Lightning Talks Session

Adoption of the OMOP-CDM
The European Health Data & Evidence Network

What is it?

Nigel Hughes
EFPIA Coordinator
Scientific Director, JCI Patient Data for Research, Janssen
All too often real world research is a challenging journey....
What we need is infrastructure and a network to conduct real world research in the 21st Century....
IMI: a unique Public-Private Partnership (PPP) between the pharmaceutical industry represented by the European Federation of Pharmaceutical Industries and Associations (EFPIA) and the European Union represented by the European Commission.
Federated distributed data networks reflect real world patients with greater phenotypic resolution: ‘big data’

A federated data network is based around a common platform to facilitate bi-directional data flow, data harmonisation (from diverse sources) and connectivity for research reuse; supported by a harmonisation fund and certified/qualified SME network.

RWD is about diversity of source as no one data source is the whole truth. A federated data network retains local provenance and governance while facilitating remote connectivity with RWD for research reuse.

There are a multitude of use cases for RWD, but the more reflective of populations, cohorts and real world patient trajectories the more reliability can be attributed to RWD/RWE.

For value and outcomes patients, their treatment episodes, and outcomes need to be tracked longitudinally across health domains.

There are a multitude of use cases for RWD, but the more reflective of populations, cohorts and real world patient trajectories the more reliability can be attributed to RWD/RWE.
Key Components of Collaborative Data Projects....
EHDEN Public & EFPIA Consortium Working on Full Project Proposal

<table>
<thead>
<tr>
<th>Institute</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erasmus MC</td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Synapse Research Management</td>
<td>Spain</td>
</tr>
<tr>
<td>Oxford University</td>
<td>UK</td>
</tr>
<tr>
<td>Tartu Ulikool</td>
<td>Estonia</td>
</tr>
<tr>
<td>University of Aveiro</td>
<td>Portugal</td>
</tr>
<tr>
<td>The Hyve</td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Odysseus Data Services</td>
<td>Czech Republic</td>
</tr>
<tr>
<td>European Patients Forum</td>
<td>Luxembourg</td>
</tr>
<tr>
<td>National Institute for Health and Care Excellence (NICE)</td>
<td>UK</td>
</tr>
<tr>
<td>Stiftelsen WHO Collaborating Centre for International Drug Monitoring</td>
<td>Sweden</td>
</tr>
<tr>
<td>International Consortium for Health Outcomes Measurement (ICHOM)</td>
<td>UK</td>
</tr>
</tbody>
</table>
Current Progress on EHDEN - 2018

- EHDEN is an IMI flagship project
- Key stakeholders and actors in this domain recognise the RWD→RWE process in the EU is inadequate
- For the coming year pre-project preparation is critical for a Q4 start

**Key Activities Q1-Q3**

- Consortium Development
- Vision Alignment
- Work Packages
  - Advancing Short Project Proposal
  - Aligning Consortium
  - Linkage with e.g. OHDSI
  - Developing ‘The Plan’

**Final Project Proposal**

EHDEN is an IMI flagship project

Key stakeholders and actors in this domain recognise the RWD→RWE process in the EU is inadequate

For the coming year pre-project preparation is critical for a Q4 start
What is the European Health Data & Evidence Network (EHDEN)?

The aim is to map 100 million health records across the EU via a common data model (OMOP), supporting research, the BD4BO IMI2 programme, and outcomes-based healthcare.

- Advance the skill base in the EU via certified & qualified SMEs
- Harmonisation fund to support mapping to CDM
- Building on prior and current (IMI) projects
- Accelerate the platform and supporting research process
- A federated network of increasing number of Data Custodians and Sources
- Evaluation and incorporation of ‘novel’ data sources
- Utilisation of FAIR Principles
- Incorporation of outcome standards (e.g. ICHOM)
- Supporting the evidence base in the EU for outcomes-based research and medicine
- EHDEN will be an IMI consortium of 12 EFPIA partners and a public consortium of 11 partners, commencing c. October 2018
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 standardised to a common data model.

Values:
EHDEN Proposed Work Packages

**WP1: Evidence Workflow Development**
Incorporating the use cases for supporting development and validation of the EHDEN socio-technical approach, inclusive of BD4BO projects

**WP2: Outcome Driven Healthcare**
Related to all activities specific to e.g. BD4BO projects outcome focus, and ICHOM standards incorporation

**WP3: Personalized Medicine**
Focusing on the support of outcomes/value based healthcare, inclusive of clinical prediction models, with the incorporation of ‘novel’ patient data

**WP4: Technical Implementation**
Key priority is socio-technical development of the EHDEN federated framework and relevant services

**WP5: Data Workflow Implementation & Service Deployment**
Development, oversight and evaluation of the ecosystem development from SME qualification/certification to data source engagement, OMOP CDM mapping and evaluation

**WP6: Outreach and Sustainability**
Ensuring the development of stakeholder mapping and engagement with regards to use of RWD, and acceptability of RWE, while developing the sustainable operational model for EHDEN during and post IMI phase

**WP7: Project Management and Dissemination**
Concentrating on intra-project project management, internal communications and external dissemination, and responding to IMI deliverables

---

**Build**  **Fuel**  **Drive**
**EHDEN is not a blank canvas project....**

- EHDEN will utilise the best from prior and current (IMI) projects and platforms to develop its own infrastructure

- It will be developing its federated framework, but also analytical tooling and processes to support observational research in the EU, closely associated with the OHDSI EU/Global Open Science/Open Source collaboration
Critical at the heart of EHDEN is the OMOP CDM and OHDSI Collaboration
What is the flow through EHDEN?

Harmonisation Fund

Engagement & Quid Pro Quo

A multitude of data sources in the EU

Qualification & Certification Process

Developing the SME domain in the EU

WP1: Evidence Workflow Development
WP2: Outcome Driven Healthcare
WP3: Personalized Medicine

Use Cases, Demonstration Projects & Developing Evidence

COMMON DATA PLATFORM

Academia
Discovery
Development
Post-Authorisation
HTA/Outcomes
Research
EHDEN: In Summary

- EHDEN is a bold step in recognising that a flagship project is needed to address conducting real world research for the 21st Century.

- At its heart is the acknowledgement that we need to develop a community via a federated network, within an ecosystem, all based on a *quid pro quo* around data for research use.

- The Open Science/Open Source community of OHDSI, and the OMOP CDM are critical enablers and partners for EHDEN in this endeavour.
The German Data Network

Martin Sedlmayr\textsuperscript{1}, Christian Maier\textsuperscript{2}, Thomas Ganslandt\textsuperscript{3}, Hans-Ulrich Prokosch\textsuperscript{2} on behalf of the MIRACUM consortium

\textsuperscript{1}Institute for Medical Informatics and Biometry, TU Dresden
\textsuperscript{2}Chair of Medical Informatics, FAU Erlangen-Nürnberg
\textsuperscript{3}Department of Biomedical Informatics, University Medicine Mannheim, Ruprecht-Karls-University Heidelberg
Motivation

• We need
  – coordinated
  – secure
  – data protected
  – governed
• access to data and tools
• based on a trustworthy framework of technologies and policies
The German Medical Informatics Funding Scheme

- Improve **research opportunities** and **patient care** through innovative IT solutions (initially at university hospitals)

- Intensify the **exchange and sharing of data** between the research community and the health care delivery system

- Position medical informatics as a progressive field in research, **teaching and continuing education**

- As a key element of the funding scheme **data integration centres** are to be set up and interlinked by the university hospitals
The initiative so far...

Announcement 2015

Conceptual Phase 2016-2017

Development and Networking Phase 2018-2021

Supplementary Funding for
• new professorships
• young researcher groups and
• non-funded university hospitals
Governance of the MI-I

- **TMF**
  - Umbrella organization for networked medical research in Germany
- **VUD**
  - Association of the 33 German University Hospitals
- **MFT**
  - Association of the 37 German Medical Faculties
- **NSG**
  - PIs of the funded consortia
- **Working Groups**
  - Members of the funded consortia

**Working Groups**
- Interoperability
- Data Sharing
- Consent
- ...

[Diagram of the governance structure, showing connections between TMF, VUD, MFT, NSG, and Working Groups.]
Working groups foster harmonisation on a national level

- Interoperability
  - Core Dataset
  - Terminologies
  - Technologies (IHE, FHIR)

- Consent / eConsent
  - Modular, research friendly

- Use and Access Policies
  - UAC Committees
  - Registry of studies
The MIRACUM consortium

• Medical Informatics in Research and Care in University Medicine
  – eight University Hospitals and Medical Faculties, two Universities of Applied Science, and one industrial partner
  – across five German States
  – associated with four German Health Research Networks

• comprising ¼ of all German University Hospitals
• handling clinical and research data of more than 10 Mio patients
### Usecases of MIRACUM

<table>
<thead>
<tr>
<th>Alerting in Care</th>
<th>From Data to Knowledge</th>
<th>From Knowledge to Action</th>
</tr>
</thead>
</table>
| • IT Support for Patient Recruitment  
  • Medical Informatics / Clinical Trials | • Clinico-molecular Predictive Knowledge Tool  
  • Biostatistics / Prediction Models / Med. Inf. | • Support for Molecular Tumor Boards  
  • Bioinformatics / Medical Informatics / Precision Medicine |
The MIRACUM ecosystem

- Medical Informatics ReusAble eCosystem of Open source Linkable and Interoperable software tools – X
  - MIRACOLIX

- MIRACOLIX is
  - pragmatic
  - modular
  - reusable
  - open source
  - interoperable
  - federated
Architecture of MIRACUM
Data Integration Centers

Clinical Module

- Reporting (including identifiable) for routine care, quality mgmt, controlling
  - Routine business intelligence

ID-Management

- Maintains structured participant consent information
  - Consent-management

Research Module

- Queries for cohort identification/feasibility studies aggregated analyses; anonymized or pseudonymized microdata
  - Research queries

Research data repository

- Maintains local master patient index; generates primary as well as project-specific pseudonyms
  - Local ID-Management

Clinical data repository

- Source system 1
- Source system 2
- Comm-Server

Clinical Decision Support
- Source system 1
- Source system 2
- Comm-Server
Example: Feasibility

- 3,3 Mio Patients
- 30 Mio Diagnoses
- 23 Mio Procedures


Summary

- The Medical Informatics Initiative (MI-I)
  - towards a German Health Data Network
  - Data integration centers and nationally harmonized policies
- MIRACUM is the largest consortium
  - 1/3 of all university hospitals
  - Strongly committed to open source tools
  - MIRACOLIX infrastructure
  - i2b2 & OMOP appliances provided to all partner sites and other consortia
- Harvest “low hanging fruits” by queries on the reduced basic core dataset
- Full basic core dataset planned for end of 2018
  - Finalization of the specification
  - Mapping of data (especially not annotated data)
  - National SNOMED license (not really) on the way
- Advancing MIRACOLIX with OMOP
  - Privacy preserving distributed computing
  - “omics” data and imaging data integration
Thank you!
Device Safety Data Network

Daniel Prieto
Associate Professor & NIHR Clinician Scientist
Oxford University, UK
DEVICE SAFETY DATA NETWORK

Dani Prieto-Alhambra, MD MSc(Oxf) PhD
Associate Professor, Theme Lead for Observational Research
Centre for Statistics in Medicine, NDORMS, University of Oxford
Scope

• To study the use and risk/benefit of medical devices, when used in the wider community (‘as far as possible’ from RCT settings/participants)

• To develop and validate clinical and/or classification tools/algorithms for outcome prediction

• To improve the existing methods for the analysis of routinely collected data for the purposes above
Scope

• To study the use and risk/benefit of medical devices, when used in the wider community (‘as far as possible’ from RCT settings/participants)

• To develop and validate clinical and/or classification tools/algorithms for outcome prediction

• To improve the existing methods for the analysis of routinely collected data for the purposes above
Scope

• To study the use and risk/benefit of medical devices, when used in the wider community (‘as far as possible’ from RCT settings/participants)

• To develop and validate clinical and/or classification tools/algorithms for outcome prediction

• To improve the existing methods for the analysis of routinely collected data for the purposes above
REMIT: Where does this fit?
REMIT: Where does this fit?

- Device invention/development
  - First-in-man trials, etc.
- Engineering
- Post-marketing surveillance
- 'Routine' / Real world studies
  - Hypothesis testing
    - Confirming RCT safety signals
    - Post-marketing safety vigilance
    - Device utilisation studies
    - ...
REMIT: Where does this fit?

HYPOTHESIS GENERATION
- New indications for old treatments
- Effect/s of off-label use
- (Adaptive pathways)
- ...

HYPOTHESIS TESTING
- Confirming RCT safety signals
- Post-marketing safety vigilance
- Device utilisation studies
- ...

ENGINEERING

FIRST-IN-MAN TRIALS, ETC

DEVICE INVENTION/DEVELOPMENT

‘ROUTINE’ / REAL WORLD STUDIES

POST-MARKETING SURVEILLANCE
EXPERTISE REQUIREMENTS

1. Pharmaco- and *Device Epidemiology*
2. Causal Inference Methods
MOTIVATION: The ‘KNOWN UNKNOWNS’

KNOWN UNKNOWNS
- EFFECTIVENESS (in elderly, multi-morbid..)
- SERIOUS (RARE) SAFETY ISSUES (MoM, AFFs..)

KNOWN UNKNOWNS
- CLINICAL EFFICACY +
- SOME SAFETY DATA ++
ORTHOPAEDIC DEVICES
(an example)
RECENT OUTPUTS
(and data used)
Outputs (1) – UK HES (hospital inpatient)
Equity: THR - Need vs Provision of care

This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown.

The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study

Bayliss L et al. Lancet 2017
Outputs (3) – Catalan Registry + linked GP EMR (SIDIAP)
Early Surrogate/s of Implant Failure

Malak T et al. [Unpublished]
Outputs (4): UK CPRD linked to NJR Data
Safety: MoM THR and Cancer Risk

<table>
<thead>
<tr>
<th>Compared with no THR</th>
<th>THR patients</th>
<th>Control patients</th>
<th>Adjusted relative rate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n cases</td>
<td>n cases</td>
<td>rate</td>
</tr>
<tr>
<td>All total hip replacements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any cancer</td>
<td>75</td>
<td>611</td>
<td>1.33 (0.74, 0.95)</td>
</tr>
<tr>
<td>By bearing surface type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stemmed metal-on-metal</td>
<td>3</td>
<td>29</td>
<td>0.84 (0.80, 0.23-2.76)</td>
</tr>
<tr>
<td>Resurfacing</td>
<td>1</td>
<td>17</td>
<td>0.26 (0.39, 0.05-3.12)</td>
</tr>
<tr>
<td>Other bearing surfaces</td>
<td>71</td>
<td>565</td>
<td>0.18 (0.08, 0.59-0.96)</td>
</tr>
<tr>
<td>Haematological cancer</td>
<td>10</td>
<td>39</td>
<td>0.12 (0.41, 0.67-2.98)</td>
</tr>
<tr>
<td>Malignant melanomas</td>
<td>1</td>
<td>19</td>
<td>0.02 (0.26, 0.03-1.99)</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>12</td>
<td>62</td>
<td>0.21 (1.14, 0.59-2.20)</td>
</tr>
<tr>
<td>Renal cancer</td>
<td>5</td>
<td>39</td>
<td>0.09 (0.82, 0.31-2.16)</td>
</tr>
<tr>
<td>Other cancer</td>
<td>47</td>
<td>452</td>
<td>0.83 (0.63, 0.46-0.86)</td>
</tr>
</tbody>
</table>

CPRD/HES = non-THR controls (counter-factual)
Outputs (5) – CPRD linked to HES

Health Economics

E Burn et al. BMJ Open 2017
OHDSI - TAKING THIS.. TO ANOTHER LEVEL
’Real World’ Data NEEDS

- European **Primary Care EMR**
  - Linked to Hospital Records
  - Lined to Mortality Registry/ies
- European **National Joint (Arthroplasty) Registry**
  - Linked to Hospital Records
  - Linked to Patient Reported Outcomes
- European **National Audit/s**
  - National Hip Fracture Database
  - National Trauma Database
  - National Renal Registry
  - ...
EU-wide Device Registries linked to patient EMR data ...

THE EU: BREADTH OF DATA FOR RESEARCH

Most EU countries

Drug Utilization (pharmacy dispensations)

Most EU countries (not always available for linkage/research)

Primary Care Records

REGISTRIES (Devices, Biologic drugs, ...)

UK, DK, SWE, NORWAY, ETC...

HOSPITAL ADMISSIONS

HOSPITAL OUTPATIENTS

UK (HES), Spain (CMBD), DK, SWE, ...

DK, SWE, NL...

PROMS

MORTALITY REGISTRY

MOST EU (not always available for linkage/research)

THE DREAM (OF A DEVICE EPIDEMIOLOGIST) ...

WITH A COMMON DATA MODEL 😊
ONGOING WORK, AND NEXT STEPS
 Uni vs Total Knee Replacement in patients with multiple co-morbidities: UTMOST

• UK NIHR funded project started Nov/2017, running for 2y

• Aim/s
  – To replicate an on going RCT comparing UKR vs TKR in patients eligible for the RCT (ASA Grade <3) using observational (NJR) data and methods
  – To use ‘validated’ methods to study the risk-benefit and comparative costs of UKR vs TKR in patients with high co-morbidity, not eligible for RCT (ASA 3+)
EHDEN & ERC Grant Applications
(Awaiting Outcome) – AIM/s

• To create a framework for post-marketing device surveillance research, including:

1. Mapped EU-wide EMR linked to national/local Device Registry/ies, all in the OMOP CDM
2. Federated network of data custodians/analysts willing to contribute to such research
3. Validated methods/analytical tools for comparative device risk-benefit and HE
THANK YOU!

Dani Prieto-Alhambra, MD MSc(Oxf) PhD
Associate Professor, Theme Lead for Observational Research
Centre for Statistics in Medicine, NDORMS, University of Oxford
Adopting OMOP in IQVIA’s scaled ecosystem

Dr. Benjamin Hughes, PhD, MBA, MRes, MSc
SVP, Global Head of Strategy & Technology
IQVIA Real World & Analytics Services (RWAS)
IQVIA’s OHDSI context (1/2)

1-2bn+ patients*

100k+ sites

400k+ HCPs

~100 EHR systems

~20 countries

Pharma R&D

Pharma RWE

Regulators (FDA, EMA)

Public Health (CDC...)

Provider Sites & HCPs

Global network for various organization’s critical decisions

x10 next biggest organization’s data ecosystem

*Fully anonymized patient level data; range indicates unique active records vs total processed
### IQVIA’s OHDSI journey

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
<th># IQVIA DBs In OMOP</th>
<th># trainings / hacks given</th>
<th># Vocab. submitted</th>
<th># network partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>Workshop’s with 6 clients on OMOP</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2015</td>
<td>CEO OMOP investment approval Christian Reich joins IQVIA</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2016</td>
<td>Build up of IQVIA OMOP team Deeping relationship with Odysseus</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2017</td>
<td>OHDSI an independent product line FDA BEST network</td>
<td>12</td>
<td>13</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2018</td>
<td>Expansion of network efforts Scaling of software investment</td>
<td>15</td>
<td>27</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

~$10m IQVIA investment in OMOP & OHDSI over 4 years
Learnings “OMOPing” 12+ datasets

<table>
<thead>
<tr>
<th>Different Healthcare Systems</th>
<th>Vocabularies</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Payers</td>
<td>– Languages</td>
</tr>
<tr>
<td>– Healthcare institutions</td>
<td>– Drugs: only 30-50% overlap among the 100-300k products</td>
</tr>
<tr>
<td>– Specialties</td>
<td>– Conditions: ICD10≠ICD10</td>
</tr>
<tr>
<td>– Visits</td>
<td>– Procedures: Wild west</td>
</tr>
<tr>
<td>– Cost</td>
<td>– Lab tests: Wild west</td>
</tr>
<tr>
<td>– Referrals</td>
<td>– Units: traditional vs. SI</td>
</tr>
<tr>
<td>– Prescriptions (days supply vs fixed packages)</td>
<td></td>
</tr>
</tbody>
</table>

“OMOPing” iterative and intensive work, with huge potential for analytics industrialization
FDA Best & Network development

50 studies:
- Simple queries
- Complex studies (e.g., Safety)
- AE Reporting

Coordination

Develop新
methods

Develop Studies

Run Studies

Hospital Master
EMR
P+ Health plan Claims
Non-adjudicated claims

Text / NLP

Exciting developments in Surveillance & networks in general
## IQVIA OMOP Status – Overall & EU

<table>
<thead>
<tr>
<th>Country</th>
<th>Type</th>
<th>Patients*</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRA</td>
<td>IQVIA Ambulatory &amp; Specialty EHR</td>
<td>14 M</td>
</tr>
<tr>
<td>GER</td>
<td>IQVIA Ambulatory &amp; Specialty EHR</td>
<td>32 M</td>
</tr>
<tr>
<td>UK</td>
<td>IQVIA Ambulatory EHR</td>
<td>16 M</td>
</tr>
<tr>
<td>BEL</td>
<td>IQVIA Ambulatory EHR</td>
<td>1.1 M</td>
</tr>
<tr>
<td>ITA</td>
<td>IQVIA Ambulatory EHR</td>
<td>2 M</td>
</tr>
<tr>
<td>CAN</td>
<td>IQVIA Ambulatory EHR</td>
<td>1 M</td>
</tr>
<tr>
<td>AUS</td>
<td>IQVIA Ambulatory EHR</td>
<td>4 M</td>
</tr>
<tr>
<td>Global</td>
<td>IQVIA Oncology Survey</td>
<td>n/a</td>
</tr>
<tr>
<td>US</td>
<td>IQVIA Ambulatory EHR</td>
<td>38 M</td>
</tr>
<tr>
<td>US</td>
<td>IQVIA P+ Health claims</td>
<td>142 M</td>
</tr>
<tr>
<td>US</td>
<td>IQVIA Oncology EHR</td>
<td>140 K</td>
</tr>
<tr>
<td>US</td>
<td>MMI Specialty EHR Enhanced</td>
<td>350 K</td>
</tr>
<tr>
<td>US</td>
<td>IQVIA Hospital Charge Master</td>
<td>16 M</td>
</tr>
<tr>
<td>US</td>
<td>IQVIA Open Source claims</td>
<td>243 M</td>
</tr>
</tbody>
</table>

Next effort in EU (or CPRD v5? Or other client ask?)

Also engaging EU actors (e.g., EMIF / EHDN) on collaborations

*Fully anonymized patient level data
Balanced view of where we are

Glass half full

Clients: Growing # of engaged & happy
Partnerships: New models & partners
Collaboration: community participation
Long-term: industrialization of analytics

Glass half empty

Cost: $100m+ for IQVIA to deploy fully
Adoption: difficult for most
Security & Privacy: not @IQVIA level
Networks: shared economic model

Optimistic on OMOP but with challenges to address
IQVIA path forward

**Definitely**
More IQVIA data conversions
Further vocab releases
OHDSI / Public interest scientific studies
Partner outreach for networks (~30 conversations)
Software investments to reduce OMOP costs

**Maybe**
Commercial research / consortium provider & pharma
Data nodes for general access / network studies
Privacy software for networks
IQVIA Oncology network inclusion
IQVIA ML algorithms on OMOP

Ongoing investment OMOP and iterative & collaborative approach
Adoption of the OMOP-CDM: Industry perspective of using the OMOP-CDM

23-March-2018

Erica A. Voss, MPH, PMP $^{1,2,3}$

$^1$Janssen Research and Development, Raritan, NJ, USA
$^2$Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA
$^3$Erasmus University Medical Center, Rotterdam, Netherlands
Who We Are

• Mission to help people everywhere live longer, healthier, happier lives.

• 3 Sectors:
  – Consumer Products
  – Medical Devices
  – Pharmaceutical Products

• Epidemiology’s focus: generating real world evidence that will benefit patients and consumers who use our company’s products every day
Opportunities for Real World Evidence across Research & Development

- Disease natural history
- Treatment pathways
- Safety Surveillance
- Comparative effectiveness
- Prediction Modeling
Many questions types with global products requires **lots of different types of data**

In total, Janssen Epidemiology Analytics conducts research on more than 400+ million patient records from across US, Europe and Asia-Pacific.
Strategy to Achieve This?

Standardized Data

Standardized Analysis

Source 1

Source 2

Source 3

Transformation to OMOP common data model

Analysis method

Analysis results
Use Case: Cohort Identification

Using OHDSI tools to conduct clinical trial feasibility

Rupa Makadia, MS¹,², Jamie B. Forlenza¹, PharmD, MS¹, Frank J. DeFalco¹,², Chris Knoll¹,², Patrick B. Ryan, PhD¹,²,³

¹Janssen Research & Development, LLC, Titusville, NJ ²OHDSI collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY ³Columbia University, New York, NY

1. Protocol evaluated
2. Concept sets created in ATLAS
3. Inclusion rules created in ATLAS
4. Results generated in ATLAS on databases
5. Results shared with clinical teams

ClinicalTrials.gov
An Efficacy and Safety Study of Sirukumab in Participants With Major Depressive Disorder
This study is currently recruiting participants. (see Currents and Locations)

Population Visualization
68721 people (21.54%), 9 criteria passed. 1 criteria failed.
3.3 | Comparative analysis

Of the 72,797 users of canagliflozin and 225,627 users of non-SGLT2 inhibitor AHAs with no history of BKLE amputation and 
≥1 day at risk who were eligible for inclusion in the comparative analysis, 63,845 pairs were formed based on matching of EPS (Figure 1).

All baseline characteristics were well balanced after EPS matching (Figure 2), and patient demographics (age and sex), key comorbid conditions (including CV disease) and medications of interest (eg, commonly reported in patients with T2DM) for the treatment cohorts are presented in Table 2. The median (interquartile range [IQR]) duration of index therapy was 0.43 (0.17, 0.94) years with canagliflozin.

**Figure 1** Attribution diagram for the comparative analysis

**Figure 2** Diagnostics of EPS matching performance
Use Case: Patient-Level Prediction

Risk Prediction for Ischemic Stroke and Transient Ischemic Attack in Patients Without Atrial Fibrillation: A Retrospective Cohort Study

Zhong Yuan MD, PhD*, & PhD, Daniel Yannicelli MD:

<table>
<thead>
<tr>
<th>CCAE</th>
<th>OPTUM</th>
<th>MDCR</th>
</tr>
</thead>
</table>

Comparing AUCs: CCAE - STROKE & 365D

Comparing AUCs: OPTUM - STROKE & 365D

Comparing AUCs: MDCR - STROKE & 365D

Comparing AUCs: CCAE - STROKE_TIA & 365D

Comparing AUCs: OPTUM - STROKE_TIA & 365D

Comparing AUCs: MDCR - STROKE_TIA & 365D
OMOP CDM and OHDSI are Critical to our Success

• Real world data:
  – increasingly important part of evidence generation
  – working with data takes substantial investment

• Our experience investing in OMOP / OHDSI has a strong return; **the more we invest the more we get out of it**
A Mental Health Data Platform

Simon Lovestone
Professor of Translational Neuroscience
Oxford University, UK
Clinical Records Interactive Search (CRIS)

• Mental Health and Dementia clinical services in the UK and the benefits of being Cinderella!

• Generating access for research

• OMOP and future directions
Clinical Records Interactive Search (CRIS)

- SLaM CRIS
  - South London and Maudsley NHS BRC implementation
- D-CRIS
  - Cambridge & Peterborough, Oxford Health, West London, Camden and Islington
  - 1 million plus patients
- UK-CRIS
  - 10 site extension
  - Connectivity to UK BioBank

Mike Denis and Simon Lovestone
CRIS – core functionality

EHR Data Source → De-identification → Processing pipeline → CRIS front end → CRIS SQL
CRIS Security

Managed by Stakeholder-led oversight committee

Project → Audit log → CRIS → Output → Findings

Reseacher → Trust contract

RIO CareNotes → Pseudonymisation

Firewall
Deriving information from data

FAST enables **information retrieval**, i.e. search and retrieval by matching against user defined strings;

GATE enables **information extraction**, i.e. extracts ‘meaning’ (structure) from free text context

Neural Networks for **advanced and automated** search and extraction for data exploration
EHR-data re-use for research
(Cholinesterase inhibitors and Alzheimer’s disease)

- Phase IV of AChEI
  > 2500 patient years of therapy
  > 8 fold dataset compared to Cochrane

- Costs and effectiveness
  precompetitive collaboration with pharma
  Text mining derivation of service utilisation and costs

- Predictors of response
  Biomarkers and clinical

Data from Robert Stewart, KCL
Real world data reutilisation for research

• EMIF publications
  – [http://www.emif.eu/results](http://www.emif.eu/results)
    • or search EMIF EU references
  – > 85 papers 2012-2017

• CRIS publications
    • or search CRIS BRC references
  – > 65 papers 2009-2017
National Clinical Records Interactive Search (UK- CRIS)

14
NHS Mental Health Trusts across the UK

2.5m+
De-identified electronic patient records

- Avon and Wiltshire Mental Health Partnership NHS Trust
- Cambridgeshire and Peterborough NHS Foundation Trust
- Camden and Islington NHS Foundation Trust
- Devon Partnership NHS Trust
- Kent and Medway NHS and Social Care Partnership Trust
- Mersey Care NHS Foundation Trust
- North East London Foundation Trust
- Nottinghamshire Healthcare NHS Foundation Trust
- Northumberland, Tyne and Wear NHS Foundation Trust
- Oxford Health NHS Foundation Trust
- Southern Health NHS Foundation Trust
- South London and Maudsley NHS Foundation Trust
- South West London and St George’s NHS Foundation Trust
- West London Mental Health NHS Trust
UK-CRIS: safe, secure and complete

Trust network
- Input files
  - EHR Backup and Data Dictionary
  - NHS Trust automated extraction
- Output files
  - Extracted NHS Trust database
  - Data Dictionary

NHS computer access

Web

SQL

UK-CRIS datacentre and secure network

Output files sent via secure FTP to UK-CRIS. Data is encrypted

N3 Network

Secure VPN tunnel

sFTP client

sFTP server

Elastic index (local)

Elastic index (OMOP)

PostgreSQL (local)

PostgreSQL (OMOP)

Data Dictionary

EHR Backup and Data Dictionary

Data de-identified and uploaded

NHS Trust secure VPN tunnel

Private network
- Access outside Trust (Authorised staff)

Web

SQL

NHS Trust secure VPN tunnel
UK-CRIS: safe, secure and complete

Trust network
- Input files
  - EHR Backup and Data Dictionary
- Output files
  - Extracted NHS Trust database
  - NHS Trust automated extraction
  - Data Dictionary
- NHS computer access
- Audit log
- Trust contract

UK-CRIS datacentre and secure network
- UK-CRIS receives encrypted files.
  - Data de-identified and uploaded
  - Elastic index (local)
  - Elastic index (OMOP)
  - PostgreSQL (local)
  - PostgreSQL (OMOP)

Private network
- Access outside Trust (Authorised staff)
- Web
- SQL

Managed by Stakeholder-led oversight committee
Federated access studies

1. Project idea conceived
2. Project proposal created at lead Trust
3. Project proposal reviewed
4. Project proposal reviewed (Trust A)
5. Project proposal reviewed (Trust C)
6. Data Analysed

Project proposal reviewed
Local oversight committee approve or reject the project proposal (Trust B)

Project proposal reviewed
Local oversight committee approve or reject the project proposal (Trust C)

Search within de-identified data

CRIS DB (Trust A)
CRIS DB (Trust C)

Audit process (Trust A)
Audit process (Trust C)

Data Analysed

NO
YES

YES
NO

Project proposal submitted to additional Trusts for review

NO
YES
UK Biobank is a national and international health resource with unparalleled research opportunities, open to all bona fide health researchers. UK Biobank aims to improve the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia. It is following the health and well-being of 500,000 volunteer participants and provides health information, which does not identify them, to approved researchers in the UK and overseas, from academia and industry. Scientists, please ensure you read the background materials before registering. To our participants, we say thank you for supporting this important resource to improve health. Without you, none of the research featured on this website would be possible.

Read more about Biobank UK
UK Biobank – Enhancements

• Web-based questionnaires for additional exposures and outcomes (cognition, mental health, occupation..)
• Wrist-worn accelerometers mailed to 100,000 participants to measure physical activity
• Multimodal imaging in 100,000
• Repeat Neuroimaging in 10,000
• Genotyping of all participants (820,000 SNPs)
• Repeat cognition, sampling
• Connectivity to EMRs for mental health
Proud to be wearing #UKBioBank movement measuring device.
Volunteer research to prevent dementia and other diseases.
UK Biobank Linkage

EHR

Trusts (Tn)

Secure data import pipeline

Generate MD5 #

PostgreSQL storage

Elastic search indexes

CRIS Trust 1

OMOP Trust n

Federation

UK Biobank data access application

Known to Biobank?

Access review by Subcommittee

Web service

Dataset generated and made available

EHR

UK Biobank

Federated search

Biobank Key Exchange

Project specific link pseudonyms exchanged

UK-CRIS search results with link pseudonyms

UK-CRIS and UKB data merged

UK Biobank data with link pseudonyms

Dataset generated and made available

EHR

UK Biobank

Dedicated analysis area

Study Publication

Federated search

Biobank Key Exchange

Project specific link pseudonyms exchanged
UK- CRIS to UK-biobank

- First ten NHS Trusts linked 2017
- >15,000 shared individuals

<table>
<thead>
<tr>
<th>NHS Trust coded diagnosis</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient numbers</td>
<td>131041</td>
<td>125239</td>
<td>238018</td>
<td>288285</td>
<td>210021</td>
<td>249117</td>
<td>110023</td>
<td>155877</td>
<td>193418</td>
<td>179389</td>
<td>330000</td>
<td>2210428</td>
</tr>
<tr>
<td>UK Biobank match</td>
<td>731</td>
<td>152</td>
<td>1,053</td>
<td>3,964</td>
<td>40</td>
<td>2,894</td>
<td>63</td>
<td>62</td>
<td>1,528</td>
<td>1,450</td>
<td>3,161</td>
<td>15,098</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Percentage</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>8%</td>
<td>56.9</td>
<td>11.8</td>
<td>82.0</td>
<td>308.8</td>
<td>1.8</td>
<td>225.4</td>
<td>4.9</td>
<td>4.8</td>
<td>119.0</td>
<td>112.9</td>
<td>246.2</td>
<td>1175</td>
</tr>
<tr>
<td>Bipolar</td>
<td>9%</td>
<td>66.1</td>
<td>13.7</td>
<td>95.2</td>
<td>358.6</td>
<td>3.6</td>
<td>261.8</td>
<td>5.7</td>
<td>5.6</td>
<td>138.2</td>
<td>131.2</td>
<td>285.9</td>
<td>1366</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>5%</td>
<td>33.1</td>
<td>6.9</td>
<td>47.6</td>
<td>179.3</td>
<td>19.0</td>
<td>130.9</td>
<td>2.8</td>
<td>2.8</td>
<td>69.1</td>
<td>65.6</td>
<td>143.0</td>
<td>700</td>
</tr>
<tr>
<td>Cognitive disorder</td>
<td>41%</td>
<td>303.1</td>
<td>63.0</td>
<td>436.5</td>
<td>1643.4</td>
<td>16.6</td>
<td>1199.8</td>
<td>26.1</td>
<td>25.7</td>
<td>633.5</td>
<td>601.1</td>
<td>1310.5</td>
<td>6259</td>
</tr>
<tr>
<td>Depression</td>
<td>22%</td>
<td>161.6</td>
<td>33.6</td>
<td>232.8</td>
<td>876.5</td>
<td>8.8</td>
<td>639.9</td>
<td>5.6</td>
<td>13.7</td>
<td>337.8</td>
<td>320.6</td>
<td>698.9</td>
<td>3330</td>
</tr>
<tr>
<td>Other</td>
<td>15%</td>
<td>110.2</td>
<td>22.9</td>
<td>158.7</td>
<td>597.6</td>
<td>6.0</td>
<td>436.3</td>
<td>9.5</td>
<td>9.3</td>
<td>230.4</td>
<td>218.6</td>
<td>476.5</td>
<td>2276</td>
</tr>
</tbody>
</table>


Mapping to the OMOP CDM

UK-CRIS Environment

- CRIS Trust 1 (T1)
- CRIS Trust 2 (T2)
- CRIS Trust n (Tn)
- OMOP T1
- OMOP T2
- OMOP Tn
- PostgreSQL storage (for SQL client access)
- Elastic search engine (for web client access)
- FEDERATION

Legend:
- Feature/Function
- PID
- Pseudonymised
- Anonymised

Local Search

T1 study result store Trust 1 (T1)

EHR
- Trust (T1)
- Trust 2 (T2)
- Trust n (Tn)

Data is extracted from Trust databases
Data deidentified and uploaded

Secure data import pipeline
• Generate pathways from complete GWAS datasets and perform clustering analysis for shared pathways

• Correlate pathway load per disease with risk relationship between disease and Alzheimer’s *from real world data*

• Perform proof of concept in human samples *using EMIF catalogue*

Nevado-Holgado et al (2017) [https://www.biorxiv.org/content/early/2017/08/21/179267](https://www.biorxiv.org/content/early/2017/08/21/179267)
Participant identification and recruitment

Readiness cohort

EPAD Registry

Integration

Cohort A

Cohort B

Cohort C

Selection criteria

Trial cohort

EPAD Registry

Selection criteria

Readiness cohort

Trial cohort

Placebo

Rx 1

Rx 2

Rx ... n

Adaptation by change in intermediate phenotype

Adaptation on cognition outcomes
**Post-licensing / ‘Phase IV’**

*Potential to deliver post-marketing data at scale and in real-world contexts*

- South London & Maudsley NHS FT **CRIS** data
- n=2460; dementia treatment with AChEIs
- MMSE derived from **coded and uncoded** data
- Improvement by 4.2 units per year in first 6 months
- Predictors of response:
  - Better early response in non-white patients
  - Worse early response in vascular dementia

Beyond research data modeling - using CRIS for real-world, individual level costs

- Destination from uncoded data and linkage
- N=3075 (5624 6m windows)
  - 25% alone; 52% ADL problems
  - 37% physical illness; 45% moderate severity
- Mean costs of severe dementia >2x that for mild
- Increased care costs associated:
  with severity, functional problems, agitation, living alone
  but not physical illness, depression or gender

CRIS consent for contact model

Assessment of capacity and consent for re-contact

Patient record re-identified in EMR

CRIS ID matched to EMR ID by Trusted Third Party

Researcher identifies eligible patient

Researcher contacts patient and informs clinical team

CRIS consent for contact model

74% agreement

20,000 consents and samples in 3 years

Next steps

• UK-CRIS expansion, renewal and sustainability
• Linkage
  – Other large scale / population cohorts
  – Primary care
  – Secondary acute care
  – Improving Access to Psychological Therapies (IAPT)
• Using UK-CRIS
  – for target identification, drug repurposing and understanding risk of disease
  – for trials platforms
  – to improve health care (algorithms and decision support)
• An ecosystem for patient empowerment
  – Patient health records / medical records
  – Connected devices for PROMs
Acknowledgements

• SLaM team: Robert Stewart and Matthew Broadbent
• UK-CRIS team: Mike Denis and David Newton
• OxCRIS: Tanya Smith and Alejo Nevado

• UK CRIS network

• Funders: NIHR and MRC
There are two types of people in this world:

Those who can extrapolate from incomplete data.
13:00 – 14:45 OHDSI Collaborator Showcase
A success story:
The adoption of the OMOP-CDM in South Korea

Rae Woong Park
Ajou University School of Medicine,
South Korea
SLIDES TO ADD
Coffee Break