OHDSI & Clinical Registries: Sanity for Health Systems

OHDSI Community Call
Aug. 22, 2023 • 11 am ET

Upcoming Community Calls

Date	Topic
Aug. 22	OMOP Supporting Clinical Registries
Aug. 29	Vocabulary Release Update
Sept. 5	DARWIN EU® Progress and Roadmap
Sept. 12	OHDSI 2023 Global Symposium Conference & Activities Preview
Sept. 19	Journal Club: 11th Revision of the ENCePP Guide on Methodological Standards in Pharmacoepidemiology
Sept. 26	Publications Presentation
Oct. 3	Workgroup Reports, pt 1
Oct. 10	Workgroup Reports, pt 2
Oct. 17	Symposium Week! Final Logistics
Oct. 24	Welcome to OHDSI

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Madeline Adelman, Torunn Sivesind, Isaac Weber, Grace Bosma, Camille Hochheimer, Chante Karimkhani, Lisa Schilling, John Barbieri, Robert Dellavalle** on the publication of **Prescribing Patterns of Oral Antibiotics and Isotretinoin for Acne in a Colorado Hospital System: Retrospective Cohort Study** in *JMIR Dermatology*.

JMIR Publications
Advancing Digital Health & Open Science

Articles | Search articles

JMIR Dermatology | Journal Information | Browse Journal | Subm

Published on 21.8.2023 in Vol 6 (2023)

Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/42883, first published October 07, 2022.

Prescribing Patterns of Oral Antibiotics and Isotretinoin for Acne in a Colorado Hospital System: Retrospective Cohort Study

Madeline J Adelman [1]; Torunn E Sivesind [1]; Isaac Weber [2]; Grace Bosma [3]; Camille Hochheimer [3]; Chante Karimkhani [4]; Lisa M Schilling [5]; John S Barbieri [6]; Robert P Dellavalle [1, 7, 8]

Article | Authors | Cited by | Tweetations (2) | Metrics

- Abstract
- Introduction
- Methods
- Results
- Discussion
- References
- Abbreviations
- Copyright

Abstract

Background:
Guidelines established by the American Academy of Dermatology recommend oral antibiotics as first-line therapy for mild, moderate, and severe acne. However, it is recommended to minimize the duration of oral antibiotic use, and there is increasing support for other systemic agents for acne.

Objective:
We sought to characterize the use of oral antibiotics and isotretinoin for the treatment of acne in the pediatric and young adult population aged 10 through 20 years and the adult population aged 21 to 45 years from 2011 to 2019.

Methods:
We conducted a retrospective, observational cohort study using electronic data from the enterprise data warehouse of the University of Colorado Anschutz Medical Campus and its affiliates, with data in the format of the Observational Health Data Sciences and Informatics (OHDSI) Observational Medical Outcomes Partnership (OMOP) common data model. Categorical values (sex, race, and ethnicity) were compared using chi-square tests, and continuous variables (age) were compared

#OHDSISocialShowcase

ohdsi.org/europe2023-showcase

#OHDSISocialShowcase

2023 Europe Symposium Collaborator Showcase

1	The EHDEN Portal – Simplifying the access to OMOP CDM databases	João Rafael Almeida, Nigel Hughes, Peter Rijnbeek, José Luís Oliveira
2	Privacy-preserving using k-anonymity and l-diversity in OMOP CDM databases	João Rafael Almeida, José Luís Oliveira
3	The Dutch ICU Data Warehouse: towards a standardized multicenter electronic health record database	Ameet Jagesar, Martijn Otten, Tariq Dam, Laurens Biesheuvel, Patrick Thoral, Armand Girbes, Harm-Jan de Grooth, Paul Elbers
4	Community Contribution to the OHDSI Vocabularies, User-Level QA and a New Entity Mapping System SSSOM	Oleg Zhuk, Anna Ostropolets, Nicolas Matentzoglu, Melissa Haendel, Alexander Davydov, Christian Reich
5	Extract, Transform, and Load of the Infectious Disease CDM for Harmonizing and Accessing Data in Real-time Infectious Disease Surveillance	Byungjin Choi, Junhyuk Chang, Soobeen Seol, Seongwon Lee, Rae Woong Park
6	Roadmap and improvement of OHDSI Vocabularies	Christian Reich, Alexander Davydov, Anna Ostropolets
7	Integrating the OMOP CDM into the AI Sandbox of the German Health Data Lab	Elham Taghizadeh, Maxim Moinat
18	Hierarchical clustering of microbial resistance profiles and ventilation protocols using the oncology extension	Jared Houghtaling, Frederic Jung, Ankur Krishnan, Marc Padros Goossens, Frank Leus, Lauren Maxwell, Tom Feusels, Freija Descamps
19	Capture and consolidation of renal specific concepts into a cohesive OMOP dataset	Jared Houghtaling, Jose Antonio Ramírez García, Clémence Le Cornec, Lore Vermeylen, Nir Assaraf, Lars Halvorsen
20	Creation of a reusable OMOP transformation workflow for Belgian electronic health record systems	Jared Houghtaling, Lore Vermeylen, Louise Vandenbroucke, Korneel Bernaert, Brecht Dekeyser, Freija Descamps
21	Construction of a central ontology platform for semantic mapping coordination and vocabulary augmentation across a multi-partner oncology consortium	Jared Houghtaling, Peter Prinsen, Maaike van Swieten, Chiara Attanasio, Lars Halvorsen
22	Application of the R-CDM extension to capture metadata and features extracted from quantitative brain MRI and CT data	Jelle Praet, Jared Houghtaling, Frederic Jung, Steve De Backer, Jeroen Pinxten and Dirk Smeets
23	NNRD-AI: a national neonatal research database for rapid insights with machine learning and artificial intelligence	Julia Lanoue, Kayleigh Ougham, Neena Modi, Sam Greenbury
24	OMOP-CDM Data conversion for the Papageorgiou General Hospital in Greece	Achilleas Chytas, Maria Bigaki, Pantelis Natsiavas
25	Development of a GA4GH Beacon for structured Clinical Data Discovery using the OMOP-CDM	Alberto Labarga, Sergi Aguiló
26	Quality Management System of the OHDSI Standardized Vocabularies	Vlad Korsik, Anna Ostropolets, Christian Reich, Alexander Davydov

Open-source analytics development

46	CDMConnector: Cross platform OMOP CDM database queries using dplyr	Adam Black, Edward Burn, Artem Gorbachev, Martí Català
47	Development of an OMOP Ontology Application – PROSA – for creation and maintenance of highly granular source concepts within the OMOP vocabulary structure	Jared Houghtaling, Emma Gesquiere, and Lars Halvorsen
48	A method to facilitate rapid stand up of OMOP research tools from validated libraries for RWE research	Jack Brewster
49	Generating Synthetic Data from OMOP-CDM databases for Health Applications	Alberto Labarga, Sergi Aguiló
50	Performance Improvement of Post-ETL in OMOP CDM	Antonella Delmestri

Clinical applications

51	Drug utilisation of valproate-containing medicinal products in women of childbearing age: a network study part of DARWIN EU®	Albert Prats-Uribe, Martí Català, Katia M Verhamme, Maria de Ridder, Carlen Reyes, Talita Duarte-Salles, Peter Rijnbeek, Edward Burn, Daniel Prieto-Alhambra, Annika M. Jödicke

ohdsi.org/europe2023-showcase

#OHDSISocialShowcase

MONDAY

Measuring multimorbidity in IPCI: An analysis of more than 1.8 million patients

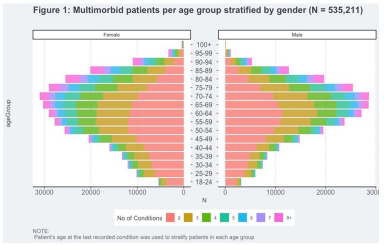
(Solomon Ioannou, Egill Fridgeirsson, Marcel de Wilde, Jan Kors, Peter Rijnbeek, Katia Verhamme)

Prevalence of **multimorbidity** in IPCI was estimated to be at **29.23%.**

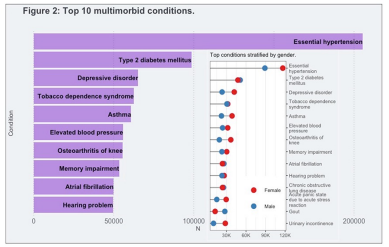
Quantifying multimorbidity in IPCI: An analysis of more than 1.8 million people.

Background: Multimorbidity, the presence of multiple chronic conditions in an individual, is a growing global health concern. Multimorbidity is associated with increased healthcare utilization, higher costs, and poorer health outcomes such as mortality and disability. Using the general practitioner database IPCI, and OHDSI's standardized analytical tools, we perform Association Rule Mining to study multimorbidity prevalence and the most common multimorbid conditions.

Result 1: Number of conditions for each age group stratified by gender.



Result 2: Top 10 multimorbid conditions stratified by gender.



Methods

1. We retrospectively extracted information from the Integrated Primary Care Information (IPCI) database, which contains electronic patient records from 350 General Practitioners in the Netherlands[1].
2. We define our cohort as "All patients in the database, above 18 years old, with an observation period starting any time before 31st December 2022 with at least one year of observation in the database."
3. We collected information based on a set of 75 conditions, which were identified by expert panels of GPs in Denmark using the ICPC-2 vocabulary[2]. We mapped those concepts to SNOMED resulting in a list of 151 distinct concept IDs and used them to quantify prevalence of multimorbidity. The conditions cover a wide range of categories that encompass various bodily systems and organs, incorporating prevalent chronic diseases such as cancer, hypertension, and dementia, as well as HIV, trauma, and lifestyle factors like tobacco and alcohol abuse.

References:

1. Maria A J de Ridder and others, Data Resource Profile: The Integrated Primary Care Information (IPCI) database, The Netherlands, *International Journal of Epidemiology*, Volume 51, Issue 6, December 2022, Pages e314–e323
2. N'Goran AA, Blaser J, Deruaz-Luyet A, Senn N, Frey P, Haller DM, Tandjung R, Zeller A, Burnand B, Herzig L. From chronic conditions to relevance in multimorbidity: a four-step study in family medicine. Fam Pract. 2016 Aug;33(4):439-44.

Limitation: To quantify multimorbidity, this study used the presence of concepts belonging to the definition of multimorbidity, instead of phenotype definitions. Additionally, the same person appearing in the study population more than once is a possibility.

Solomon Ioannou, Egill Fridgeirsson, Marcel de Wilde, Jan Kors, Peter Rijnbeek, Katia Verhamme

#OHDSISocialShowcase

TUESDAY

A Simple Standard for Sharing Ontological Mappings (SSSOM) Workshop for OHDSI Europe

(Nicolas Matentzoglu, Melissa Haendel)

Synergizing Simple Standard for Sharing Ontology Mappings (**SSSOM**) and the Observational Health Data Sciences **and** Informatics (**OHDSI**)

PRESENTERS:
Nicolas Matentzoglu
Polina Talapova

INTRO

The OHDSI community is always interested in improving the quality of mappings across vocabularies because they significantly influence the reliability of evidence during analysis.

SSSOM is a metadata standard for sharing semantic entity mappings. Entity mappings are crucial for data integration efforts such as OMOP and define the relationships between entities in vocabularies, including terminologies and ontologies. Many entity mappings currently produced are imprecise in some way; for example OMOP:44499396 ("Myelodysplastic syndrome with single lineage dysplasia") is mapped to OMOP:4003185 ("Refractory anemia") in the OMOP Vocabulary. When this mapping is applied, clinically relevant content ("with single lineage dysplasia") is "lost".

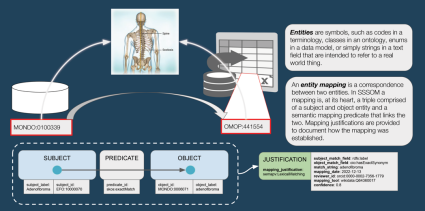
The implementation of SSSOM-derived mapping metadata into the operational model of the OMOP Standardized Vocabularies has recently begun to explicitly address the lack of semantic mapping predicates and provenance for the crucial "maps to" relation in OMOP. As more data and vocabularies are mapped to OMOP, the time has come for the broader OHDSI community to get familiar with SSSOM and semantic entity mappings to improve their overall quality.

METHODS

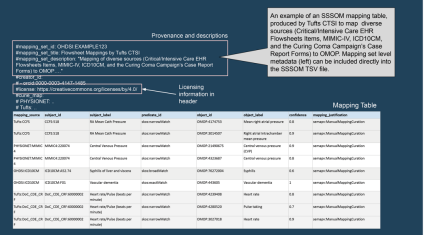
1. Established goals, focusing on mapping standardization and dissemination within the collaboration between SSSOM and OHDSI.
2. Explored use cases to identify the potential benefits for OHDSI, assessing the value proposition offered by SSSOM.
3. Expert discussions between SSSOM and OHDSI took place to achieve the objectives.
4. SSSOM integration by Tufts CTSI to create mappings to the OMOP Standardized Vocabularies from diverse sources (Critical/Intensive Care EHR Flowsheets Items, MIMIC-IV, ICD10CM, and the Curing Coma Campaign's Case Report Forms).
5. Endured accuracy of the mappings by following SSSOM specifications and adhering to OHDSI principles.
6. Mapping QA using SQL heuristics, ready for implementation within an ETL pipeline.
7. Thorough documentation of the entire process.

Increase mapping precision, enhance dissemination, facilitate integration, and promote interoperability across OHDSI using SSSOM

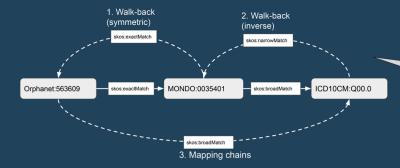
What is a semantic entity mapping?



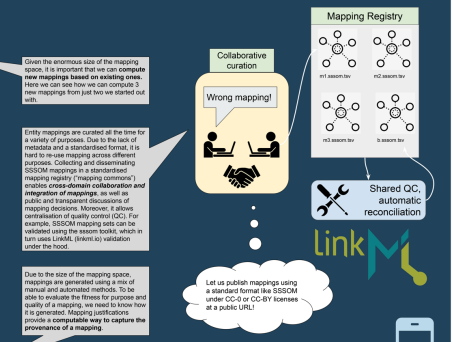
Example SSSOM TSV file



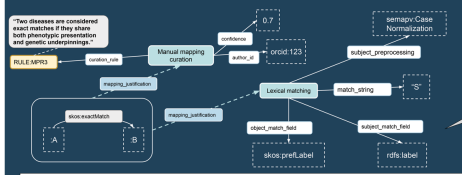
Semantic mapping predicates enable mapping chaining



Enhanced re-usability and collaborative mapping curation reduces cost



Mapping justifications provide clear provenance and increase trust



Take a picture to download the full paper

OHDSI | Tufts CTSI | EMBL-EBI | BOSCH | monarch INITIATIVE

*SSSOM can be pronounced "sessom"
Website: https://w3id.org/sssom
Email: monarchinitiative@gmail.com

RESULTS

In OMOP, only "Maps to" and "Maps to value" relationships are used for converting non-standard expressions to standard terms. There is no additional mapping metadata available, such as provenance or mapping precision (predicate).

Our initial analysis of ICD10CM-to-OMOP mappings using SSSOM reveals that only 10.9% of "Maps to" mappings demonstrate genuine exact matches, requiring further validation.

Problems with mappings include lack of metadata, ambiguity in mapping standards, mapping complexity, and the constant evolution of multiple terminologies.

Sharing mappings in SSSOM format improves mapping quality and overcomes limitations in the OMOP Vocabulary data model.

The OHDSI community contribution system has been modernized to include additional SSSOM-borrowed mapping metadata.

CONCLUSION

Well-documented and reusable mappings can be created effectively through SSSOM, improving overall mapping quality.

- The collaborative efforts have yielded a significant return on investment (ROI) by advancing mapping creation, dissemination, and coordination within the OHDSI community.
- Future plans include continued engagement with the OHDSI community, refining mapping practices, and exploring new use cases.
- Ongoing validation of ICD10CM-to-OMOP mappings using SSSOM will be pursued.
- Integration of Human Phenotype Ontology (HPO) as a standard vocabulary to OMOP CDM will be explored using the OMOP Vocabulary contribution system.
- Transforming OMOP2OBO mappings for machine-readable representation is another objective.
- Advocacy for the inclusion of a mapping metadata table in the Themis Working Group at the level of OMOP CDM conventions will be pursued.
- Promotion of SSSOM integration into OHDSI tools like Usagi, Jackalope, Perseus, and Atlas is planned.

SSSOM Contributor Community

Nicolas Matentzoglu, Polina Talapova, Andrew Williams, Alex H. Wagner, Anita Caron, Cássia Trojhan, Charlie Hoyt, Chris Mungall, Damien Goutte-Gattat, David Osumi-Sutherland, Emily Hartley, Ernesto Jimenez-Ruiz, Harshad Hegde, Henriette Harmse, Hyeongsik Kim, Ian Braun, Ian Harrow, James McLaughlin, Jim Balhoff, John Graybeal, Melissa Haendel, Nicole Vasilevsky, Nomi Harris, Núria Queralt Rosinach, Simon Jupp, Sophie Aubin, Thomas Liener, Tiffany Callahan, Tim Putman, William Duncan

sssom
SIMPLE STANDARD FOR SHARING ONTOLOGY MAPPINGS

#OHDSISocialShowcase

WEDNESDAY

Transforming EBMT Registry to the OMOP Common Data Model

(Maria Paula Busto, Marina Atlija, Freija Descamps, Ben Burke)

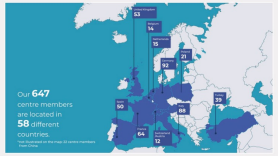
Transforming the EBMT dataset to OMOP-CDM 5.3 has been challenging. We have encountered technical and vocabulary challenges.

Ongoing Transformation of the EBMT Registry to the OMOP CDM 5.3

Background: In April 2021, EBMT was granted funds from the EHDEN - Data Partner Call to map it's registry data onto the OMOP CDM. Members of the EBMT are centres and individuals, active in the field of transplantation of any kind of haematopoietic cells, or any other organisation involved in the care of donors and recipients of haematopoietic cells.

Patients 702,579

Transplants 826,418



Results & Discussion

- ETL design is completed (see Figure 1) and is being tested and iterated on.
- ETL design driven by data size (> 4 million records)
- Mappings through SQL queries (memory-efficient)

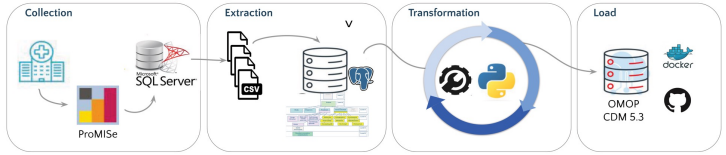
Encountered Issues

- Handling special characters in free text
- Custom concepts for very specific collected items
- Technical challenges due to DB size
- Genomic concepts
- Dates

Methods

The EBMT data is transformed to OMOP-CDM 5.3 through a custom-designed Extract-Transform-Load (ETL) process written primarily in the Python (v3.10) programming language. All transformations (source to PERSON, source to DEATH, etc.) are implemented in SQL and performed sequentially in an order predetermined based on table constraints and transformation dependencies.

Figure 1 ETL Process



Challenges: We have encountered technical challenges as well as vocabulary mapping challenges along the way. Dedicated functions and tools were developed to handle the technical hurdles. Where needed, custom vocabulary concepts are being created to ensure that the granularity of the source data can be maintained. We look forward to finalizing the work and collaborating with other EHDEN data-partners.

EBMT: Maria Paula Busto, Marina Atlija
Edence Health: Freija Descamps, Ben Burke

#OHDSISocialShowcase

THURSDAY

A new route of administration hierarchy derived from dose forms supporting standardised drug dose calculations

(Theresa Burkard, Artem Gorbachev, Kim Lopez-Güell, Daniel Prieto-Alhambra, Martí Català, Christian Reich)

This **new route of administration hierarchy**, derived from and linked to dose forms of drugs will **enable the use of route information in standardised analytics.**

A new route of administration hierarchy derived from dose forms supporting standardised drug dose calculations

Background: Few observational databases provide route of administration records. In addition, the current route vocabulary, which was adopted from sources with different use cases in mind, proved insufficient to make clinically relevant categories for dose estimation and to our knowledge have largely been ignored in analytical use cases. The current system also provides no hierarchical relationships between different routes.

Results: We created a route of administration hierarchy with "systemic", "local", and "other" (undefinable through dose form) as top classifiers. Subclassifications and their hierarchy are shown in Figure 1.
Some examples are listed in Table 1, where we can see the new route categorisation per dose form.

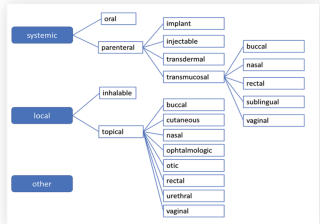


Figure 1. Suggested names and hierarchy of new route vocabulary

Table 1. Examples of suggested categorisation among the 214 existing dose form

Dose form concept id	Dose form concept name	Assigned route
19103220	12 hour Extended Release Capsule	oral
46234466	Auto-Injector	injectable
19082049	16 Hour Transdermal Patch	transdermal
40164192	Buccal Film	transmucosal - buccal
40220762	Sublingual Powder	transmucosal - sublingual
19127579	Dry Powder Inhaler	inhalable
19135446	Augmented Topical Gel	topical - cutaneous
43563498	Nasal Pin	topical - nasal
779945	Drug-Eluting Contact Lens	topical - ophtalmologic
19082194	Otic Ointment	topical - otic
19082574	Rectal Foam	topical - rectal
21014179	Vaginal delivery system	topical - vaginal

The frequency and proportion of unique drug concept ids per newly suggested routes in CPRD GOLD and CPRD AURUM are depicted in Table 2.

Table 2. Most frequent unique drug concept ids in the drug strength table of CPRD GOLD and CPRD AURUM that would be linked to the newly suggested route

Route	CPRD GOLD unique drug concept ids, n (%)	CPRD AURUM unique drug concept ids, n (%)
oral	906'064 (48.9%)	929'935 (48.7%)
injectable	365'685 (19.7%)	378'184 (19.8%)
has no dose form	261'108 (14.1%)	272'888 (14.3%)
topical - cutaneous	171'165 (9.2%)	174'156 (9.1%)
topical - ophtalmologic	42'344 (2.3%)	43'260 (2.3%)
inhalable	28'121 (1.5%)	29'297 (1.5%)
transdermal	13'752 (0.7%)	14'050 (0.7%)
topical - vaginal	11'917 (0.6%)	12'258 (0.6%)

Methods: We obtained all existing dose forms from ATHENA (searching for "Drug" domain, "dose form" concept class, "valid" flag and "RxNorm Extension" and "RxNorm" vocabularies). Based upon their name and looking at the actual drugs linked to them, we suggest a route for each dose form. TB (pharmacist) and AG (medical doctor) did this review independently and met for a consensus meeting with CR in which also the level, hierarchies, categories, and names of routes were defined.

Limitation: Most dose forms can be unambiguously assigned to a route of administration, but there are exceptions where we had to make a choice. Our review had a focus on systemic administrations because they are more relevant for dose estimations – topical administration faces surface tissue and dosing is therefore less stringent.

Theresa Burkard[1], Artem Gorbachev[2], Kim Lopez-Güell[1], Daniel Prieto-Alhambra[1,3], Martí Català[1], Christian Reich[3]

[1] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[2] Odysseus Inc
[3] Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

#OHDSISocialShowcase

FRIDAY

Pattern of long COVID symptoms and conditions: clustering analysis based on large multinational cohorts as part of an EHDEN Study-A-Thon

(Kim López-Güell#, Martí Català#, Daniel Dedman, Talita Duarte-Salles, Raivo Kolde, Raúl López-Blasco, Álvaro Martínez, Lourdes Mateu, Gregoire Mercier, Alicia Abellan, Johnmary T. Arinze, Theresa Burkard, Edward Burn, Zara Cuccu, Dominique Delseny, Chungsoo Kim, Ji-woo Kim, Kristin Kostka, Francesco Lapi, Cora Loste, Ettore Marconi, Miguel A. Mayer, Jaime Meléndez-Cardiel, Núria Mercadé-Besora, Mees Mosseveld, Hedvig ME Nordeng, Jessie O Oyinlola, Roger Paredes, Juan Manuel Ramírez-Anguita, Nhung TH Trinh, Anneli Uusküla, Bernardo Valdivieso, Junqing Xie, Annika M Jödicke*, Daniel Prieto-Alhambra)

Pattern of Long COVID Symptoms and Conditions: Clustering Analysis Based on Large Multinational Cohorts from an EHDEN Study-A-Thon

Kim López-Güell, Martí Català, Daniel Dedman, Talita Duarte-Salles, Raivo Kolde, Raúl López-Blasco, Álvaro Martínez, Lourdes Mateu, Gregoire Mercier, Alicia Abellan, Johnmary T. Arinze, Theresa Burkard, Edward Burn, Zara Cuccu, Dominique Delseny, Chungsoo Kim, Ji-woo Kim, Kristin Kostka, Francesco Lapi, Cora Loste, Ettore Marconi, Miguel A. Mayer, Jaime Meléndez-Cardiel, Núria Mercadé-Besora, Mees Mosseveld, Jessie O Oyinlola, Roger Paredes, Juan Manuel Ramírez-Anguita, Anneli Uusküla, Bernardo Valdivieso, Junqing Xie, Annika M Jödicke, Daniel Prieto-Alhambra

Abstract

We aimed to categorize and validate clusters of long COVID symptoms, as defined by WHO's "post-COVID-19 condition", through a broad international study. Results revealed numerous single-symptom clusters. However, when analyzing patients with multiple concurrent symptoms, we identified repeated clusters (anxiety-depression, cough-dyspnea, gastrointestinal-abdominal pain) across various databases. Still, considerable variation was noted among different databases and healthcare settings.

Introduction and Methods

Background
Long COVID, a condition with wide-ranging and heterogeneous symptoms, remains vaguely defined and difficult to diagnose. Clustering algorithms could potentially classify patients based on symptom combinations, aiding in understanding the disease's mechanisms and requirements.

Objectives
Our goal was to identify and confirm clusters of conditions and symptoms among patients with long COVID, using the WHO's clinical definition of "post-COVID-19 condition" in an extensive international network study.

Data
We conducted a multinational network study including databases form different healthcare settings: primary care, hospital databases and a claims-based database.

CPRD GOLD (UK)	CPRD AURUM (UK)	eDOL (France)	PSMAR (Spain)
IPCI (Netherlands)	CORIVA (Estonia)	Pharmetrics Plus (USA)	IISLAFE (Spain)

Long COVID definition
We studied those with COVID-19 (positive SARS-CoV-2 test or diagnosis) and no prior diagnosis/test within 42 days, then identified any post-acute symptoms from WHO's list of 25. A symptom was classified as long COVID if it appeared 90-365 days post-infection and wasn't in the 180-day pre-infection history.

Statistical analysis
We performed latent class analysis on these symptoms, considering patient's age and sex. We characterized clusters, analyzing patient baseline characteristics, healthcare use, vaccination status, and co-morbidities. We calculated the analyses for 2-7 clusters, choosing the optimal number based on statistical and clinical relevance.

Post-hoc analyses included clustering among subgroups with at least 2 or 3 different long COVID symptoms, conducted only in databases with over 500 individuals in the respective subgroup.

We performed all the analyses in R v4.2.1. Code available in https://github.com/oxford-pharmacoepi/LongCovidStuyathon_W3.

Results

Figure 1

Figure 2

Clustering results for individuals with 1+ symptoms (upper left), 2+ symptoms (upper right) and 3+ symptoms (lower left)

Figure 3

- The clustering solutions are heterogeneous, particularly for databases of different settings
- Databases pf the UD (CRPD GOLD and AURUM) show strikingly similar results for the subsets of >1 symptoms
- Some clusters seem to get repeated across countries (3+ symptoms): anxiety-depression, gastrointestinal-abdominal pain, cough-dyspnea

Conclusions

Using consistent clustering algorithms, symptom lists, and data structures across databases, we found diverse symptom combinations among long COVID patients. The prevalence of single-symptom clusters may reflect common symptoms recorded post-COVID-19, or persistent complications like olfactory deficits.

Analyzing subsets with 2+ or 3+ symptoms, clusters were more reflective of long COVID patient characteristics in specialist clinics, albeit results varied across databases and healthcare settings.

For more information, please email **kim.lopez@spc.ox.ac.uk** !

Scan to explore the Shiny App

OHDSI Shoutouts!

Any shoutouts from the community? Please share and help promote and celebrate OHDSI work!

Do you have anything you want to share? Please send to sachson@ohdsi.org so we can highlight during this call and on our social channels.
Let's work together to promote the collaborative work happening in OHDSI!

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Wednesday	12 pm	Latin America
Wednesday	7 pm	Medical Imaging
Thursday	9:30 am	Data Network Quality
Thursday	7 pm	Dentistry
Friday	9 am	GIS – Geographic Information Systems General
Friday	9 am	Phenotype Development and Evaluation
Friday	11 am	Clinical Trials
Monday	10 am	Healthcare Systems Interest Group
Monday	4 pm	Eyecare & Vision Research
Monday	6 pm	OMOP & FHIR

Sept. 15: HowOften Phenotype Library Contribution Deadline

OHDSI

HowOften: Community contributions wanted

General

Patrick_Ryan 2h

Friends:

As we discussed on the 20June2023 and 15August2023 community calls, @hripcsa and I would like to encourage our community to think big and collaborate together in a effort toward large-scale incidence characterization. HowOften is be a community-wide study to define a broad set of target cohorts T that'll serve as denominators, and another broad set of outcome cohorts O that'll serve as numerators. And for a defined list of time-at-risk windows (e.g. 30d, 1yr, all-time), stratified by age/sex/index year, we will compute the incidence of O in T for all T-O combinations within each database in our participating network, and then meta-analyze the results to produce composite summaries.

As with all OHDSI network studies, we will use GitHub to share study materials, including protocol and source code, which should be based where possible off of existing HADES packages. And we intend to make the full resultset publicly available through an interactive website, likely initially taking advantage of the RShiny modules built by the HADES team as part of the Strategus workflow. As we've seen with prior OHDSI work, background incidence rates can be used for a wide range of clinical applications, including providing disease natural history, providing context for pharmacovigilance by quantifying the magnitude of risk for known effects, and reporting digital quality measures (see @bnhamlin 's talk here).

Titan Award Nominations Are Open!

To recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI's mission, the OHDSI Titan Awards were introduced at the 2018 Symposium and have been handed out at the Global Symposium each year since.

bit.ly/2023TitanNominations

Global Symposium

Oct. 20-22 • East Brunswick, NJ, USA

Hilton East Brunswick Hotel & Executive Meeting Center

ohdsi.org/OHDSI2023

Global Symposium Weekend Agenda

	Friday, Oct. 20	Saturday, Oct. 21	Sunday, Oct. 22
7:00 am	Registration/Lite Breakfast	Lite Breakfast	Lite Breakfast
8:00 am	Welcome to OHDSI2023!	Intro to OHDSI Tutorial & OHDSI Workgroup Activities	OHDSI collaborative workshop: HowOften (part 2)
9:00 am	State of the Community		
10:00 am	Community Networking		
11:00 am	Plenary Session		
12:00 pm	Buffet Lunch	Buffet Lunch + Collaborator Showcase: Posters & Demos	Buffet Lunch + Collaborator Showcase: Posters & Demos
1:00 pm	Panel: Network Studies	OHDSI collaborative workshop: HowOften (part 1)	OHDSI workgroup activites
2:00 pm	Collaborator Showcase: Lightning Talks		
2:45 pm	Collaborator Showcase: Posters & Demos		
3:30 pm	Collaborator Showcase: Lightning Talks		
4:15 pm	Collaborator Showcase: Posters & Demos		
5:00 pm	Closing Talk & Titan Awards	Free time	We'll see you again in 2024!
6:00 pm	Networking Reception		
7:00 pm	OHDSI Got Talent!		

** this agenda is tentative and subject to change*

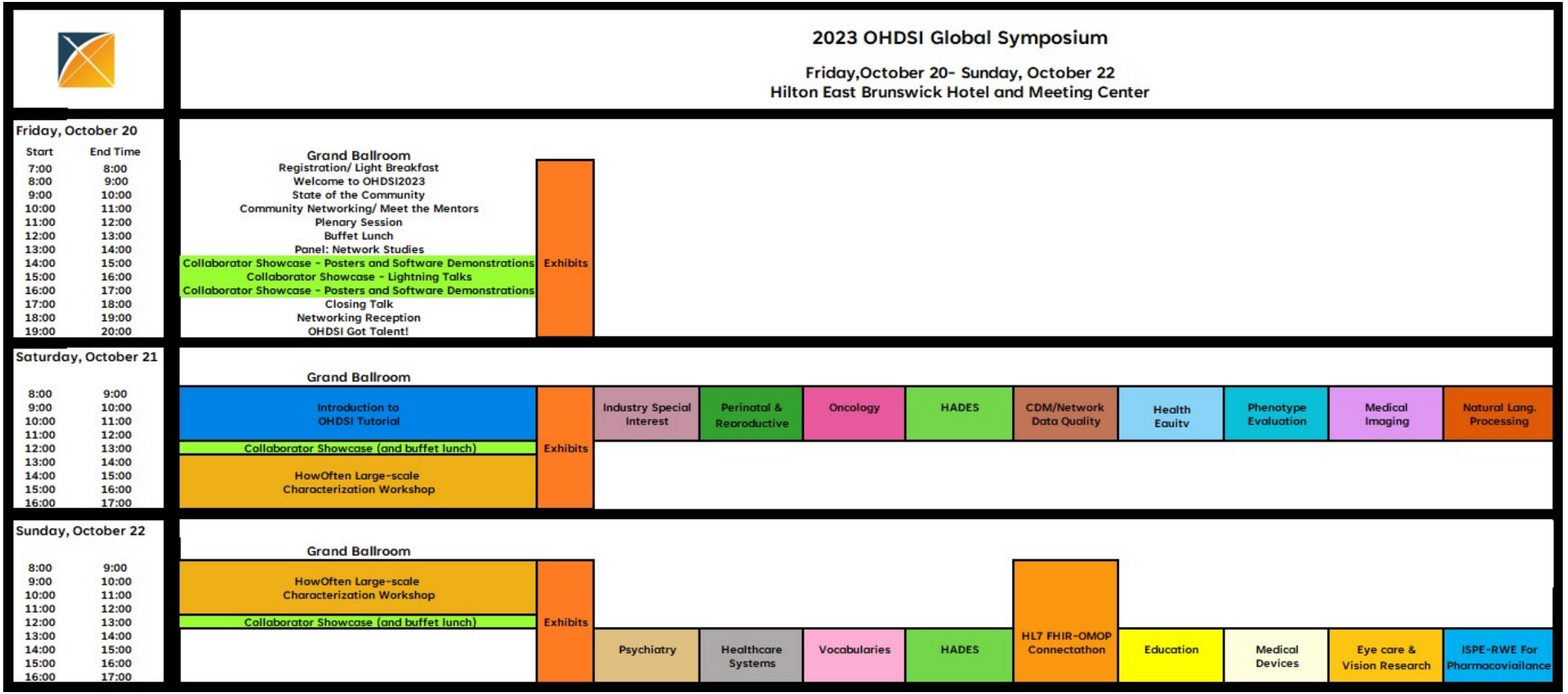
OHDSI Got Talent!

Please join us for the first **OHDSI Got Talent!** competition at our 2023 Global Symposium.

We are looking for anybody with a special talent – singing, dancing, playing an instrument, comedy, magic, etc. – to join us for this fun event in October. Please use the link below to share your interest in participation!

bit.ly/OHDSIGotTalent2023

Global Symposium



Where Are We Going?

Any other announcements of upcoming work, events, deadlines, etc?

Three Stages of The Journey

Where Have We Been?

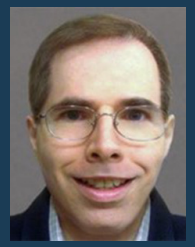
Where Are We Now?

Where Are We Going?

Aug. 22 — OHDSI and Clinical Registries:
Sanity for Health Systems

Paul Nagy

Program Director for Graduate Training in Biomedical Informatics and Data Science, Deputy Director of the Johns Hopkins Medicine Technology Innovation Center



Lee Evans

Founder, LTS Computing LLC

DuWayne Willett

Chief Medical Informatics Officer, University of Texas Southwestern Health System

Jeff Weaver

Director of Data Solutions for Emory University