European Medical Information Framework

Bart Vannieuwenhuyse
Janssen R&D
November 2017
Defining Real World Data (RWD)

RWD is generated using data typically collected in usual health care settings.
RWE is most commonly generated using a range of non-interventional (observational) studies, including:

- **Primary data** collections such as registries collecting prospective and/or retrospective data, or **surveys** collecting cross-sectional or retrospective information.
- Analyses of **secondary data that includes (electronic) medical records, insurance claims data, and government databases** which provide data typically used for retrospective analyses.
Why is EMIF needed?
Secondary use of health data to enrich research

The “burning platform” for life sciences

Challenge

Today, Pharma doesn’t have ready access to this data, yet insights for safety, CER and other areas are within this clinical domain, which includes medical records, pharmacy, labs, claims, radiology etc.

The value of healthcare data for secondary uses in clinical research and development — Gary K. Mallow, Merck, HIMSS 2012
Acetylcholinesterase inhibitor (Galantamine) therapy in Alzheimer’s disease

Results from placebo-controlled clinical study

- Mini-Mental State Examination (MMSE) score provides a global impression of cognitive functioning
- Galantamine Tx => 8 months delay of disease progression (vs. placebo)
- Are similar treatment effects also obtained in a “real life” setting?
Successful example of data re-use for research

Cholinesterase inhibitors and Alzheimer’s disease

• RWE of AChEi
  > 2500 patient years of therapy documented in EHR-system
  > 8 fold dataset compared to Cochrane

• Cost effective
  Text mining derivation of service utilisation and costs.
  Created in one month

• Resembles cognitive decline curve derived from clinical trials

S. Lovestone et al, unpublished 2012
Project overview

14 European countries combining 58 partners

€56 million worth of resources

3 projects in one

5 year project (2013 – 2017)

ACADEMIC PARTNERS

SME PARTNERS

EFPIA PARTNERS

PATIENT ORGANISATION

58 Partners
Our vision

To become the trusted European hub for health care data intelligence, enabling new insights into diseases and treatments
Project objectives

**EMIF-Platform**

Develop a framework for evaluating, enhancing and providing access to human health data across Europe, support EMIF-Metabolic and EMIF-AD (the specific topics below) as well as support research using human health data in general.

**EMIF-Metabolic**

Identify predictors of metabolic complications in obesity.

**EMIF-AD**

Identify predictors of Alzheimer’s Disease (AD) in the pre-clinical and prodromal phase.
EMIF system overview

Data Sources
- 1° care
- Hospital
- Admin
- Regional
- Registries & cohorts
- Biobanks
- 2° care
- Paediatric

Governance
- EMIF platform solution

Data owners
- Site Y
- Site Z

Common Ontology / De-identification

Extract

User admin
- User admin
- User admin

Remote user 1
- Remote user 2

Researchers

Data Sources

1° care
Hospital
Admin
Regional
Registries & cohorts
Biobanks
2° care
Paediatric
Available data types

Large variety in “types” of data

- Primary care data sets
- Hospital data
- Administrative data
- Regional record-linkage systems
- Registries and cohorts (broad and disease specific)
- Biobanks
- Secondary care data sets
- Paediatric data sets

Data is available from more than 40 million subjects from six EU countries, and in addition:

- more than 60,000 subjects in AD cohorts
- more than 94,000 subjects in metabolic cohorts
Available data sources

EMIF-Available Data Sources; EXAMPLES

Approximate total (cumulative) number of subjects

- THIN: 12M
- PHARMO: 10M
- SIDIAP: 6M
- ARS: 3.6M
- IPCI: 2.8M
- AUH: 2.3M
- HSD: 1.6M
- IMASIS: 1M
- SCTS: 475K
- PEDIANET: 400K
- EGCUT: 52K
- SDR: 2K
- MAAS: 1K

Status Jan 2016
Starts with knowing where the data is ...

EMIF Catalogue (request access via emif.eu)
EMIF – Full suite of tools
Prevalence amyloid positivity in non-demented subjects

Jansen et al JAMA 2015
Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University.

www.ohdsi.org
OHDSI-tools
Power of distributed data --

Collaboration with 11 data sets representing 255MM subjects

Characterizing treatment pathways at scale using the OHDSI network
High collaboration – high impact

FIGURE 6.4.1 COLLABORATION INDEX VERSUS CITATION IMPACT PER PROJECT

Citation impact (Field-normalised)

EMIF

Collaboration Index
Follow-up project IMI-EHDN

European Health data network (EHDN)

Objectives:

- The first goal of the EHDN is to implement the approaches pioneered in earlier research projects (e.g. EMIF) and develop a standard methodology.

- The second goal of EHDN is to help mature both the supply side and the demand side of this ‘health data ecosystem’ in compliance with robust privacy and ethics governance.

- The third goal of EHDN is to stimulate development of new and augmented health services through available and expanded technologies, in the interest of health outcomes.

An operational network of hospital and other quality data sets covering up to 20% of the EU population or approx. 100 million persons in support of existing and new BD4BO and other health outcome related initiatives;

Enable inclusion of novel data types and use of the data by new stakeholder groups;

Define and clarify value proposition for data providers including an agreed strategy that will maximise the number of data providers that are willing to put data into the EHDN;

Disease specific BD4BO projects will have resource and capability to enable access to the multiple data repositories that these projects identify, including guidance on mapping to the common data model;
Conclusion

• Large scale project with multiple stakeholders and data providers. This would be very hard for one pharma company to set up

• Scale brings complexity – project took a while to get to “full steam” but results are now coming at much faster rate

• Reuse of existing data is cost (and time) efficient.

• High impact publications
More Information

• EMIF general
  • Bart Vannieuwenhuyse (bvannieu@its.jnj.com)
  • Simon Lovestone (simon.lovestone@psych.ox.ac.uk)
  • Johan van der Lei (j.vanderlei@erasusmc.nl)

• EMIF-Platform
  • Johan van der Lei (j.vanderlei@erasusmc.nl)
  • Nigel Hughes (nhughes@its.jnj.com)

• EMIF-Metabolic
  • Ulf Smith (ulf.smith@medic.gu.se)
  • Dawn Waterworth (dawn.m.waterworth@gsk.com)

• EMIF-AD
  • Pieter Jelle Visser (pj.visser@maastrichtuniversity.nl)
  • Johannes Streffer (jstreffe@its.jnj.com)

EMIF is operating under IMI Grant Agreement n°115372