OHDSI "Speed Dating"

OHDSI Community Call
Aug. 16, 2022 • 11 am ET

Upcoming OHDSI Community Calls

Date	Topic
Aug. 23	Workgroup Updates
Aug. 30	EHDEN Academy/EHDEN Portal
Sept. 6	OHDSI Studies
Sept. 13	Registries And Their Adoption To OMOP
Sept. 20	OHDSI2022 Preview
Sept. 27	HTA Challenge

Upcoming OHDSI Community Calls

Date	Topic
Aug. 23	Workgroup Updates
Aug. 30	EHDEN Academy/EHDEN Portal
Sept. 6	OHDSI Studies
Sept. 13	Registries And Their Adoption To OMOP
Sept. 20	OHDSI2022 Preview
Sept. 27	HTA Challenge

Aug. 23 Community Call: Workgroup Updates

Registry

Tina Parciak

PhD Student • UHasselt/BIOMED

Latin America

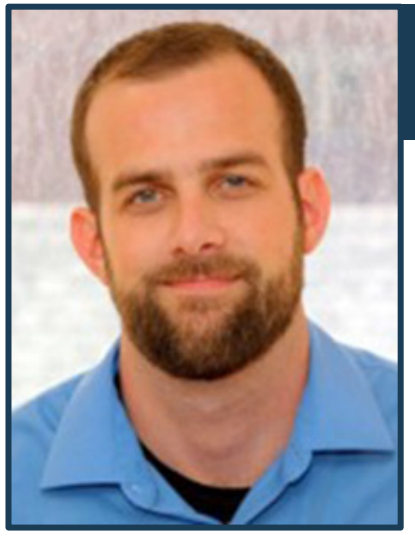
Jose Posada

Assistant Professor • Universidad del Norte

Health Equity

Jake Gillberg

Software Developer • Tufts Clinical and Translational Science Institute



GIS

Robert Miller

Software Development Analyst • Tufts Clinical and Translational Science Institute

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Antoine Lamer, Mouhamed Djahoum Moussa, Romaric Marcilly, Régis Logier, Benoit Vallet and Benoît Tavernier** on the publication of **Development and usage of an anesthesia data warehouse: lessons learnt from a 10-year project** the Journal of Clinical Monitoring and Computing.

Journal of Clinical Monitoring and Computing
https://doi.org/10.1007/s10877-022-00898-y

ORIGINAL RESEARCH

Development and usage of an anesthesia data warehouse: lessons learnt from a 10-year project

Antoine Lamer[1,2] · Mouhamed Djahoum Moussa[3] · Romaric Marcilly[1,4] · Régis Logier[4] · Benoit Vallet[1] · Benoît Tavernier[1,5]

Received: 16 March 2022 / Accepted: 12 July 2022

Abstract
This paper describes the development and implementation of an anesthesia data warehouse in the Lille University Hospital. We share the lessons learned from a ten-year project and provide guidance for the implementation of such a project. Our clinical data warehouse is mainly fed with data collected by the anesthesia information management system and hospital discharge reports. The data warehouse stores historical and accurate data with an accuracy level of the day for administrative data, and of the second for monitoring data. Datamarts complete the architecture and provide secondary computed data and indicators, in order to execute queries faster and easily. Between 2010 and 2021, 636 784 anesthesia records were integrated for 353 152 patients. We reported the main concerns and barriers during the development of this project and we provided 8 tips to handle them. We have implemented our data warehouse into the OMOP common data model as a complementary downstream data model. The next step of the project will be to disseminate the use of the OMOP data model for anesthesia and critical care, and drive the trend towards federated learning to enhance collaborations and multicenter studies.

Keywords Anesthesia information management system · Data warehouse · Data reuse · Database · Retrospective studies

1 Introduction

The Anesthesia Information Management System (AIMS) has gradually replaced the paper records in the last decades [1, 2]. This electronic version of the anesthesia record captures data automatically from clinical monitors and facilitates the documentation of the anesthesia procedure such as the main stages of the surgical intervention and anesthesia management. The benefits of an AIMS include improved readability of the anesthesia record and greater efficiency in documentation efforts [3].

Anesthesia records are mainly used for reviewing single patient record but also help searching, querying and retrieving data from thousands of cases to conduct retrospective analyses for clinical research or evaluate compliance with guidelines [4–6]. The main characteristics of the data recorded in the operating room are their sampling frequency and accuracy, with one measurement every 30 s for signals

OHDSI Shoutouts!

Congratulations to the team of **Aurore Nishimwe, Charles Ruranga, Clarisse Musanabaganwa, Regine Mugeni, Muhammed Semakula, Joseph Nzabanita, Ignace Kabano, Annie Uwimana, Jean N. Utumatwishima, Jean Damascene Kabakambira, Annette Uwineza, Lars Halvorsen, Freija Descamps, Jared Houghtaling, Benjamin Burke, Odile Bahati, Clement Bizimana, Stefan Jansen, Celestin Twizere, Kizito Nkurikiyeyezu, Francine Birungi, Sabin Nsanzimana and Marc Twagirumukiza** on the publication of **Leveraging artificial intelligence and data science techniques in harmonizing, sharing, accessing and analyzing SARS-COV-2/COVID-19 data in Rwanda (LAISDAR Project): study design and rationale** in BMC Medical Informatics and Decision Making.

RESEARCH Open Access

Leveraging artificial intelligence and data science techniques in harmonizing, sharing, accessing and analyzing SARS-COV-2/COVID-19 data in Rwanda (LAISDAR Project): study design and rationale

Aurore Nishimwe[1*], Charles Ruranga[2], Clarisse Musanabaganwa[3], Regine Mugeni[4], Muhammed Semakula[3], Joseph Nzabanita[5], Ignace Kabano[2], Annie Uwimana[2], Jean N. Utumatwishima[4], Jean Damascene Kabakambira[6], Annette Uwineza[6], Lars Halvorsen[7], Freija Descamps[7], Jared Houghtaling[7], Benjamin Burke[7], Odile Bahati[8], Clement Bizimana[8], Stefan Jansen[1], Celestin Twizere[9], Kizito Nkurikiyeyezu[9], Francine Birungi[1], Sabin Nsanzimana[3] and Marc Twagirumukiza[1,10]

Abstract

Background: Since the outbreak of COVID-19 pandemic in Rwanda, a vast amount of SARS-COV-2/COVID-19-related data have been collected including COVID-19 testing and hospital routine care data. Unfortunately, those data are fragmented in silos with different data structures or formats and cannot be used to improve understanding of the disease, monitor its progress, and generate evidence to guide prevention measures. The objective of this project is to leverage the artificial intelligence (AI) and data science techniques in harmonizing datasets to support Rwandan government needs in monitoring and predicting the COVID-19 burden, including the hospital admissions and overall infection rates.

Methods: The project will gather the existing data including hospital electronic health records (EHRs), the COVID-19 testing data and will link with longitudinal data from community surveys. The open-source tools from Observational Health Data Sciences and Informatics (OHDSI) will be used to harmonize hospital EHRs through the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). The project will also leverage other OHDSI tools for data analytics and network integration, as well as R Studio and Python. The network will include up to 15 health facilities in Rwanda, whose EHR data will be harmonized to OMOP CDM.

Expected results: This study will yield a technical infrastructure where the 15 participating hospitals and health centres will have EHR data in OMOP CDM format on a local Mac Mini ("data node"), together with a set of OHDSI open-source tools. A central server, or portal, will contain a data catalogue of participating sites, as well as the OHDSI tools that are used to define and manage distributed studies. The central server will also integrate the information

OHDSI Shoutouts!

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

Have a study published? Please send to sachson@ohdsi.org so we can share 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
Tuesday	1 pm	Common Data Model
Wednesday	9 am	FHIR and OMOP Data Model Harmonization Subgroup (ZOOM)
Wednesday	9 am	Africa Chapter
Wednesday	10 am	FHIR and OMOP Digital Quality Measurements Subgroup (ZOOM)
Wednesday	12 pm	Health Equity
Thursday	12 pm	HADES
Thursday	12 pm	FHIR and OMOP Oncology Subgroup Vocabulary/Development Subgroup
Thursday	1 pm	OMOP CDM Oncology
Thursday	6 pm	FHIR and OMOP Terminologies Subgroup (ZOOM)
Friday	9 am	GIS – Geographic Information System Development
Friday	10:30 am	Clinical Trials
Tuesday	4 pm	OMOP CDM Oncology Genomic Subgroup

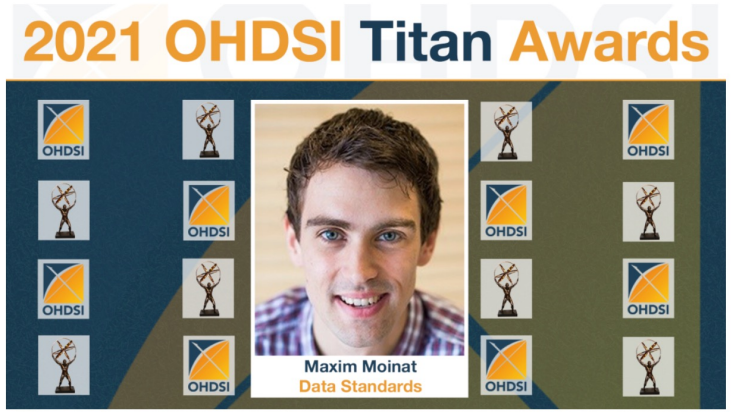
www.ohdsi.org/upcoming-working-group-calls

Titan Awards Nominations Are Open

Nominations for the 2022 Titan Awards are now OPEN!
Please use the form below to nominate an individual or institution for a top contribution to the OHDSI community this past year!

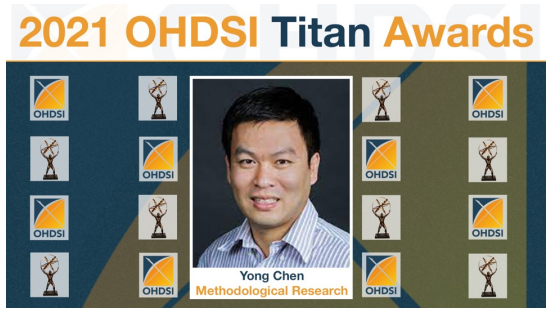
2022 Nomination Form

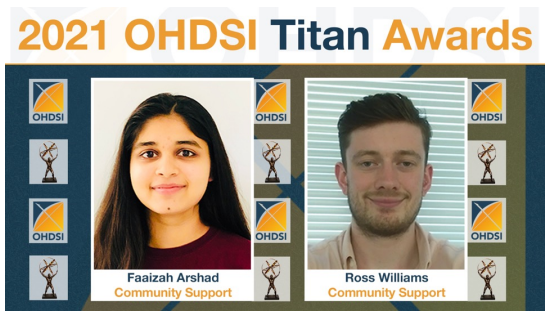
To recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI's mission, the OHDSI Titan Awards were introduced at the 2018 Symposium and have been handed out at the U.S./Global Symposium each year since. Annually, community members are invited to nominate individuals or institutions they feel have made significant contributions towards advancing OHDSI's mission, vision and values. Once nominations are submitted, the OHDSI Titan Award Committee will select the award winners. Award winners will be announced before the networking reception at the annual symposium. The award categories, as well as all previous recipients, can be found below.



Titan Award for Data Standards – to recognize extraordinary contributions by an individual, organization, or team in development or evaluation in community data standards, including OMOP common data model and standardized vocabularies

- **2021 –** Maxim Moinat, The Hyve/Erasmus University Medical Center
- **2020 –** Clair Blacketer, Janssen Research and Development
- **2019 –** Oncology Workgroup (Michael Gurley, Northwestern Univ.; Rimma Belenkaya, Memorial Sloan Kettering Cancer Center; Robert Miller, Tufts CTSI)
- **2018 –** Vocabulary team (Christian Reich, IQVIA; Anna Ostropolets, Columbia Univ.; Dmitry Dymshyts, Odysseus Data Services)





ohdsi.org/titan-awards

2022 OHDSI Symposium

Registration is OPEN for #OHDSI2022!





The 2022 OHDSI Symposium will be held Oct. 14-16 at the Bethesda North Marriott Hotel & Conference Center.

www.ohdsi.org/ohdsi2022symposium

2022 OHDSI Symposium

OHDSI Community Calls | Events & Past Collaborations | Learn About & Join OHDSI Workgroups | This Week In OHDSI | EHDEN Academy
Annual Report: Our Journey | Publications | Support & Sponsorship | OHDSI2022 Symposium | Newsletters | Follow OHDSI on Social

2022 OHDSI Symposium

Oct. 14-16 · Bethesda North Marriott Hotel & Conference Center

We are thrilled to announce that registration for the 2022 OHDSI Symposium, which will be held Oct. 14-16 at the Bethesda North Marriott Hotel & Conference Center, is now open!

It is so exciting to bring our community back together this fall. Our collaborator showcase will return; please click the link to see how you can take part in our poster presentations, software demos and lightning talks. The full agenda for our conference is still being developed, so please continue to check the OHDSI website (www.ohdsi.org) and our social platforms for updates as we plan for the 2022 Symposium.

The main conference will be held Friday, Oct. 14. A full-day tutorial will be held Saturday, Oct. 15, while other community activities will be held both Oct. 15 and Oct. 16.

Symposium Registration Details

Friday, Oct. 14 – Main Conference

Registration Fee: $500*
** this is an open and inclusive event; if the registration fee represents a burden to you, please contact symposium@ohdsi.org.*

Should you need to make changes or cancel your registration ticket, please follow the instructions you will receive on your Eventbrite confirmation upon registration completion. Please note that tickets can be refunded up until 7 days prior to the event; Eventbrite fees are not refundable.

Register For The Main Conference · Friday, Oct. 14

Saturday, Oct. 15 – Full-Day Tutorial: An Introductory Journey From Data To Evidence

Registration Fee: $300*
** this is an open and inclusive event; if the registration fee represents a burden to you, please contact symposium@ohdsi.org.*

Should you need to make changes or cancel your registration ticket, please follow the instructions you will receive on your Eventbrite confirmation upon registration completion. Please note that tickets can be refunded up until 7 days prior to the event; Eventbrite fees are not refundable.

Register For The Full-Day Tutorial · Saturday, Oct. 15

What Will Be Taught At This Tutorial?

Saturday, Oct. 15 and Sunday, Oct. 16 — Community Activities

A highlight of the OHDSI Symposium will be a full weekend of workgroup activities and meetings within the Bethesda North Marriott Hotel & Conference Center. You are now able to register for any workgroup sessions as long as there is no overlap between any two sessions; registration is free, but please do so early as this will be first-come, first-served due to room capacity.

See The Schedule & Agenda For Workgroup Activities · Weekend of Oct. 15-16

Register For Workgroup Activities · Weekend of Oct. 15-16

Hotel Information and Sleeping Room Block

Hotel: Bethesda North Marriott Hotel & Conference Center
Address: 5701 Marinelli Road, Rockville, Maryland, 20852
Hotel Main Number: (301) 822-9200
Reservations Toll Free: (877) 212-5752
Reservations Local Phone: (301) 822-9200

This year, OHDSI is holding a sleeping room block for the nights of Oct. 13 and 14 with a special room rate of $179 plus taxes. Please note that all sleeping rooms are on a first-come, first-served basis. To help us in the planning process, we ask that you do not cancel your hotel room ordered through the OHDSI Room Block. If you must cancel, please let us know prior to Thursday, Sept. 1, so that we can offer the room to others who may need one. Once the room block is full, or if specific nights are sold out, you may make additional room reservations on the hotel's website or by calling the hotel phone number above. Please note that OHDSI is not holding any sleeping rooms on Saturday, Oct. 15. Therefore, please call the hotel phone number or make this reservation online should you plan to stay Saturday night.

ohdsi.org/ohdsi2022symposium

#OHDSISocialShowcase This Week

Conversion of Estonian health data into the OMOP CDM: *insurance claims, prescription data and electronic health records*

PRESENTER: **Marek Oja**

INTRO:

- Estonia needs a research database where all health data is standardized and ready to use for observational research.

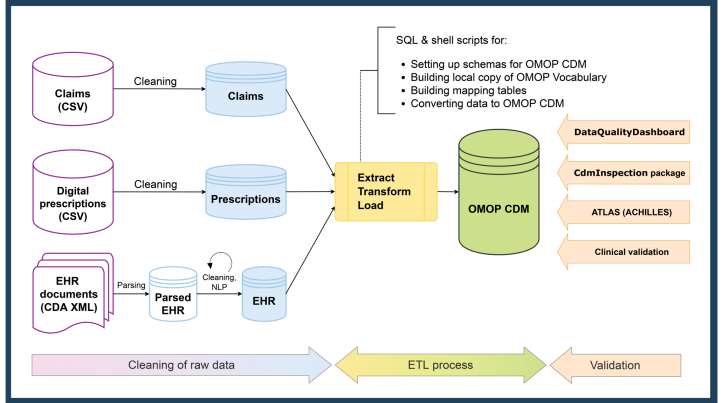
METHODS

- 10% random sample of the Estonian population (n=149K patients) from 2012 to 2019
- Dataset included three national data sources:
  - insurance claims (n=6.2M)
  - digital prescriptions (n=9.6M)
  - electronic health records (EHR) (n=4.97M)
- OMOP CDM v5.3
- Technologies used:
  - Git for version control
  - PostgreSQL for database
  - Python (luigi), SQL, bash for ETL pipelines
  - EstNLTK for NLP
- Translation of local vocabularies to standard vocabularies

RESULTS

- All three different data sources were combined successfully into one OMOP CDM. With this, we have a full view of patient data over the observation period.
- Process is repeatable and used for different datasets and projects in Estonia:
  - Asthma specific dataset
  - COVID specific dataset
  - Estonian Biobank health data
- Participation in network studies:
  - Prostate cancer study - PIONEER

Repeatable ETL process to transform Estonian health data to OMOP CDM



Statistics on the ETL procedure to convert Estonian health data to OMOP CDM

Table	Total number of records	Mapped records	Mapping rate	Number of source codes	Number of codes mapped	Mapping rate
location	1					
care_site	1,820					
person	149,351					
death	8,277					
observation_period	149,351					
visit_occurrence	18,194,512	18,194,512	100.0%	10	10	100.0%
visit_details	48,002	48,002	100.0%	8	8	100.0%
condition_occurrence	20,238,707	20,220,877	99.9%	9,240	9,215	99.7%
procedure_occurrence	6,950,829	5,386,453	77.5%	5,999	1,506	25.1%
drug_exposure	8,596,491	8,151,931	94.8%	1,153	915	79.4%
device_exposure	77,842	47,296	60.8%	159	4	2.5%
observation	15,159,794	14,721,482	97.1%	755	548	72.6%
measurement	30,440,242	27,457,547	90.2%	3,524	2,904	82.4%
measurement-value	1,544,690	1,544,690	100.0%	49	49	100.0%
measurement-unit	11,193,028	11,193,028	100.0%	79	79	100.0%
drug_era	4,313,993	4,313,993	100.0%	774	774	100.0%
condition_era	9,347,695	9,347,695	100.0%	6,401	6,401	100.0%

Source vocabulary mappings to standardized vocabularies

Source vocabulary	Target vocabulary	Count	Percent
ATC	RxNorm	5,111	87.7%
ATC	RxNorm Extension	368	6.3%
ATC (mostly combination drugs)	0	352	6.0%
Cancer related findings (TNM codes, cancer stages, etc.)	Cancer Modifier	9	8.0%
Cancer related findings (TNM codes, cancer stages, etc.)	NCIt	36	32.1%
Cancer related findings (TNM codes, cancer stages, etc.)	SNOMED	67	59.8%
Pathology findings, body measurements	SNOMED	18	100.0%
Drug administration routes	SNOMED	87	100.0%
Local codes from claims (procedures, drugs, measurements, etc)	LOINC	32	0.9%
Local codes from claims (procedures, drugs, measurements, etc)	OMOP Extension	3	0.1%
Local codes from claims (procedures, drugs, measurements, etc)	RxNorm	85	2.5%
Local codes from claims (procedures, drugs, measurements, etc)	RxNorm Extension	14	0.4%
Local codes from claims (procedures, drugs, measurements, etc)	SNOMED	1,188	34.6%
Local codes from claims (procedures, drugs, measurements, etc)	0	2,110	61.5%
ICD10	Cancer Modifier	21	0.1%
ICD10	OMOP Extension	3	0.0%
ICD10	SNOMED	20,016	99.9%
ICD10	0	1	0.0%
LOINC	LOINC	81,452	97.9%
LOINC	SNOMED	731	0.9%
LOINC (mostly local and temporary LOINC codes)	0	1,019	1.2%
NOMESCO Classification of Surgical Procedures (NCSP)	RxNorm	2	0.0%
NOMESCO Classification of Surgical Procedures (NCSP)	SNOMED	728	9.8%
NOMESCO Classification of Surgical Procedures (NCSP)	0	6,671	90.1%
Measurement units	UCUM	922	100.0%

Marek Oja[1], Sirli Tamm[1], Sulev Reisberg[1,2,3], Raivo Kolde[1], Sven Laur[1], Hendrik Šuvalov[1], Harry-Anton Talvik[1,2,3], Jaak Vilo[1,2]

[1] Institute of Computer Science (ICS), University of Tartu, Tartu, Estonia
[2] STACC, Tartu, Estonia
[3] Quretec, Tartu, Estonia

Contact: marek.oja@ut.ee

Acknowledgements This work was supported by the Estonian Research Council grants number PRG1095, RITA1/02-96; the European Union through the European Regional Development Fund grant number EU48684; and the European Social Fund via IT Academy programme. The whole conversion was carried out in the High Performance Computing Center of the University of Tartu.

MONDAY

Conversion of Estonian health data into the OMOP CDM: insurance claims, prescription data and electronic health records

Lead: Marek Oja

#OHDSISocialShowcase This Week

Developing a frailty concept in the OMOP CDM among sexual and gender minority older adults (age 50+) in the All of Us database

Brianne Olivieri-Mui1,2, Chelsea Wong2, Michael Wilczek1, Jordan Bosse3
1.The Roux Institute, Northeastern University, 2. Marcus Institute for Aging Research, Harvard Medical School, 3. School of Nursing, Northeastern University

INTRODUCTION

- Deficit accumulation frailty measures have prognostic value, are comprehensive and can be applied across many data sources
- Frailty is not a standardized concept in many common data models, including the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM)
- Frailty in the older sexual and gender minority (OSGM) population has not been studied
- The All of Us (AoU) Research Program provides an opportunity to study frailty among OSGM and to create a frailty concept for the OMOP CDM

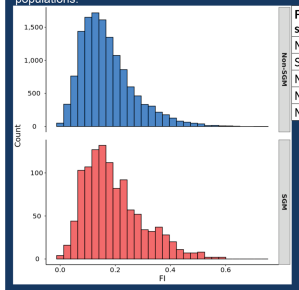
METHODS

- n = 13,357 non-OSGM; n = 1,118 OSGM; Aged 50+ with complete data
- Using AoU baseline surveys, developed a 35-item deficit accumulation frailty index (AOU-FI) based on validated FI's[1,2,3]
- Deficit items included concepts spanning comorbidities (18 concepts), physical functioning (9 concepts), mental health (6 concepts), and cognition (2 concepts)
- Compared AOU-FI to two known FI distributions using t-tests
- Performed principal components analysis of the 35-items

OSGM potentially have **higher frailty** at younger ages

Frailty would be a valuable concept **to add to the OMOP CDM** for AoU users

Figure 1. Comparing distributions of the All of Us – Frailty Index for the OSGM and general older All of Us populations.



FI Summary statistics	Non-OSGM (n=13,357)	OSGM (n=1,118)
Mean	0.17	0.19
SD	0.10	0.11
Median	0.15	0.17
Min	0	0
Max	0.74	0.60

Figure 2. PCA Scree plot for dimension 1 of each group

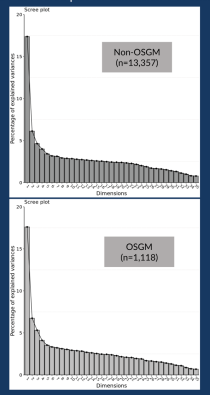
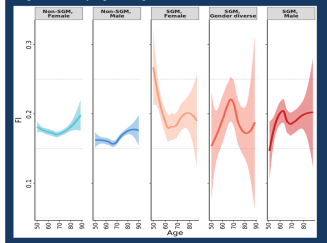


Figure 3. FI by age and gender



1. Mitnitski AB, Song X, Rockwood K. The estimation of relative fitness and frailty in community-dwelling older adults using self-report data. J Gerontol A Biol Sci Med Sci. 2004;59(6):M627-32.
2. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC geriatrics. 2008;8(1):24.
3. Schoufour JD, Mitnitski A, Rockwood K, Evenhuis HM, Echteld MA. Development of a frailty index for older people with intellectual disabilities: results from the HA-ID study. Res Dev Disabil. 2013;34(5):1541-50.

The Roux Institute Northeastern University

Marcus Institute for Aging Research | Harvard Medical School Affiliate
Hebrew SeniorLife

RESULTS

- The AOU-FI is a ratio (range 0-1) with a maximum of 35-items worth up to 1 point each
- Both AOU-FI distributions had expected gamma shapes (Figure 1)
- The non-OSGM mean was higher (p<.01) than the known Canadian FI distribution (mean=0.164; sd=0.098)
- The OSGM mean was higher than the known Canadian distribution, but lower (p<.01) than the FI for people with intellectual disabilities (mean=0.27; sd=0.13)
- 35-items are each independently contributing to the AOU-FI, justifies our choice of the items (Figure 2)
- Both groups were >80% white. Non-OSGM were 42% male, 61% age 60 or younger <1% had HIV. OSGM were 54% male, 70% age 60 or less, 5% had HIV.
- Compared to non-OSGM, mean age of OSGM was significantly lower (65 [sd=8] vs 66 [sd=9]), but the AOU-FI was significantly higher (p<.01)
- Age trends for FI were as expected for non-OSGM (Figure 3)

DISCUSSION

- AOU-FI is consistent with shape and behavior of established FI distributions
- OSGM potentially have higher frailty at younger ages compared to a general older population
- Adding the AOU-FI as a concept to the OMOP CDM for AoU users will be critical to maximizing the utility of these data for studying vulnerable subpopulations of older adults

TUESDAY

Developing a frailty concept in the OMOP CDM among sexual minority older adults (age 50+) in the All of Us database

Lead: Brianne Olivieri Mui

#OHDSISocialShowcase This Week

CohortsExport:
A Shiny app to explore and export data from the OMOP Common Data Model

Vittoria Ramella

INTRO:

- One of the first and fundamental steps in observational research is defining suitable cohorts of patients with shared characteristics of interest. Being able to explore raw data of patients belonging to those cohorts could be helpful for researchers to verify whether a cohort definition meets the needs of their clinical questions.

METHODS

1. *CohortsExport* is a Shiny app built using the open-source packages and tools developed by the R/Shiny community, as well as tools developed by the OHDSI community
2. Requires access to a database containing data in OMOP format
3. Uses HADES packages to connect to database platforms, to retrieve cohort definitions and concept sets from ATLAS, and to query the database

RESULTS

- The Shiny app provides an interface to select a cohort and visualize record count for the main CDM tables
- User can select tables to export and, optionally, filter them using concept sets
- The app shows a preview of the data and allows the user to apply an additional filter on each table to add/remove columns from the output
- User can export and download data in different formats (currently Microsoft Excel, R, SAS, SPSS and Stata)

BIOMERIS

Export patient-level data from the OMOP CDM.

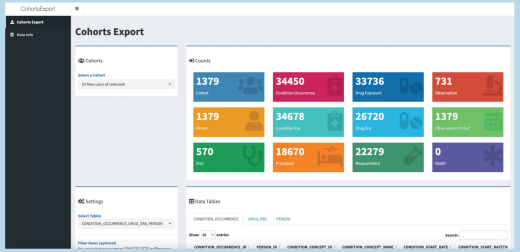


Figure 1. Main window of the *CohortsExport* app. First, the user selects a cohort and then the record count for the main CDM tables are shown.

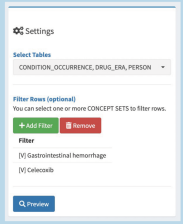


Figure 2. Example of export settings.
Here the user selects the tables to export and, optionally, one or more concept sets to filter them.

Figure 3. Example of *Data preview* dialog box.
The dialog box shows as many tabs as the selected tables and, for each one of these, the user can add/remove each single column from the output.

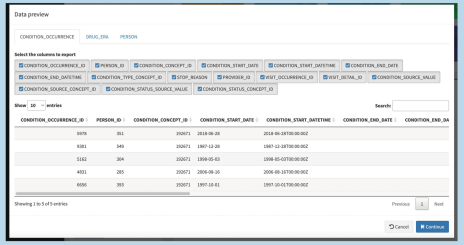
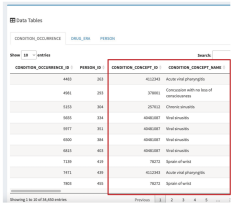


Figure 4. Available export formats

More about the app
In the *Data Tables* box the user can see a preview of the data to be exported. Additional columns have been added beside each *"concept_id"* column making explicit the concept name, to help researchers and improve readability.



Future developments

- Login page
- Integrate the WebApi/ATLAS user permissions
- Save cohort definitions and export settings for reproducibility
- Export patient-level data grouped by visit ID
- Additional filters (e.g., date range, ...)

Vittoria Ramella,
Matteo Gabetta,
Nicola Barbarini

OHDSI

WEDNESDAY

CohortsExport: A Shiny app to explore and export data from the OMOP Common Data Model

Lead: Vittoria Ramella

#OHDSISocialShowcase This Week

Title:Analyzing the impact of COVID-19 on the healthcare system: an OMOP-CDM framework applied to Northern Italy

PRESENTER: **Sara Conti**

INTRO

- **The VICES-SMIRE project (University of Milano-Bicocca in partnership with the Health Protection Agency of Bergamo, Northern Italy, epicenter of COVID-19 first outbreak) analyzed the impact of COVID-19 outbreak on mortality and hospital admissions.**
- **Aim is to use the same statistical approach using databases harmonized to OMOP-CDM and to compare results.**

METHODS

Study period: 01/01/2017-31/12/2021

Population: All residents assisted by the HPA

OMOP-CDM tables: PERSON, LOCATION, DEATH, VISIT_OCCURRENCE

Outcomes: deaths and hospital admissions

Statistical model:

- Generalized additive model applied to the daily time-series of each outcome[1].
- Estimation of the absolute and percent variation of each outcome during the different waves of the outbreak.

RESULTS

Table 1 – Percent variation in the outcomes (95% Confidence interval)

Population subgroup	% Variation deaths	% Variation hospital admissions
Total	34.54 (28.91 to 39.78)	-19.76 (-21.92 to -17.52)
Gender		
Female	30.75 (23.89 to 37.61)	-19.58 (-21.91 to -17.40)
Male	38.19 (30.67 to 46.58)	-19.96 (-22.46 to -17.57)
Age class		
0-18	13.63 (-54.15 to 79.91)	-29.62 (-33.14 to -26.36)
19-59	16.36 (-1.21 to 31.64)	-17.08 (-19.77 to -14.34)
60-69	38.13 (21.19 to 56.44)	-17.56 (-21.02 to -14.14)
70-79	41.17 (29.72 to 53.02)	-19.92 (-22.80 to -17.08)
80-89	41.24 (32.76 to 50.77)	-22.06 (-24.83 to -19.39)
90+	23.08 (13.92 to 32.24)	-29.64 (-34.16 to -25.24)

Estimates of COVID-19 impact derived from OMOP-CDM data sources **reproduced those obtained on the original ones.**

This approach can be used to **compare different contexts within the EHDEN network.**

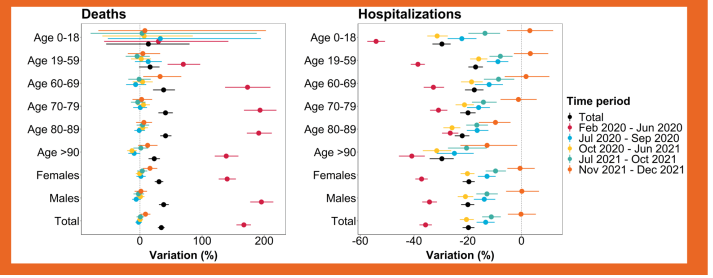


Table A1 – Main results from the Data Quality Dashboard on the CDM instance used for the analyses

Table S2 –Detail of the impact analyses. Estimated absolute variation in the frequency of death and hospital admission

Population subgroup	Observed deaths	Absolute deaths variation	Observed hospital admissions	Absolute hospital admissions variation
Total	24,852	6,380 (5,574 to 7,073)	178,678	-43,994 (-50,166 to -37,948)
Gender				
Female	12,168	2,862 (2,346 to 3,326)	91,148	-22,198 (-25,574 to -19,202)
Male	12,684	3,506 (2,977 to 4,031)	87,530	-21,824 (-25,361 to -18,656)
Age class				
0-18	65	8 (-77 to 29)	12,423	-5,229 (-6,158 to -4,446)
19-59	1,535	216 (-19 to 369)	73,028	-15,038 (-17,994 to -12,228)
60-69	2,259	624 (395 to 815)	28,417	-6,053 (-7,562 to -4,682)
70-79	5,340	1,557 (1,224 to 1,850)	34,142	-8,493 (-10,085 to -7,032)
80-89	9,799	2,861 (2,418 to 3,300)	25,541	-7,230 (-8,437 to -6,145)
90+	5,854	1,098 (716 to 1,427)	5,127	-2,160 (-2,660 to -1,731)

Scan QR to download the methodological references

Sara Conti[1,2], Matteo Spezia[2], Carlo Franceschini[2], Elvira Beato[3], Roberta Ciampichini[3], Giacomo Crotti[3], Dario Montermini[2,4], Lorenzo Giovanni Mantovani[1], Alberto Zucchi[3]

[1] Research Centre on Public Health, University of Milano – Bicocca, Monza, Italy
[2] P.G.M.D.Consulting srl, Milano, Italy (EHDEN accredited SME)
[3] Health Protection Agency of Bergamo (ATS Bergamo), Bergamo, Italy. (EHDEN data partner)
[4] CUTE srl, Milano, Italy

Sistema Socio Sanitario Regione Lombardia ATS Bergamo

THURSDAY

Analyzing the impact of COVID-19 on the healthcare system: an OMOP-CDM framework applied to Northern Italy

Lead: Sara Conti

#OHDSISocialShowcase This Week

Mapping UK Biobank to the OMOP CDM: challenges and solutions

PRESENTER: **Sofia Bazakou**

Background

UK Biobank[1] (UKB) is a large-scale registry containing medical and genetic data from 500,000 consented participants from the UK's general population, aged between 40 and 69 years (Figure 2). UKB is an extraordinary resource for human health research, accessible to approved research initiatives worldwide.

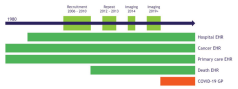


Figure 2. UKB data structure and timeline. The data include multiple baseline assessments (light green), such as surveys, samples, and imaging, linkages to electronic health records (EHR) from different sources (dark green), and information on COVID-19 testing (red). Picture adapted from Prof. Denaxas, UCL.

As part of the European Health Data Evidence Network[2] (EHDEN), The Hyve collaborated with University College London (UCL) to map the UKB data to the OMOP CDM v5.3. The Hyve performed the technical part of the mapping, whilst UCL provided the source data expertise.

The main goal of the collaboration was to make the dataset available for research related to the COVID-19 pandemic.

The UKB data conversion effort came with several challenges:

- Conversion of a large wide format table to long format. For each patient, a wide variety of variables and time points needed to be extracted from a 500,000 by 9,000 table.
- Large heterogeneity of source terms amongst data providers.
- Terms in free text or captured using a mix of ontologies (some of which have now been deprecated).
- Developing the ETL scripts relying entirely on synthetic data from the WhiteRabbit scan report.
- Working with an evolving data source.

Methods

We initially used White Rabbit to generate synthetic datasets. Next, we created the syntactic mappings to the OMOP CDM and the documentation with Rabbit in a Hat (RiaH), and the semantic mappings of source codes to standard concepts using Usagi (free text fields).

A powerful and flexible **ETL framework** together with existing **open-source tools** from the OHDSI suite allowed us to perform the conversion and to deliver a high-coverage mapping **without direct access** to the UKB data.

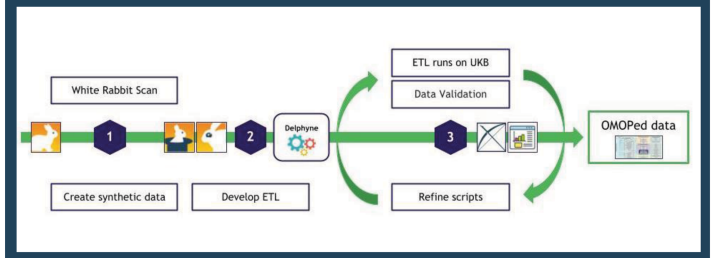


Figure 1. Data conversion workflow using existing OHDSI tools together with our internal ETL pipeline- Delphyne.

Scan QR to **see our blogpost.**

We implemented tests with the R testing framework functionality of RiaH. The code was then deployed at UCL to be executed on the original data. Lastly, we performed quality assessments with both Achilles and the Data Quality Dashboard (DQD).

Results

Our effort to convert the UK Biobank data to the OMOP CDM is an excellent example of successful collaboration and community engagement. Besides working with UCL to enable the UKB data to be used for research, the Hyve actively participated in the OHDSI community, by founding and leading the UKB working group, as well as initiating and taking part in OHDSI Forums discussions on issues that arose from the mapping and would benefit the community as a whole.

In total, the ETL codebase includes 35 table to table transformation scripts, and makes use of 19 OMOP vocabularies for the semantic mapping. The code has been executed successfully at UCL with the DQD currently achieving a 99% pass rate.

Such a high rate is a direct result of regular conversations with UCL, from which we gained useful insights for code adjustments, as well as the use of Delphyne's informative execution logs and summary reports to facilitate the investigation of data quality issues.

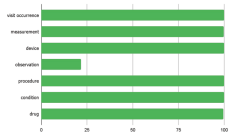


Figure 3. Percentage of UKB source codes mapped to a standard OMOP concept per domain by record frequency based on the full dataset). We achieved a near full or mapping coverage for most tables and the lowest mapping coverage was for the observation.

References

1. UK Biobank. Available from: https://www.ukbiobank.ac.uk/
2. European Health Data Evidence Network (EHDEN). Available from: https://www.ehden.eu/

Sofia Bazakou, Maxim Moinat, Alessia Peviani, Anne van Winzum, Stefan Payralbe, Vaclav Papez, Spiros Denaxas

The Hyve UCL EHDEN

FRIDAY

Mapping UKB to the OMOP CDM: Challenges and Solutions

Lead: Sofia Bazakou

Call For Participation

Call for Participation: Phenotyping Metadata Framework Survey

Researchers

mattspotnitz 3 18h

Dear OHDSI Friends,

We are looking for participants to fill out a brief survey to assess the comprehensiveness of a novel phenotyping metadata framework. As you may know, many electronic phenotype algorithms have been developed in the past decade. It is imperative to develop a metadata framework to catalogue and categorize these electronic phenotypes to help researchers better discern their differences and select suitable phenotypes for certain tasks. This metadata framework defines about 40 metadata elements. We would like researchers to grade how useful each data element is. The survey will take 5-10 mins to complete. We'd greatly appreciate your input and support by 08/25. We welcome anyone interested in participating on this project as a collaborator or co-author.

Here is a link to the survey:
https://cumc.co1.qualtrics.com/jfe/form/SV_8hUQx18rsDiGvhs 1

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?

OHDSI Speed Dating!

- Groups of 7-8 people
- 8-minute sessions
- 4 sessions total

What can you share?

OHDSI Speed Dating!

Name:

Organization:

Where have you been on your OHDSI journey?

Where do you want to be on the OHDSI journey?

Will you be coming to the OHDSI Symposium in October in Bethesda, Md., USA?

OHDSI Speed Dating!

+ 1 Special Question

Round 1: What's the OHDSI standard/tool/best practice you've most recently used?

pollev.com/patrickryan800

OHDSI Speed Dating!

+ 1 Special Question

Round 2: Where is the ideal location for the next OHDSI collaborative event?

pollev.com/patrickryan800

OHDSI Speed Dating!

+ 1 Special Question

Round 3: What is the best swag you've ever gotten at a scientific conference?

pollev.com/patrickryan800

OHDSI Speed Dating!

+ 1 Special Question

Round 4: You've decided to form a new OHDSI workgroup that you'll all be founding members of, what is it called?

pollev.com/patrickryan800