

Meeting Report

Observational Medical Outcome Partnership (OMOP) Special Interest Group: Inaugural Meeting

Introduction

The inaugural meeting of the UK Health Data Research Alliance OMOP (Observational Medical Outcome Partnership) Special Interest Group was held on Zoom on Tuesday, 27th June.

The Alliance convened this Special Interest Group devoted to the OMOP Common Data Model (CDM)¹ to bring together data user communities, including those who are transforming data to OMOP and those who are using transformed data for research. We hope that the group will provide a forum to share knowledge, experience, and tools. We are also working in collaboration with related groups, such as the [Observational Health Data Sciences and Informatics UK Node](#) and the [European Health Data Evidence Network \(EHDEN\)](#).

The meeting was co-chaired by [Geoff Hall](#) (Professor of Digital Health, University of Leeds, and Chief Clinical Data Officer, HDR UK) and [Dani Prieto Alhambra](#) (Professor of Pharmaco- and Device Epidemiology, University of Oxford).

Presentations

Geoff Hall: Introduction to OMOP

[\[link to presentation\]](#)

Geoff opened the meeting and highlighted the importance of bringing communities together to share knowledge around adoption of standards, including the OMOP CDM. He suggested that this group could help drive alignment and reduce duplication of effort, as well as provide a forum to share experiences about OMOP transformation.

Dani Prieto-Alhambra: Local data, global evidence: federated learning

[\[link to presentation\]](#)

Dani stated that collaboration is key to get all the expertise, diverse data, registries, and electronic records to cross fertilise, share codes, and interact with partners. Federated learning using OMOP is significant for transparency, to maximise interoperability, reproducibility, and trustworthiness. He also spoke about the need to transition from the EHDEN project, which is finishing, to sustainable funding for OMOP activities.

¹ The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) is an open community data standard, designed to standardize the structure and content of observational data such as Electronic Health Records (EHR). It is developed and maintained by the [Observational Health Data Sciences and Informatics \(OHDSI¹\)](#) programme.

Erica Voss (Johnson & Johnson): Large-scale studies using OMOP: Characterizing adverse events in COVID-19 infected patients

[<https://linkinghub.elsevier.com/retrieve/pii/S2589537023001098>]

[[link to presentation](#)]

Erica presented a multinational retrospective study: “Contextualising adverse events of special interest to characterise the baseline incidents rates in 24 million patients with Covid-19 across 26 databases”. This large network study was conducted utilising data from OHDSI, looking at the serious adverse events of special interest for COVID-19 vaccines and COVID-19 infections in eight countries.

John Birkinshaw and Kuldeep Sohal (Connected Bradford) Case Study 1: Connected Bradford

[[link to presentation](#)]

John and Kuldeep gave a presentation on Connected Bradford, a platform hosting linked routine electronic data across primary, secondary and social care data from over 700,000 individuals from the Bradford population, hosted by Bradford hospitals.

Suzy Gallier (PIONEER) Case Study 2: DECOVID Project

Suzy’s case study presented the DECOVID project: “De-coding COVID-19”, which focussed on how to utilise and access highly granular real time data on COVID from two digitally mature NHS Trusts, to provide information for research and reduce burden on frontline staff.

Daniel Dedman (MHRA) Case Study 3: CPRD

[[link to presentation](#)]

Dan presented the OMOP conversion of the CPRD data set (Clinical Practice Research Datalink) which contains over 60 million individuals and has been widely used for over 30 years. With a large database, OMOP has been invaluable to visualise these data sets, develop analytics, and reduce barriers to observational research.

Discussion

The presentations were followed by a lively Q&A and discussion. Some key points are summarised below.

Longevity/Popularity of OMOP

- The audience highlighted the existence of various standards (e.g. i2b2) and the uncertainty around which one to use.
- It was acknowledged that the OMOP common data model is the model currently being adopted across Europe and this has driven adoption across European organisations. It is also one of the most used common data models in the United States and will be increasingly adopted across NHS settings in the UK, as Secure Data Environments are developed.

- In addition, there are multiple initiatives supporting and funding the use of OMOP internationally, including many EU-funded projects (Horizon Europe, Innovative Medicine Initiative/Innovative Health Initiative), industry-funded ones, and regulatory (see www.darwin-eu.org).
- However, there is a need to explore sustainable models nationally in the UK.

Sharing the load

- It is helpful to know whether other organisations have already done something similar, used certain tools or what data have been mapped. Re-using protocols and tools that others have successfully used is important to reduce duplication. The EHDEN and OHDSI portals provide some information: <https://www.ehden.eu/datapartners/> or <https://www.ohdsi.org/who-we-are/collaborators/>. However, it is not easy to find information about what datasets have been mapped. See below for information on the OMOP slack channel and other HDR UK resources.
- It was commonly agreed that sharing of ETL code resources is crucial to enable researchers to collaborate and learn from each other and enable research that makes a difference.

NHS data

- Participants were interested to explore the complexities around mapping NHS data sets. There was discussion of the need for “NHS extensions” to OMOP and converting the NHS dictionary to OMOP. The need to share mappings around common Electronic Health Records systems such as EPIC was also discussed. There was emphasis on sharing effort and reaching out to the community for help, for example on the OHDSI forum.
- There was an agreement on the need to focus on the **national picture**. There is an advantage of a single language and more homogeneity in healthcare organisations. Development of an OHDSI UK node will offer an opportunity for sharing experiences across organisations and providing information around mappings for core national data sets. It was noted that there are also specific disease registries such as the national cancer data set which are being mapped into OMOP.

Variable mappings

- There were questions around what data can be mapped to OMOP and what the advantages are, and to what extent we can rely on results and knowledge generated by large scale ‘OMOPified’ data analysis as opposed to seeing this approach as a hypothesis generating strategy.
- It was suggested that the OMOP CDM contains “source codes” so even if not using use standardized terminologies, it is still possible to take advantage the improved analytic layout of the data while still using source data.
- There are many data sources that can be mapped to this model. Where data sources cannot be mapped, the community has worked together to generate updates to CDM to improve handling the data. The main challenges highlighted include finding the skilled people to

perform the mapping, having proper data governance set up, and working on mapping source codes to standard terminologies.

- The benefit gained from converting data is improved federated network-based research as well as leveraging standardised tools to perform characterisations, estimation studies, and prediction studies.
- Ethnicity was mentioned as a real challenge in mapping to OMOP as the CDM and vocabularies have a strong US bias, and ethnicity data varies by country. There is an opportunity for the community to think about how to improve on this.

Next steps

Most participants agreed that there is a need to bring together communities and organisations, particularly from across the UK, to share knowledge, tips and experiences around adoption of the OMOP common data model for data.

The Alliance secretariat will soon circulate details for a second Special Interest Group meeting to delve into some of the aspects highlighted to date, including the need for mapping the UK landscape to understand what data has already been mapped and the need for an accessible repository/resource to share documentation, protocols and outputs derived from use of OMOP data.

If you would like to suggest topics for discussions at the next meeting, please email ukalliance@hdrug.ac.uk.

A survey to explore adoption of OMOP across the UK is open until 28 July:

<https://www.surveymonkey.co.uk/r/BRG538V>. You are invited to participate, as we are keen to hear from people who do not use OMOP as well as those who do.

Useful links

- Athena: <https://athena.ohdsi.org/>
- European EHDS data partners: <https://www.ehden.eu/datapartners/>
- OHDSI YouTube content: <https://www.youtube.com/@OHDSI/playlists>. See the "Features" playlist as well as the "Tutorials & Workshops"
- OHDSI UK National Node <https://ohdsi-europe.org/index.php/national-nodes/uk>
- DARWIN EU Project: <https://darwin-eu.org/>
- OHDSI Forum: <https://forums.ohdsi.org/>
- OMOP Slack channel: email alex.knight@hdrug.ac.uk to be added
- HDR UK OMOP site on GitHub Pages: <https://hdrug.github.io/OMOP/>
- Alliance work on data standards: <https://ukhealthdata.org/projects/data-standards-and-quality/>
- Alliance survey on OMOP adoption: <https://www.surveymonkey.co.uk/r/BRG538V>