

# **How Did OHDSI** Do in 2025?

**OHDSI Community Call** Dec. 9, 2025 • 11 am ET









## **Upcoming Community Calls**

Date	Topic
Dec. 9	How Did OHDSI Do This Year?
Dec. 16	Holiday Farewell To 2025
Dec. 23	No Meeting
Dec. 30	No Meeting
Jan. 6	No Meeting
Jan. 13	Where Can We Go Together in 2026?









## Three Stages of The Journey

Where Have We Been? Where Are We Now? Where Are We Going?









### **OHDSI Shoutouts!**



Congratulations to the team of Patricia Pedregal-Pascual, Carlos Guarner-Argente, **Eng Hooi Tan, Asieh Golozar, Talita Duarte-**Salles, Andreas Weinberger Rosen, Antonella Delmestri, Wai Yi Man, Edward Burn, Daniel Prieto-Alhambra, and Danielle Newby on the publication of Incidence and Survival of Colorectal Cancer in the United Kingdom From 2000 to 2021: A Population-Based Cohort Study in the American Journal of Gastroenterology.

### Open

### Incidence and Survival of Colorectal Cancer in the United Kingdom From 2000 to 2021: A Population-Based Cohort Study

Patricia Pedregal-Pascual, MD1:2:\*, Carlos Guarner-Argente, PhD, MD1:\*, Eng Hooi Tan, PhD3, Asieh Golozar, PhD4:5 Talita Duarte-Salles, PhD67, Andreas Weinberger Rosen, PhD8, Antonella Delmestri, PhD3, Wai Yi Man, MSc3, Edward Burn, PhD3, Daniel Prieto-Alhambra, PhD, MD3,7 and Danielle Newby, PhD3

INTRODUCTION: The management of colorectal cancer (CRC) is evolving, with advances in screening and treatment. The

purpose of this study was to leverage population-based data to generate up-to-date UK estimates of agespecific and sex-specific incidence and overall survival for the period 2000-2021.

METHODS: We analyzed nationally representative primary care records from Clinical Practice Research Datalink GOLD and replicated in Clinical Practice Research Datalink Aurum. We calculated incidence rates, and

short-term and long-term survival stratified by age, sex, and diagnosis year.

RESULTS: Overall incidence was 67.4/100,000 person years, increasing in 2000-2011 to drop slightly in 2011-2014, and then plateauing. By contrast, early-onset CRC raised uninterruptedly throughout the

study period, from 8.33 to 19.07/100,000 person-years. Overall survival was 78.3%, 51.4%, and 38.5% at 1. 5, and 10 years, respectively, lower in men compared with women. Modest improvements in survival were observed over the study period, particularly for 60-69-year-old patients.

Although the overall incidence in the population has plateaued, a worrying increasing trend of earlyonset CRC was observed. Moreover, the slight improvement in overall survival suggests that significant

progress is still needed. These findings highlight the urgent need for continued research and resource allocation to improve the diagnosis and management of CRC.

KEYWORDS: colorectal cancer; incidence; cancer survival; population-based study

SUPPLEMENTARY MATERIAL accompanies this paper at http://links.lww.com/AJG/D630

Am J Gastroenterol 2025;120:2909-2917. https://doi.org/10.14309/ajg.000000000003460











## 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	12 pm	ATLAS/WebAPI
Tuesday	12 pm	Generative AI and Analytics
Tuesday	3 pm	Oncology Outreach/Research Subgroup
Wednesday	8 am	Psychiatry
Wednesday	9 am	Patient-Level Prediction
Wednesday	2 pm	Natural Language Processing
Wednesday	7 pm	Eye Care and Vision Research
Wednesday	7 pm	Medical Imaging
Thursday	7 am	Europe Community Call
Thursday	8 am	OpenEHR and OMOP
Thursday	9:30 am	Network Data Quality
Thursday	10 am	Rare Diseases
Thursday	10 am	Rehabilitation
Thursday	10:30 am	Evidence Network
Thursday	7 pm	Dentistry









### **Upcoming Workgroup Calls**



Date	Time (ET)	Meeting
Friday	9 am	Phenotype Development and Evaluation
Friday	10 am	GIS-Geographic Information System
Friday	10 am	Transplant
Friday	11 am	Clinical Trials
Friday	11:30 am	Steering
Friday	11 pm	China Chapter
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Common Data Model
Monday	11 am	Data Bricks User Group
Monday	2 pm	Electronic Animal Health Records
Tuesday	9 am	Data2Evidence









# India, APAC Symposiums











# December Newsletter is Available



### The Journey Newsletter (December 2025)

The final newsletter of 2025 focuses on the progress being made around global collaboration, including in local events that are taking place across four continents between November and December, New collaborators have multiple resources to help begin their journeys, including the Early-Stage Workgroup and a new tutorial homepage. Those who want to help educate the community can consider sharing plenary and tutorial proposals for the 2026 OHDSI Global Symposium. We look at all of that and more in the latest newsletter, #JoinTheJourney

### Podcast: Global Events, DARWIN EU Progress



highlight new community events in Sweden, Africa and Canada that took place in

November. They discuss the recent DARWIN EU update and its progress assisting

the EMA, and they announce a call for both plenaries and tutorials for the 2026

Global Symposium, which will take place Oct. 20-22 in New Brunswick, N.J. (If video

does not appear, please click 'view this email in your browser.)

### are, and to support regulatory decision-making. Leaders from the initiative provided a 2025 update recently. Where Are We Now?

• The 2025 OHDSI India Symposium will be held Dec. 2 at SVM Hospitals in Bangalore. The registration link and more information, including the symposium agenda, is now available.

**Community Updates** 

 November community calls were highlighted by numerous #OHDSI2025 presentations, including from our Best Community Contribution honorees and

three of our Early-Stage Researchers. All of the talks can be found in the

· The first Africa Symposium and new events in both Sweden and Canada

highlighted a busy November of collaboration opportunities. Learn more about

· DARWIN EU is a European Medicines Agency-led network that uses real-

world healthcare data, standardized to the OMOP CDM, across Europe to

generate evidence on how medicines are used, how safe and effective they

'November Presentations' section of this newsletter.

- The 2025 OHDSI Asia-Pacific Symposium will be held Dec. 6-7 in Shanghai, China. Registration is open, and agenda details will be shared when available.
- The Dec. 9 OHDSI Community Call will reflect on our community's impact in 2025. Workgroup leads have been asked to provide a brief self-assessment to assist that call by Dec. 5.

### Where Are We Going?

Where Have We Been?

each within this newsletter.

- The 2026 OHDSI Europe Symposium will be held April 18-20 in Rotterdam, Netherlands. While registration has not opened yet, the deadline for abstract submissions is Feb. 6, 2026.
- The 2026 OHDSI Global Symposium will be held Oct. 20-22 at the Hyatt Regency Hotel in New Brunswick, N.J. There is a call for plenary and tutorial submissions; more details can be found later in this newsletter.
- . The #OHDSISocialShowcase is currently highlighting research from the 2025 Europe Symposium, and will soon begin showcasing research from the Global Symposium. Please follow our LinkedIn, Twitter/X, Bluesky and Instagram feeds to learn more about the research happening in our

### **November Publications**

### Semaglutide and diabetic retinopathy: an OHDSI network study

Nicole Welskopf, <sup>8</sup> Hannah Morgan-Cooper, <sup>28</sup> Priya Desai, <sup>28</sup> Diep Tran, <sup>1</sup> Zainab Rustam, <sup>1</sup> Gina Zhu, <sup>1</sup> Joel Swerdel, <sup>28</sup> Anthony Sena, <sup>28,27</sup> Paul Nagy, <sup>2</sup> Marc Suchard, <sup>28,26</sup> Martiin Schuemie, <sup>28,36</sup> George Hripcsak, <sup>18</sup> Patrick Ryan <sup>18,26</sup>

Cai CX, Nishimura A, Baxter S, Goetz K, Hribar M, Toy B, Barkmeier A, Wang S, Swaminathan S, Flowers A, Brown E, Xu B, Chen J, Chen A, Leng T, Boland M, Alshammari T, Bu F, Falconer T, Martin B, Westlund E, Mathioudakis N, Zhang L, Fan R, Wilcox A, Lai A, Stocking JC, Xie Y, Lee LH, Dorr D, Humes I, McCov D, Adibuzzaman M, Areaux R Jr, Brash J, Weiskopf N, Morgan-Cooper H, Desai P, Tran D, Rustam Z, Zhu G, Swerdel J, Sena A, Nagy P, Suchard M, Schuemie M, Hripcsak G, Ryan P. Semaglutide and diabetic retinopathy: an OHDSI network study. BMJ Open Diabetes Res Care. 2025 Nov 4:13(6):e005424, doi: 10.1136/bmidrc-2025-005424, PMID: 41192935; PMCID:

Datzmann T, Lang C, Tesch F, Spoden M, Dröge P, Ehm F, Schuler E, Krogias C. Günster C. Schmitt J. Gumbinger C. Barlinn J. Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction. Sci Rep. 2025 Nov 7;15(1):38994. doi: 10.1038/s41598-025-25979-1. PMID: 41203731; PMCID: PMC12594769.

Cuyàs B. Alvarado-Tapias E. Tan EH. Golozar A. Duarte-Salles T. Delmestri A. Argemi J, Man WY, Burn E, Guarner-Argente C, Prieto Alhambra D, Newby D. Trends in incidence, prevalence, and survival of primary liver cancer in the United Kingdom (2000-2021). Eur J Public Health. 2025 Nov 10:ckaf153. doi: 10.1093/eurpub/ckaf153. Epub ahead of print, PMID: 41212072.

Finster M, Wenzel M, Taghizadeh E. Common data models and data standards for tabular health data: a systematic review. BMC Med Inform Decis Mak. 2025 Nov 13:25(1):422. doi: 10.1186/s12911-025-03267-2. PMID: 41233809: PMCID: PMC12616946.

Guerrero P, Ernebjerg M, Holst T, Weese D, DiBello H, Ibing S, Schmidt L, Ungaro R, Renard B, Lippert C, Alleva E, Quinn TD, Kovatch P, Antao EM, Heyneke E, Rasheed A, Kalabakov S, Arnrich B, Charney A, Wieler LH, Nadkarni G. The AIR-MS data platform for artificial intelligence in healthcare. JAMIA Open. 2025 Nov 11;8(6):ooaf145. doi: 10.1093/jamiaopen/ooaf145. PMID: 41267854: PMCID: PMC12629540.

2026 Global Symposium Dates Set; Community Seeks Plenary, Tutorials By Jan. 30, 2026





The 12th annual OHDSI Global Symposium will return to the Hyatt Regency Hotel in New Brunswick, N.J., Oct. 20-22, 2026. All pertinent information will be added to this page when available. Currently, the OHDSI steering group is seeking proposals for both plenaries and tutorials. The deadline for both is January 30, 2026.

Symposium plenaries provide opportunities to share innovative, communitydeveloped content to empower researchers to generate reliable real-world evidence. The community is currently seeking proposals for our #OHDSI2026 plenaries. These sessions will be 60 minutes in duration and must touch on at least two of following pillars of our community: open community data standards, methodological research, open-source development, and clinical applications.

Plenary sessions must also involve three or more on-stage participants across at least two organizations. Sessions may include a combination of keynote talks, panel discussions, interactive activities, and more. We strongly encourage using multiple formats and synthesizing completed research. current perspectives and future calls-to-action to maximize community engagement.

Tutorial sessions aim to deliver educational content, led by community members who wish to train our global collaborators on scientific, technical, and other skills that can support advancing OHDSI's mission and the effective use of real-world data and the generation and dissemination of reliable real-world evidence. Tutorial sessions are 4 hours in duration. Sessions may include a combination of talks, interactive activities, and more. We strongly encourage using multiple formats to maximize community engagement. Your session must include at least three people from at least two different organizations.

Learn more and share your submissions before the Jan. 30 deadline using the links below.

Plenary/Tutorial Submission Details

Submit Your OHDSI 2026 Plenary Proposal

Submit Your OHDSI 2026 Tutorial Proposal

mailchi.mp/ohdsi/december2025













# December Newsletter is Available



mailchi.mp/ohdsi/december2025











# 2026 Global Symposium

The 2026 OHDSI Global Symposium will return to the Hyatt Regency Hotel in New Brunswick, N.J., on Oct. 20-22.











# 2026 Global Symposium

### 2026 OHDSI Global Symposium Call for Plenary Sessions

Symposium plenaries provide opportunities to share innovative, community-developed content to empower researchers to generate reliable realworld evidence. The community is currently seeking proposals for our #OHDSI2026 plenaries. These sessions will be 60 minutes in duration and must touch on at least two of following pillars of our community:

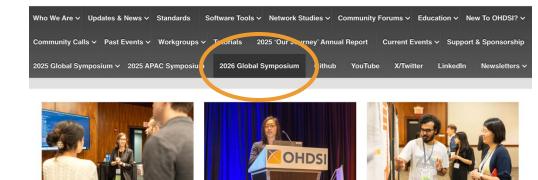
- · Open community data standards
- Methodological research
- Open-source development
- Clinical applications

Plenary sessions must also involve three or more on-stage participants across at least two organizations. Sessions may include a combination of keynote talks, panel discussions, interactive activities, and more. We strongly encourage using multiple formats and synthesizing completed research, current perspectives and future calls-to-action to maximize community engagement.

The deadline for proposal submissions is January 30, 2026. Please use the link below to submit your proposal by answering the following questions:

- Name(s) of plenary session organizers:
- · Your email address(es):
- Short (2,500 character max) description / abstract of your proposed session:
- · Which pillars are you targeting:
- One sentence "pitch" of your session to excite the community:
- Names and roles of individuals who have tentatively agreed to participate in your session:

**Deadline to submit** proposals for #OHDSI2026 plenaries or tutorials is Jan. 30, 2026!



### 2026 OHDSI Global Symposium

Oct. 20-22 · New Brunswick, N.J. · Hyatt Regency Hotel

### 2026 OHDSI Global Symposium Call for Tutorials

Tutorial sessions aim to deliver educational content, led by community members who wish to train our global collaborators on scientific, technical, and other skills that can support advancing OHDSI's mission and the effective use of real-world data and the generation and dissemination of reliable real-world evidence. Examples of prior tutorials offered are provided here: https://www.ohdsi.org/tutorials.

Tutorial sessions are 4 hours in duration. Registrants for your tutorial will be requested to pay a registration fee. The fees will be used to offset the costs of the symposium and other OHDSI expenses. Sessions may include a combination of talks, interactive activities, and more. We strongly encourage using multiple formats to maximize community engagement. Your session must include at least three people from at least two different organizations.

The deadline for tutorial proposal submissions is January 30, 2026. Please use the link below to submit your proposal by answering the following questions:

- · Name(s) of tutorial session organizers:
- Your email address(es):
- Short (2,500 character) description / abstract of your proposed session:
- Names and roles of individuals who have tentatively agreed to participate in your session:













# 2026 Europe Symposium

The 2026 OHDSI Europe Symposium returns to Rotterdam next year and will be held April 18-20.

The deadline for abstract submissions will be Feb. 6, 2026.









### Monday

**Characterizing the OHDSI Evidence** Network – A Global **Snapshot of Real-World Data Partners** 

(Clair Blacketer, Evanette Burrows, Ben Gerber, Vojtech Huser, Paul Nagy)

Characterizing the OHDSI Evidence Network: A Global Snapshot of Real-World Data Partners

♣ PRESENTER: Clair Blacketer

- Reliable real-world evidence (RWE) requires diverse, high-quality data
- · Federated research networks. preserve patient privacy and data control while enabling
- · US (PCORnet, Sentinel) and EU examples (EHDEN, DARWIN EU) have shown the power of this
- OHDSI Evidence Network launched in 2024 to build a global. open, opt-in federated network.
- · Understanding the scope, diversity, and structure of the participating data sources is essential for evaluating the network's capacity to support a broad research agenda.

### METHODS

- · Partners generate DbProfiles comprised of aggregate summary statistics using Database Diagnostics R package
- DbProfiles include: demographics. domain coverage, observation time, record density
- · Aggregate data characterized by:
- · Source type (EHR, claims, other)
- · Visit setting (outpatient, inpatient, ER, etc.)
- Population size
- · Domain completeness

- 48 databases from 28 partners across 4 continents
- Database types: 75% EHRs, 20% claims, 5% registries/other
- · Visit settings: ~98% outpatient, 84% inpatient; half report combined ER+inpatient visits

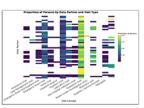
The **OHDSI Evidence Network** brings together a diverse, global collection of 48 real-world data sources, supporting scalable, open, and federated research





### RESULTS, CONT:

- Population size: 2.6k → 829M across databases; mean 53M, median 5.3M
- Gender distribution: 52% female overall
- Domain coverage
- · Condition & observation domains nearly universally available in all
- Device least represented by proportion of persons
- Death domain lowest coverage; 7 databases with no death data



### DISCUSSION

- · The OHDSI Evidence Network is a diverse set of global real-world data
- It is well positioned to support
- international research questions. Heterogeneity of the network underscores the need to generate
- evidence using multiple data sources. One data source only tells part of the

Next steps: better communicate the breadth of data available in the network to the community, work on database similarity and quality assessments

Clair Blacketer<sup>1,2,4</sup>, Evanette Burrows<sup>1,4</sup>, Ben Gerber<sup>1,5</sup>, Voitech Huser<sup>1,6</sup>, Paul Nagy<sup>1,7</sup>















### **Tuesday**

**Real-World** Implementation of the Medical Imaging CDM: An Alzheimer's **Disease Use Case** 

(Woo Yeon Park, Teri Sippel Schmidt, **Blake Dewey, Paul Nagy)** 

Real-World Implementation of the Medical Imaging CDM: An Alzheimer's Disease Use Case

♣ PRESENTER: Jen Park

- · Healthcare data is siloed. Clinical outcomes live in the EHR. Imaging details live in PACS or VNA. Without integration, we can't fully understand patients, build reproducible phenotypes, or scale multimodal
- · The Medical Imaging Common Data Model (MI-CDM) bridges this gap by bringing standardized imaging metadata into the OMOP CDM, making images computable alongside labs, meds, and diagnoses.
- · Institutions unlock the value of millions of images already sitting in

### METHODS

- · Population: We included patients aged 50 years or older with at least one clinical encounter within the Johns Hopkins health system. The clinical data for the cohort were already mapped to a standard OMOP CDM instance but lacked imaging features.
- · Imaging Data: We extracted a sample of 1M DICOM Series for the cohort. The extraction and transformation steps are modality-agnostic.
- Imaging Concepts: We harvested a complete list of DICOM tags and select coded values from DICOM Standard.

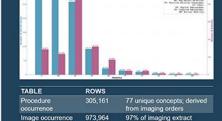
From Imaging files	Counts	
Patients	108,250	
DICOM Studies	311,810	
DICOM Series	999,972	
Average Series per Patient	9	
DICOM metadata elements	73,965,730	
Average elements per Series	73	
Most popular imaging modalities	CT, MR, SR PR, KO, PT	

Large-scale real-world implementation of MI-CDM demonstrates **feasibility**, **scalability**, and support for computable imaging phenotype



### **DATA ETL & LINKING RESULTS**

5183 tags, 3628 coded



### PHENOTYPE GENERATION





JHU Results: 13,924 patients; 47,977

### Lessons Learned

The real-world implementation of MI-CDM depends on how your institution has set up OMOP CDM. What works smoothly at one site might need extra steps at another.

### Considerations:

- Data Governance: Make sure you have IRB approval that covers linking imaging data to clinical tables like Procedure occurrence. If you only have de-identified DICOM headers from your PACS or VNA, integration can be difficult. Since OMOP tables don't store raw IDs. you'll also need intermediate registry tables or local mappings to bridge DICOM identifiers with clinical
- ETL Environment: Even though MI-CDM focuses on DICOM headers and not pixel data, the metadata can still be very large. Extracting, cleaning, transforming, and loading requires storage and compute power. Partner early with your Research IT team so you can plan for the right environment and resources.
- Procedure to DICOM Files: If the Procedure occurrence is derived from billing codes, then you should strategize how you can map a billing code to imaging files. We included the imaging orders to maintain the data lineage

### Future Studies:

- Data validity and variability check: study how well imaging metadata aligns with the official DICOM Standard
- Private tags: reassess excluded private attributes and test if they capture research-relevant details
- Image findings: incorporate vocabularies like RadLex or CDEs to describe what's
- Free-text mapping: standardize long text fields (e.g., series description) to improve searchability and reproducibility

- Breaking data silos: incorporating the DICOM imaging standard into the OMOP CDM to enable multimodal research doi:10.1093/jamia/ocaf091
- https://github.com/paulnagy/DICOM2O MOP/tree/main/files/OMOP%20CDM%2

Jen Park, Teri Sippel Schmidt, Blake Dewey, Paul Nagy















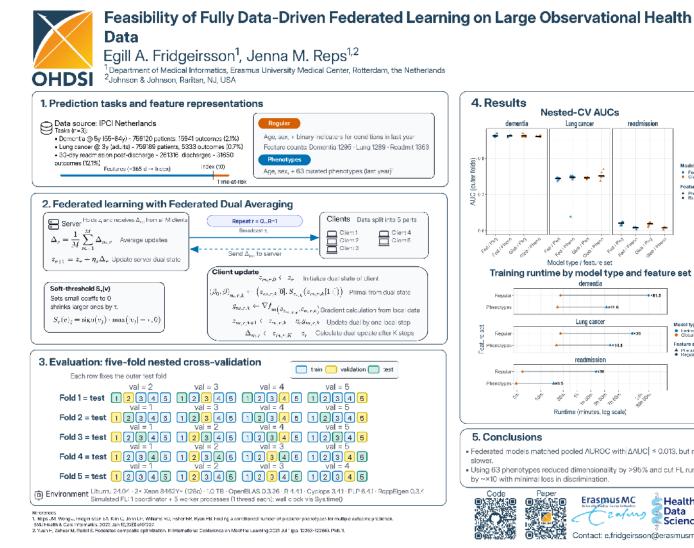
Concept



### Wednesday

**Feasibility of Fully Data-Driven Federated Learning** on Large **Observational Health Data** 

(Egill Fridgeirsson, Jenna Reps)





Nested-CV AUCs

Lung cancer







Feature set





### **Thursday**

**Characterizing Patients** with Type 2 Diabetes **Mellitus Treated with Anti-Diabetic Medication:** A Feasibility Study to **Enable Future Multi-Database Application** 

(Jiwon Um, Subin Kim, Seng Chan You)

Characterizina Patients with Type 2 Diabetes Mellitus Treated with Anti-Diabetic Medication:

A Feasibility Study to Enable Future Multi-Database Application

♣ PRESENTER: Jiwon Um

### Global Burden:

- · 537 million people with T2DM; China and India account for 40% of global cases.
- · Yet the evidence base remains predominantly Western-derived

### Evidence Gan:

· Major diabetes RCTs enroll only 24% non-White participants, leaving regional patterns underrepresented

- Aligned with one of the 2025 OHDSI APAC study led by Peking Univ.
- "Gastrointestinal Risk of GLP-1 Receptor Agonist versus SGLT-2 and DPP-4 Inhibitor in Type 2 Diabetes".

· Validate cohort feasibility and

### Data Source

 YUHS OMOP CDM v. 5.4 (2006-2023) ≥ 6.4 million longitudinal records.

### Study population

- Adults (≥18 years) with at least 365 days of prior observation and at least one
- Index/exposure: ≥ 90 days prior metformin use; no prior use to GLP-1RA, DPP-4i, or SGLT-2i before index.
- · Exclusions: T1DM or secondary diabetes.

 Outcomes assessed for feasibility were gastroparesis, intestinal obstruction, non-alcoholic fatty liver disease (NAFLD), and acute pancreatitis.

### Feasibility Today, Multi-Database Evidence Tomorrow: A collaborative path to bridge the T2DM evidence gap









APAC 2025 Studies-Your join accelerate

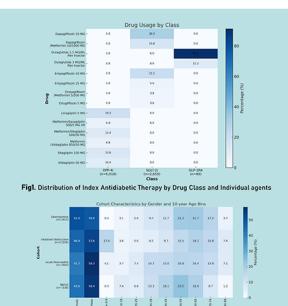


Fig 2. Demographic Composition of Clinical Outcome Cohorts

### Statistical Analysis

 We used CohortDiagnostics to evaluate phenotype validity, internal consistency, and feasibility, leveraging its standardized diagnostics for cohor characterization, misclassification detection, and cross-database

### RESULTS:

- Analytic consistency: Primary ITT and sensitivity PP analyses produced identical sample counts in both exposure groups
- Exposure profile: Initiation clustered around was concentrated in dapagliflozin/empagliflozin (SGLT-2i), sitagliptin/linagliptin (DPP-4i), and dulaglutide (GLP-1RA)
- DPP-4i markedly predominant (N=6,018), supporting an activecomparator design.
- Cohort demographic/validity: The number of patients in the GLP-1RA target cohort was relatively small. which warrants further discussion in future analyses.

### CONCLUSION:

- This is a preliminary analysis based on a single-source dataset (YUHS), so cross-country comparisons and generalization are limited. Additiona multi-site, multi-country data are needed to better characterize patterns across outcomes and settings
- In collaboration with the OHDSI APAG Network, multi-country diabetesrelated studies are underway, with three planned for presentation at the 2025 OHDSI APAC Symposium.

- This research was supported by the National Research of Korea (NRK) grant funded by the Korea government (MIST) (No. RS-2023-00263899).
- Jiwon Um<sup>1,2</sup>, Subin Kim<sup>1,2</sup>, Seng Chan You<sup>1,2</sup>
- in Dept. Of Biomedical Systems Informatics, Yonse niversity College of Medicine, Seoul, Korea <sup>2</sup>Yonsei Institute for Digital Health, Yonsei













## **Friday**

# **Barista: Brewing A New** Methodology for **Governing Study** Execution

(Ajit Londhe, Martin Lavallee, Katy Sadowski, Carmen Ng, Casey Tilton) Barista: Brewing A New Methodology for Governing Study Execution

▲ PRESENTER: Ajit Londhe

- · Within a study, it is challenging to collect, maintain, and correctly leverage assets, such as concept sets and cohort definitions.
- Atlas and Capr are useful for designing those study assets, but the identification and lifecycle of each asset at scale is a gap.
- · CohortGenerator provides cohort organization, subsets, and definition hashing, but the complexity of a study repository can grow as a result.
- Ulvsses establishes a framework for organized study execution but is limited in its governance and enforcement of a standardized
- We sought to address these challenges with the new R package

### METHODS

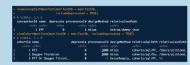
We established the idea of "manifests" as lists to track all study assets with the following principles:

- 1. Cohort and concept set ids are meaningless: they need to be unique and consistent.
- 2. Accurate manifest management is only possible with a strongly governed system.
- 3. Ulysses is the basis for how the study repository is organized and
- 4. A study asset can be added to a manifest but can never be modified: it can be deprecated but
- CohortGenerator is the exclusive method for generating cohorts.
- 6. Metadata on each study asset is critical to transparency and quality

Barista will help ensure your study repository pours over all concept sets, cohort definitions, and analyses to reduce the protocol grind and achieve that perfect brew!



Architecture diagram of a Barista SQLite Database and the Manifest tables within, each with types of assets that could be maintained



within SQLite manifest tables in Barista. Each manifest type follows assumptions to ensure quality and transparency of all study



Dependencies are handled via the Dependency Manifest, with graphs used to not only illustrate the the full impact of an asset

 Ajit Londhe, Martin Lavallee, Katy Sadowski, Carmen Ng, Casey Tilton



RESULTS

Barista leverages R6 classes to

· A SOLite database is used to

Any item added to a manifest is

ensuring no false assumptions

Instead, all actions taken with a

about study asset content or

manifest are explicit and

Barista integrates with

Analyses, Migrations, Dependencies, and Tagging.

intentional via user-facing

CohortGenerator via an active

binding of a cohort definition set

Within a Barista database, there are

· Concept Sets, Cohorts, File Paths,

· The Dependency manifest allows

between assets to be created

· The Tag Manifest supports a oneto-many approach, so that any

for a graph of logical dependencies

manifest item can be tagged with a name-value pair that can be used

With the first principle rendering

ids meaningless, it is important to

provide efficiencies for identifying

study assets during a study to provide transparency and avoid

Barista is currently in development,

but it will introduce a new paradigm

and a practical implementation for

delivering results from an OHDSI study repository.

maintaining study assets and

overreliance on ids.

generation.

functions

Manifest tables for

for filtering

permanent but can be deprecated,

handle primary keys and constraints

coordinate all workflows







QR code to visit the Barista GitHub Repo!













## 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?









## Dec. 9: How Did We Do In 2025?

















## The weekly OHDSI community call is held every Tuesday at 11 am ET.

**Everybody** is invited!

Links are sent out weekly and available at: ohdsi.org/community-calls-2025





