Athon

"Why had we not joined the journey earlier?"

Finished in the 1600s

Finished in 1430



BACKGROUND

eular

fighting rheumatic & musculoskeletal diseases together

We knew this was coming…

After 7 years of silence …

EULAR recommendations for the management of rheumatoid arthritis – 2019 Update
Smolen J et al. Ann Rheum Dis 2020 Feb

QUESTIONS ???