Unlock Causal Insights with Evidence Synthesis

OHDSI Community Call
May 6, 2025 • 11 am ET

Upcoming Community Calls

Date	Topic
May 6	Evidence Synthesis
May 13	Maternal Health Fellowship Review
May 20	Guideline-Driven Evidence Study Review
May 27	Collaborator Showcase Brainstorm (Deadline is July 1)

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

Congratulations, Dr. Tom Seinen

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
Wednesday	8 am	Psychiatry
Thursday	9:30 am	Network Data Quality
Thursday	10 am	Rare Diseases
Thursday	10:30 am	Evidence Network
Thursday	7 pm	Dentistry
Friday	9 am	Phenotype Development and Evaluation
Friday	10 am	GIS-Geographic Information System
Friday	11 am	Clinical Trials
Friday	11:30 am	Steering
Friday	11 pm	China Chapter
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Africa Chapter
Monday	10 am	The Getting Started Subgroup
Monday	12 pm	Book of OHDSI

ATLAS Usage & Feedback Survey

Atlas Survey

General

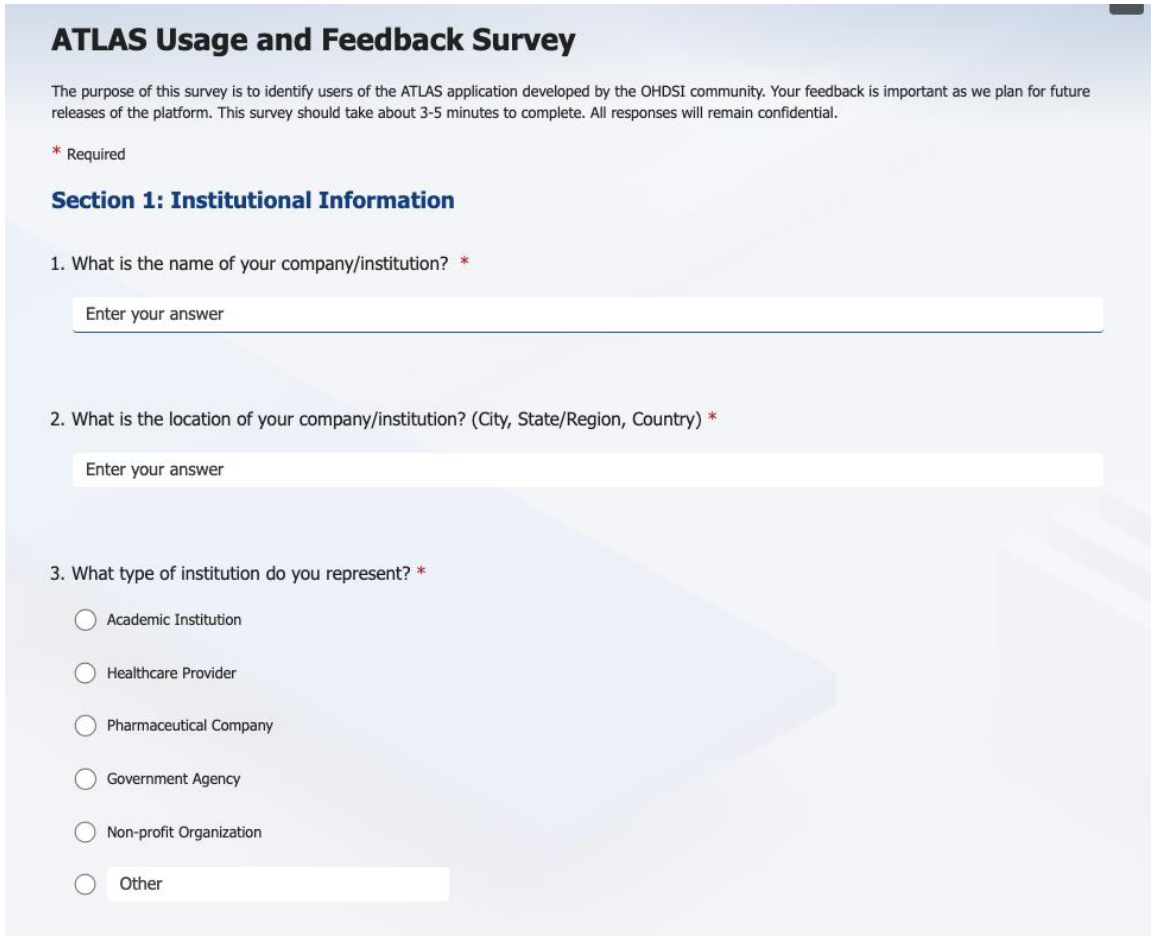
Chris_Knoll 5d

The ATLAS working group has put together a short survey (Microsoft Forms 6) to help us identify who is using ATLAS in our community. If you are not using ATLAS, we'd still ask you to fill in this survey so you can help us to identify any barriers for adoption in your company/institution.

Additionally, this survey will ask if you'd like to be interviewed for feedback on your usage of ATLAS. Data4Life (https://www.data4life.care/ 1) is working closely with our working group to conduct interviews (~1hr) that will help inform the future direction of the application.

We'd appreciate if you could fill out this survey and consider speaking with Data4Life regarding your experiences with ATLAS.

Tagging @anthonysena



ATLAS Usage and Feedback Survey

The purpose of this survey is to identify users of the ATLAS application developed by the OHDSI community. Your feedback is important as we plan for future releases of the platform. This survey should take about 3-5 minutes to complete. All responses will remain confidential.

* Required

Section 1: Institutional Information

1. What is the name of your company/institution? *

Enter your answer

2. What is the location of your company/institution? (City, State/Region, Country) *

Enter your answer

3. What type of institution do you represent? *

- Academic Institution
- Healthcare Provider
- Pharmaceutical Company
- Government Agency
- Non-profit Organization
- Other

Link on community calls page

Latest Newsletter is Available

OHDSI

The Journey Newsletter (May 2025)

Open-source tools are developed and maintained within the community to create a pathway for OHDSI's global research mission. Several tools—and the overall open-source environment—were a main focus last month, and they are highlighted in this newsletter. We also look at upcoming collaboration opportunities, 14 published studies in April, and more in the latest newsletter. #JoinTheJourney

Podcast: Open-Source Tools, Ecosystem

Analysis April was the theme last month, so Patrick Ryan and Craig Sachson discussed some of the tools that were featured in recent community calls, including Strategus and Treatment Patterns. They also looked at recent research around estimation and prediction, and they discuss the value of the fourth annual DevCon, which was held in late April. *(if video does not appear, please click 'View this email in your browser')*

Community Updates

Where Have We Been?

- Our fourth annual DevCon was held April 25. This event brought together developers and innovators to explore the latest tools, technologies, and strategies shaping the future of open-source software in healthcare and data science. All three sessions, as well as the introductory talk by **Paul Nagy,** are all available on the event page.
- April community calls highlighted numerous tools that were developed and are being used by the community for global research, including Strategus and TreatmentPatterns. They also highlighted some of the methods and tools being innovated for both estimation and prediction research. You can find those presentations at the bottom of this newsletter or on our community calls page.

Where Are We Now?

- The ATLAS working group has put together a short survey to help identify who is using ATLAS in our community. If you are not using ATLAS, please also fill in this survey to help identify any barriers for adoption in your company/institution. Additionally, this survey will ask if you'd like to be interviewed for feedback on your usage of ATLAS.
- **Christian Reich** and **Sarah Seager** are leading an effort to publish a second edition of the Book of OHDSI, which will include updates to previous text, as well as new material. This work is taking place within the Education workgroup; if you would like to join this effort, please sign up here.
- The Industry Workgroup is hosting a two-day in-person studyathon, though it may also have a virtual component, May 29-30 in the Gilead Office in Troy Hills, N.J. You can register here, and more information the event is available here.

Where Are We Going?

- **The submission deadline for the** 2025 Collaborator Showcase **is July 1.** The showcase will be accepting both posters and software demos, as well as interest in hosting lightning talks. Details on the collaborator showcase are now available, and you can submit your brief report(s) here.
- The Columbia Summer School on OHDSI, which will be held July 14-18 at Columbia University, provides health professionals, researchers, and industry practitioners with an immersive, hands-on training to working with real-world health data and generating real-world evidence (RWE). Participants will explore the types of healthcare data captured during routine clinical care—such as electronic health records and administrative claims—and learn how to standardize these data using the OMOP Common Data Model to support

From Health Data to Discovery: Learn Real-World Evidence Generation at Columbia University Summer School on OHDSI

Leaders within the OHDSI community will host a Summer School in Observational Health Data Science & Informatics, AI, and Real World Evidence this July 14-18 within the Department of Biomedical Informatics at Columbia University in New York City.

George Hripcsak, Patrick Ryan, Anna Ostropolets and Karthik Natarajan will serve as faculty for this session, which will be limited to 30 participants to ensure a high level of interaction and personal support. Registration is now open for the event, which you can access using the button below.

The Columbia OHDSI Summer School provides health professionals, researchers, and industry practitioners with an immersive, hands-on training to working with real-world health data and generating real-world evidence (RWE). Participants will explore the types of healthcare data captured during routine clinical care—such as electronic health records and administrative claims—and learn how to standardize these data using the OMOP Common Data Model to support collaborative, distributed research as part of a data network.

Columbia Summer School on OHDSI Details/Registration

April Publications

Pedregal-Pascual P, Guarner-Argente C, Tan EH, Golozar A, Duarte-Salles T, Rosen AW, Delmestri A, Man WY, Burn E, Prieto-Alhambra D, Newby D. Incidence and survival of colorectal cancer in the United Kingdom from 2000-2021: a population-based cohort study. Am J Gastroenterol. 2025 Apr 1. doi: 10.14309/ajg.0000000000003460. Epub ahead of print. PMID: 40167184.

Sanz F. Integrative Data Science in Drug Safety Research: Experiences, Challenges, and Perspectives. Annu Rev Biomed Data Sci. 2025 Apr 1. doi: 10.1146/annurev-biodatasci-103123-095506. Epub ahead of print. PMID: 40169005.

Lee M, Lee KJ, Kim J, Lee DY, Park RW, Rhee SY, Cha JM, Yang HJ, Jang JW, Jung S, Lee J, Lee SH, Kim C, Bae JS, Kim YJ, Lee JH, Bae H, Kim Y. Low-density lipoprotein cholesterol levels and risk of incident dementia: a distributed network analysis using common data models. J Neurol Neurosurg Psychiatry. 2025 Apr 1:jnnp-2024-334708. doi: 10.1136/jnnp-2024-334708. Epub ahead of print. PMID: 40169354.

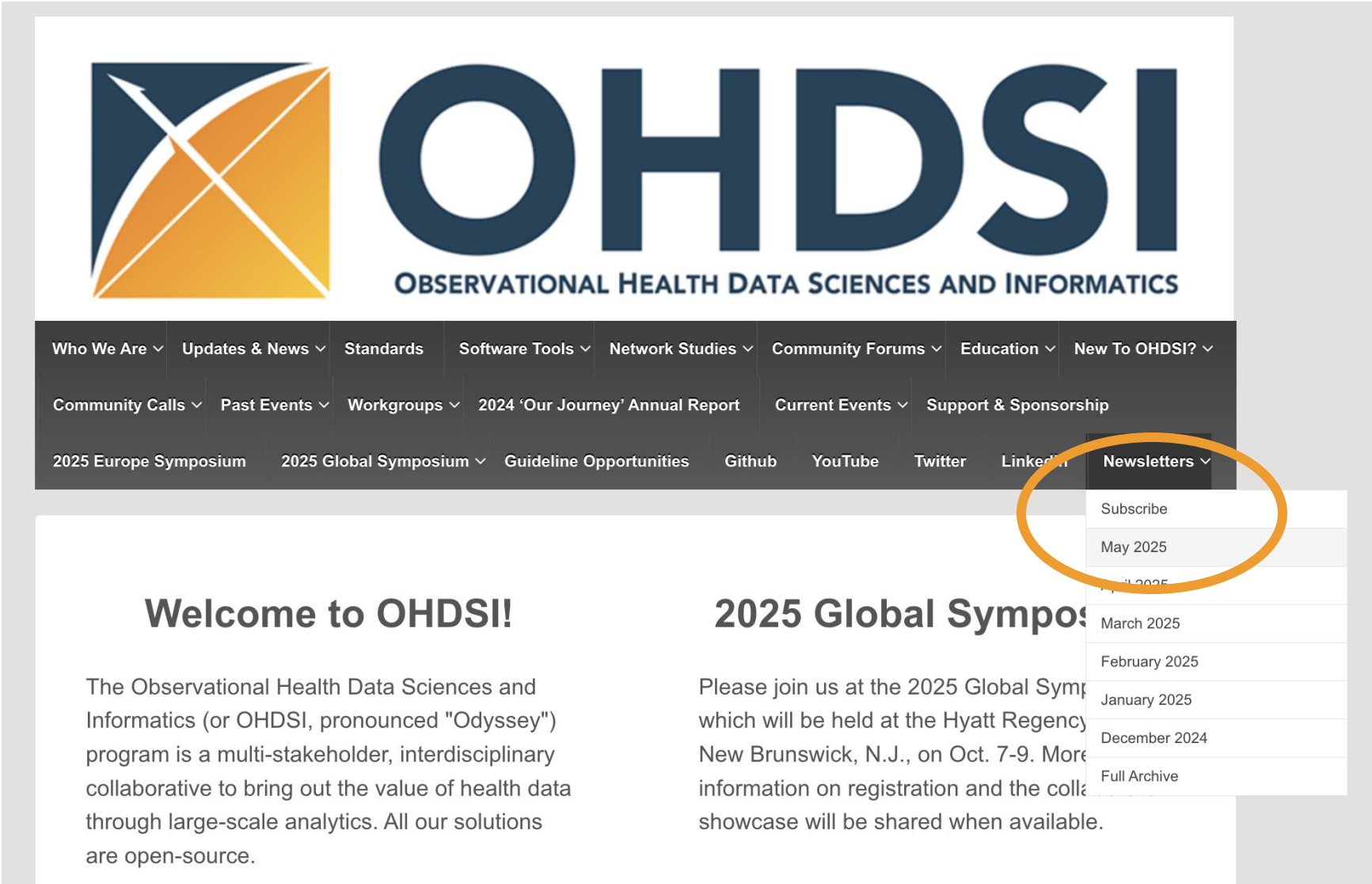
Cox S, Masood E, Panagi V, Macdonald C, Milligan G, Horban S, Santos R, Hall C, Lea D, Tarr S, Mumtaz S, Akashili E, Rae A, Urwin E, Cole C, Sheikh A, Jefferson E, Quinlan PR. Conversion of Sensitive Data to the Observational Medical Outcomes Partnership Common Data Model: Protocol for the Development and Use of Carrot. JMIR Res Protoc. 2025 Apr 2;14:e60917. doi: 10.2196/60917. PMID: 40173432; PMCID: PMC12004012.

Wang L, Wen A, Fu S, Ruan X, Huang M, Li R, Lu Q, Lyu H, Williams AE, Liu H. A scoping review of OMOP CDM adoption for cancer research using real world data. NPJ Digit Med. 2025 Apr 7;8(1):189. doi: 10.1038/s41746-025-01581-7. PMID: 40189628; PMCID: PMC11973147.

Hallaj S, Halfpenny W, Radgoudarzi N, Boland MV, Swaminathan SS, Wang SY, Xu BY, Amarasekera DC, Stagg B, Chen A, Hribar M, Thakoor KA, Goetz KE, Myers JS, Lee AY, Christopher MA, Zangwill LM, Weinreb RN, Baxter SL. Gap Analysis of Standard Automated Perimetry Concept Representation in Medical Terminologies. J Glaucoma. 2025 Apr 8. doi: 10.1097/IJG.0000000000002575. Epub ahead of print. PMID: 40193233.

Henke E, Lorenz S, Zoch M, Sedlmayr M, Peng Y. Mapping National Vocabularies to International Standards Using OHDSI Standardized Vocabularies. Stud Health Technol Inform. 2025 Apr 8;323:349-353. doi:

Latest Newsletter is Available



2025 OHDSI Focus Areas

Guideline-driven Evidence Generation	Evidence-driven Data standardization	Evidence-driven Open Source Development	Evidence-driven Collaborative Education

Dry January: Guideline review to determine evidence needs where RWE could potentially contribute	**Phenotype Phebruary:** Develop/evaluate cohorts needed to support filling the evidence gaps	**March to Data Fitness:** Evidence network to determine which partners are appropriate to generate which evidence
Analysis April: Prepare protocol and analysis specification to initiate network execution	**Meta-analysis May:** Collaborative interpretation of results from across network	**Journey to June:** Mid-year reflection on evidence generation process and progress
Spread-the-Word Second Half: Focus on Evidence Dissemination		
July: OHDSI EU	August:	September:
October: OHDSI Global (tbd)	November:	December: OHDSI APAC

🎙️ **Monthly Ask Me Anything (AMA)**
Presenting, Mui in May

📣 The **Women of OHDSI** workgroup is thrilled to feature **Mui Van Zandt** as May's guest responder in our AMA series!

💬 Got questions about OHDSI, data, career paths, or just curious about Mui's journey?

Now's the time to ask — literally anything!
✅ Just keep it:
Clean
Positive
And flowing!

👇 Post your questions below and let's get the conversation going!

The Center for Advanced Healthcare Research Informatics (CAHRI) at Tufts Medicine welcomes:

Georgie Kennedy, PhD

Senior Research Fellow, University of New South Wales and the Ingham Institute

'Learning from Real-World Cancer Data: Maturing data pipelines to support research that impacts clinical care'

May 29, 2025, 9-10am EDT

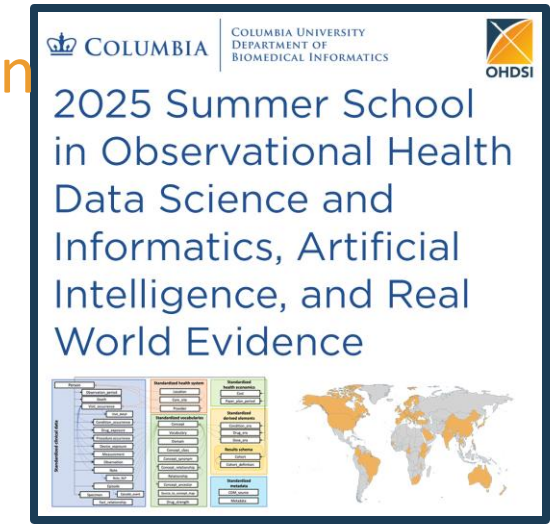
Virtually via Zoom

Please contact Marty Alvarez at malvarez2@tuftsmedicalcenter.org for calendar invite or questions.

TuftsMedicine
Tufts Medical Center

Columbia Summer School on OHDSI

Registration is open for the first ever Columbia Summer School on OHDSI, held July 14-18, 2025,
at the Columbia University Department of Biomedical Informatics in New York City.



The Columbia Summer School in Observational Health Data Science and Informatics, Artificial Intelligence, and Real World Evidence (RWE) offers health professionals, researchers and industry practitioners the opportunity to gain familiarity and hands-on experience with real world data and generating real world evidence. Participants will learn about the different types of healthcare data captured during routine clinical care, including electronic health records and administrative records, and how these data can be standardized to the OMOP Common Data Model to enable distributed data network research.

Meet Our Faculty

George Hripcsak, MD MS
Vivian Beaumont Allen Professor of Biomedical Informatics

Patrick Ryan, PhD
Adjunct Assistant Professor of Biomedical Informatics

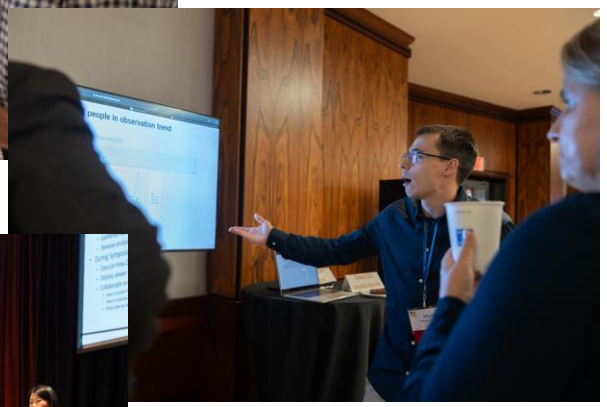
Anna Ostropolets, MD PhD
Adjunct Assistant Professor of Biomedical Informatics

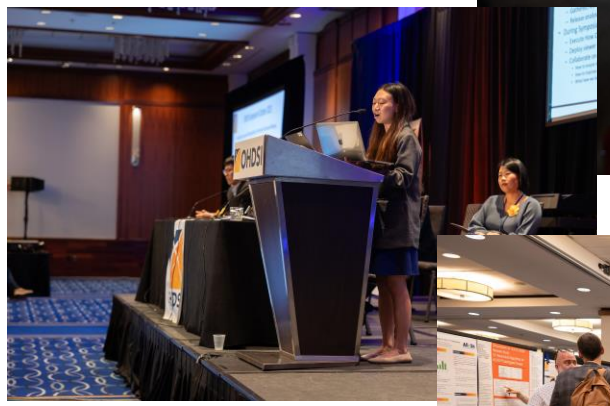
Karthik Natarajan, PhD
Assistant Professor of Biomedical Informatics

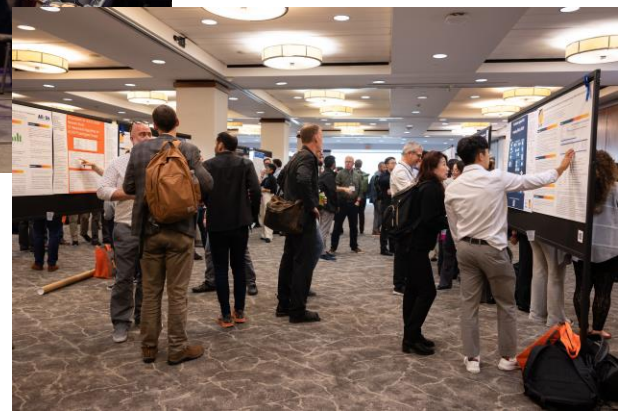
Save The Date!

The submission deadline for the 2025 Global Symposium Collaborator Showcase is July 1.

The showcase will be accepting both posters and software demos, as well as interest in hosting lightning talks. More information on the symposium, including abstract submission and registration links, will be available soon.





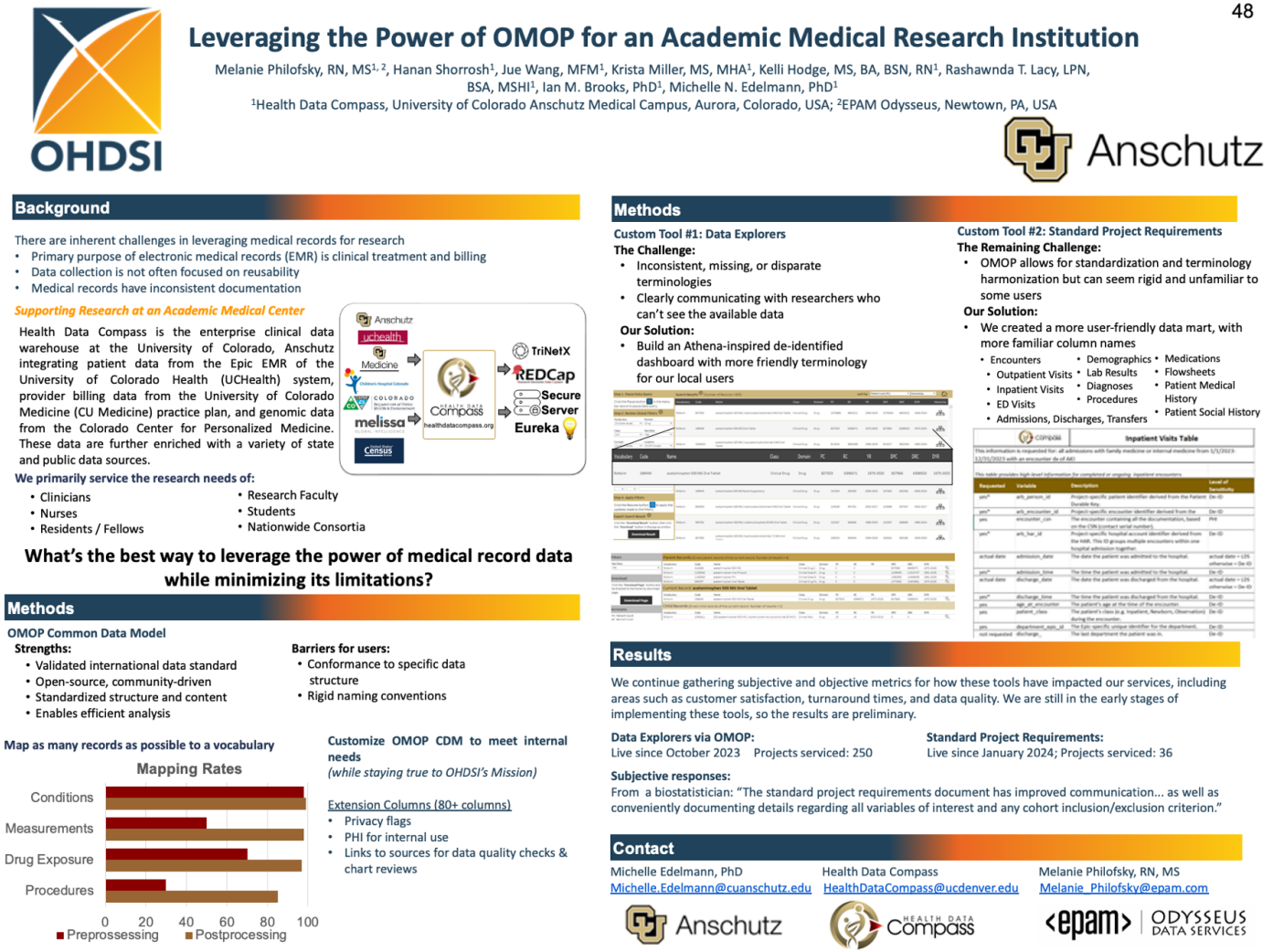


#OHDSISocialShowcase This Week

Monday

Leveraging the Power of OMOP for an Academic Medical Research Institution

(Melanie Philofsky, Hanan Shorrosh, Jue Wang, Krista Miller, Kelli Hodge, Rashawnda T. Lacy, Ian M. Brooks, Michelle N. Edelmann)



48

Leveraging the Power of OMOP for an Academic Medical Research Institution

Melanie Philofsky, RN, MS[1, 2], Hanan Shorrosh[1], Jue Wang, MFM[1], Krista Miller, MS, MHA[1], Kelli Hodge, MS, BA, BSN, RN[1], Rashawnda T. Lacy, LPN, BSA, MSHI[1], Ian M. Brooks, PhD[1], Michelle N. Edelmann, PhD[1]
[1]Health Data Compass, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA; [2]EPAM Odysseus, Newtown, PA, USA

Background

There are inherent challenges in leveraging medical records for research
- Primary purpose of electronic medical records (EMR) is clinical treatment and billing
- Data collection is not often focused on reusability
- Medical records have inconsistent documentation

Supporting Research at an Academic Medical Center

Health Data Compass is the enterprise clinical data warehouse at the University of Colorado, Anschutz integrating patient data from the Epic EMR of the University of Colorado Health (UCHealth) system, provider billing data from the University of Colorado Medicine (CU Medicine) practice plan, and genomic data from the Colorado Center for Personalized Medicine. These data are further enriched with a variety of state and public data sources.

We primarily service the research needs of:
- Clinicians
- Nurses
- Residents / Fellows
- Research Faculty
- Students
- Nationwide Consortia

What's the best way to leverage the power of medical record data while minimizing its limitations?

Methods

OMOP Common Data Model

Strengths:
- Validated international data standard
- Open-source, community-driven
- Standardized structure and content
- Enables efficient analysis

Barriers for users:
- Conformance to specific data structure
- Rigid naming conventions

Map as many records as possible to a vocabulary

Customize OMOP CDM to meet internal needs
(while staying true to OHDSI's Mission)

Extension Columns (80+ columns)
- Privacy flags
- PHI for internal use
- Links to sources for data quality checks & chart reviews

Methods

Custom Tool #1: Data Explorers
The Challenge:
- Inconsistent, missing, or disparate terminologies
- Clearly communicating with researchers who can't see the available data

Our Solution:
- Build an Athena-inspired de-identified dashboard with more friendly terminology for our local users

Custom Tool #2: Standard Project Requirements
The Remaining Challenge:
- OMOP allows for standardization and terminology harmonization but can seem rigid and unfamiliar to some users

Our Solution:
- We created a more user-friendly data mart, with more familiar column names
  - Encounters
    - Outpatient Visits
    - Inpatient Visits
    - ED Visits
    - Admissions, Discharges, Transfers
  - Demographics
  - Lab Results
  - Diagnoses
  - Procedures
  - Medications
  - Flowsheets
  - Patient Medical History
  - Patient Social History

Results

We continue gathering subjective and objective metrics for how these tools have impacted our services, including areas such as customer satisfaction, turnaround times, and data quality. We are still in the early stages of implementing these tools, so the results are preliminary.

Data Explorers via OMOP:
Live since October 2023 Projects serviced: 250

Standard Project Requirements:
Live since January 2024; Projects serviced: 36

Subjective responses:
From a biostatistician: "The standard project requirements document has improved communication... as well as conveniently documenting details regarding all variables of interest and any cohort inclusion/exclusion criterion."

Contact

Michelle Edelmann, PhD
Michelle.Edelmann@cuanschutz.edu
Health Data Compass
HealthDataCompass@ucdenver.edu
Melanie Philofsky, RN, MS
Melanie_Philofsky@epam.com

#OHDSISocialShowcase This Week

Tuesday

Characterizing the Temporality of OMOP CDM Concepts in a Mastectomy Phenotype

(Matthew Spotnitz, Yechiam Ostchega, Stephanie L. Goff, Lakshmi Priya Anandan, Emily Clark, John Giannini, Lew Berman)

Characterizing the Temporality of OMOP CDM Concepts in a Mastectomy Phenotype

Matthew Spotnitz, MD, MPH[1], Yechiam Ostchega, PhD, RN[1], Stephanie L. Goff, MD[2], Lakshmi Priya Anandan, MPH[1], Emily Clark, MPH[1], John Giannini, PhD[1], Lew Berman, PhD, MS[1]
1. *All of Us* Research Program, Office of the Director, National Institutes of Health, Bethesda, MD; 2. Surgery Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD

Background

Standardized approaches to characterizing temporal electronic health record (EHR) data are not well established[1].

Phenotype

The first occurrence of an OMOP CDM Mastectomy procedure concept.

NCCN Phases

Mastectomy Related Concepts

↓

Concept Sets

↓

NCCN Phases

Contact e-mail: matthew.spotnitz@nih.gov

Results

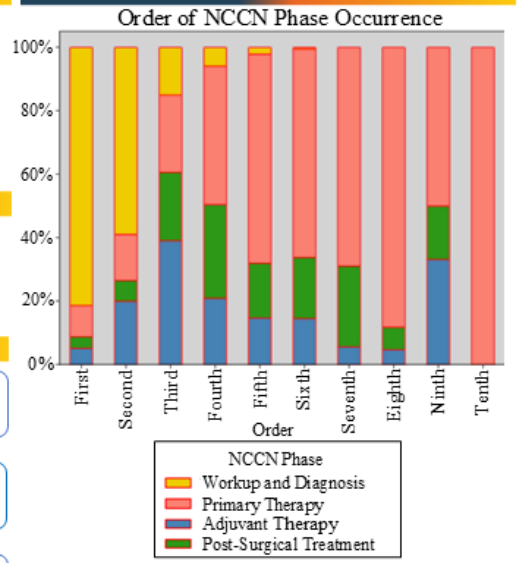


Figure 1: Stacked Bar Chart of OMOP CDM Mastectomy Related Concept Sets by National Comprehensive Cancer Network (NCCN) Phase and Sequence Order. Data Source: The *All of Us* Research Program, Controlled Tier Dataset, Version 7 Release.

Discussion

- The inability to differentiate neoadjuvant from adjuvant therapy based on concept names was a limitation of our analysis.
- The consistency of our temporal patterns with NCCN guidelines was indeterminate.
- Other representations of temporal information may be alternatives to bar charts.

Conclusions

- We have presented a novel method for characterizing temporal patterns with OMOP CDM concepts.
- Further refinement and research with OHDSI software packages may be warranted.

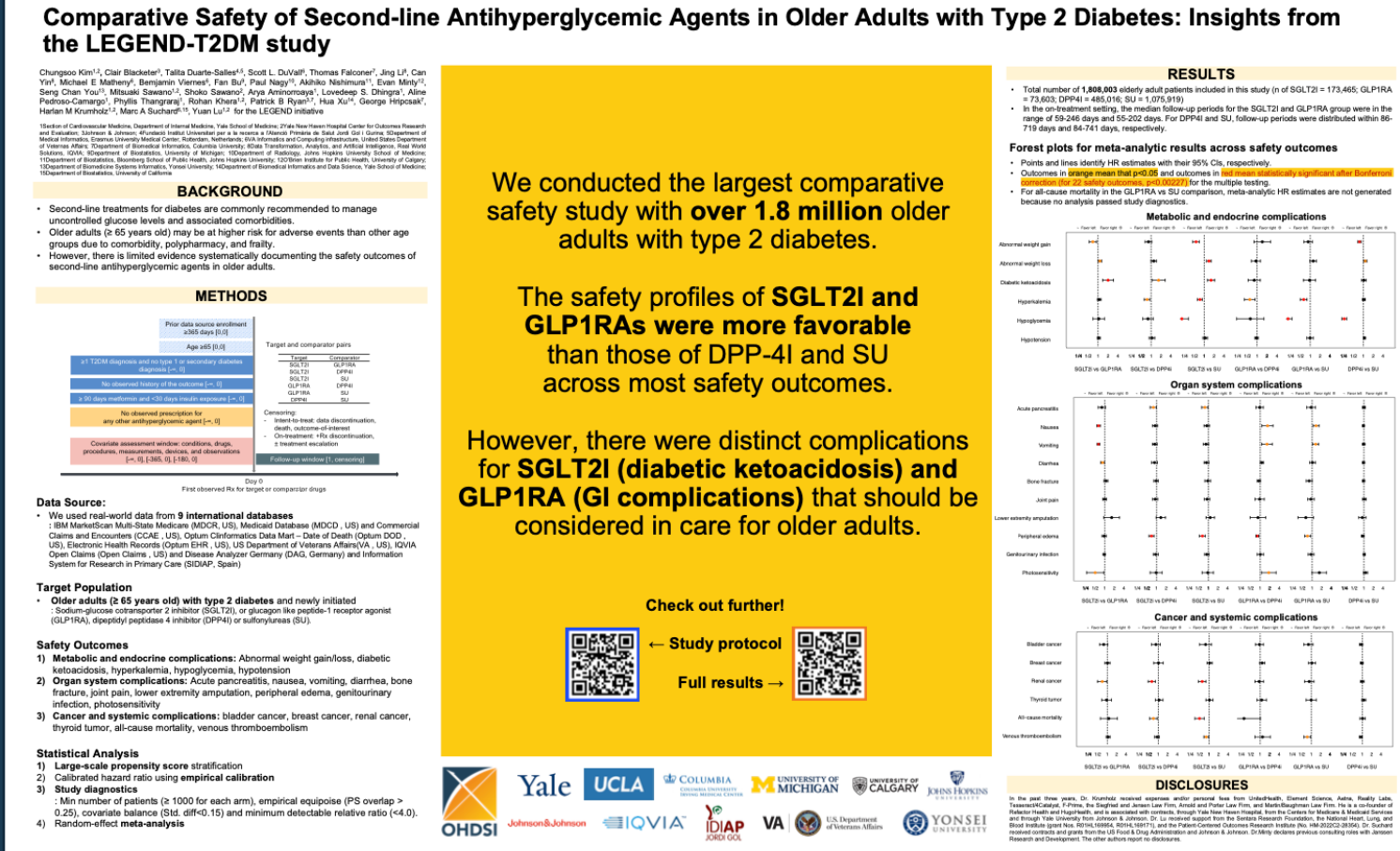
References:
[1] Berman L, Ostchega Y, Giannini J, et. al. Application of a Data Quality Framework to Ductal Carcinoma in Situ Using Electronic Health Record Data from the All of Us Research Program. JCO Clinical Cancer Informatics. 2024 Aug:8:e2400052. doi: 10.1200/CCI.24.00052

#OHDSISocialShowcase This Week

Wednesday

Comparative Safety of Second-line Antihyperglycemic Agents in Older Adults with Diabetes: Insights from the LEGEND-T2DM study

(Chungsoo Kim, Clair Blacketer, Talita Duarte-Salles, Scott DuVall, Thomas Falconer, Jing Li, Can Yin, Michael Matheny, Bemjamin Viernes, Fan Bu, Paul Nagy, Akihiko Nishimura, Evan Minty, Seng Chan You, Mitsuaki Sawano, Shoko Sawano, Arya Aminorroaya, Lovedeep Dhingra, Aline Pedroso-Camargo, Phyllis Thangraraj, Rohan Khera, Patrick Ryan, Hua Xu, George Hripcsak, Harlan Krumholz, Marc Suchard, Yuan Lu)



#OHDSISocialShowcase This Week

Thursday

The Impact of Evolving Diagnostic Guidelines on Clinical Characterization of Endometriosis

(Harry Reyes Nieva , Aparajita Kashyap, Erica A. Voss, Anna Ostropolets, Adit Anand, Mert Ketenci, Frank J. DeFalco, Karthik Natarajan, Young Sang Choi, Yanwei Li, Monica N. Allen, Stephanie Guang, Noémie Elhadad)

OHDSI

The Impact of Evolving Diagnostic Guidelines on Clinical Characterization of Endometriosis

Harry Reyes Nieva, MPhil, MAS, MA,[1,2] Aparajita Kashyap, MA,[1] Erica A. Voss, MPH,[3] Anna Ostropolets, MD, PhD,[1,3] Adit Anand,[1] Mert Ketenci, MA,[4] Frank J. DeFalco,[3] Karthik Natarajan, PhD,[1] Young Sang Choi, MS, MA,[1] Yanwei Li,[1] Monica N. Allen, MD,[5] Stephanie Guang, MD,[5] Noémie Elhadad, PhD[1,4]

[1] Department of Biomedical Informatics, Columbia University; [2] Department of Medicine, Harvard Medical School; [3] Global Epidemiology, Janssen Research & Development, LLC; [4] Department of Computer Science, Columbia University; [5] Department of Obstetrics and Gynecology, Columbia University

Background

- Endometriosis is a chronic estrogen-dependent inflammatory condition with no known etiology, no diagnostic biomarker, and no cure to date.
- Marked by a heterogeneous set of systemic symptoms across women, and often involves chronic pain, dysmenorrhea, dyspareunia, dysuria, and fatigue.
- In 2022, clinical diagnosis guidelines shifted from diagnostic surgery to a more multimodal approach, emphasizing assessment of indicative symptoms and use of diagnostic imaging as complement to surgery.
- Shift in clinical guidelines acknowledges high variability in diagnosing endometriosis and aims to enhance accuracy and expedite diagnosis. Also suggests prior characterizations may not accurately reflect full composition, care patterns, and spectrum of experiences associated with endometriosis.
- This study investigates the impact of changes in clinical guidelines on phenotype definitions and the corresponding patient cohorts considered for diagnosis.

Methods

- Observational cohort study of women aged 15-49 years diagnosed with endometriosis between January 2013 and December 2023, with at least one year of continuous observation prior to cohort entry.
- Data sources included United States (US) insurance claims from the Merative™ MarketScan® Commercial Database (CCAE), Merative™ MarketScan® Multi-State Medicaid Database (MDCD), Optum® de-identified Electronic Health Record (Optum® EHR) data set, and Columbia University Irving Medical Center electronic health record (CUIMC EHR).
- Created five cohort definitions of women diagnosed with endometriosis according to several phenotype definitions:
  - **Cohort A**: surgical confirmation
  - **Cohort B**: imaging and guideline-recognized symptoms
  - **Cohort C**: guideline-recognized symptoms
  - **Cohort D**: guideline-recognized symptoms and/or pelvic pain
  - **Cohort E**: guideline-recognized symptoms, pelvic pain, and/or abdominal pain
- Two gynecological experts conducted chart review on a sample (n=100) from the CUIMC EHR dataset. Performance of phenotype definition assessed according to precision, recall, and F1 scores.
- Tested pairwise differences between cohorts using X^2 tests and z-tests with Bonferroni correction to account for multiple comparisons.

Acknowledgments

The research was supported in part by grants from the National Library of Medicine and National Heart, Lung, and Blood Institute at the National Institutes of Health (T15-LM007079 [HRN, AK, AA]; R01HL148248 [MK]; R01LM013043 [NE]) and a Computational and Data Science Fellowship from the Association for Computing Machinery Special Interest Group in High Performance Computing (HRN).

Results

Overall, 491,048 women with endometriosis across all data sources (Fig. 1). Only 15-20% of cases were captured by all 5 phenotypes. More than one-fourth (26-30%) were only captured by symptom-based definitions (Cohorts C-E). Study findings were largely consistent across each data source.

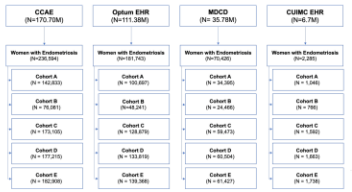


Figure 1. STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) Flow Diagram of Women Diagnosed with Endometriosis Based on Surgery, Imaging, and/or Symptom Presentation.

Table 1. Participant Characteristics, Optum® EHR Dataset

Characteristic	Cohort A (N = 100,697)	Cohort B (N=48,241)	Cohort C (N = 128,879)	Cohort D (N = 133,819)	Cohort E (N = 139,368)
Age Group					
15-19 years	1,290 (1.3%)	1,169 (2.4%)	2,827 (2.2%)	2,937 (2.2%)	3,023 (2.2%)
20-24 years	5,051 (5.0%)	4,242 (8.8%)	9,568 (7.4%)	10,024 (7.5%)	10,362 (7.4%)
25-29 years	9,785 (9.7%)	7,239 (15%)	16,600 (13%)	17,397 (13%)	18,028 (13%)
30-34 years	15,720 (16%)	9,504 (20%)	23,381 (18%)	24,276 (18%)	25,120 (18%)
35-39 years	20,619 (20%)	10,006 (21%)	27,294 (21%)	28,391 (21%)	29,546 (21%)
40-44 years	24,973 (25%)	9,376 (19%)	27,549 (21%)	28,459 (21%)	29,784 (21%)
45-49 years	23,259 (23%)	6,705 (14%)	21,660 (17%)	22,335 (17%)	23,505 (17%)
Race					
Asian	1,976 (2.0%)	866 (1.8%)	2,155 (1.7%)	2,246 (1.7%)	2,370 (1.7%)
Black / African American	11,231 (11%)	5,726 (12%)	14,875 (12%)	15,308 (11%)	15,866 (11%)
White	81,420 (81%)	38,597 (80%)	103,353 (80%)	107,365 (80%)	111,729 (80%)
Unknown	6,070 (6.0%)	3,052 (6.3%)	8,496 (6.6%)	8,900 (6.7%)	9,403 (6.7%)
Ethnicity					
Hispanic / Latina	6,460 (6.4%)	3,403 (7.1%)	8,520 (6.6%)	8,810 (6.6%)	9,114 (6.5%)
Not Hispanic / Latina	87,584 (87%)	41,713 (86%)	111,929 (87%)	116,076 (87%)	120,767 (87%)
Unknown	6,653 (6.6%)	3,125 (6.5%)	8,430 (6.5%)	8,933 (6.7%)	9,487 (6.8%)

Figure 2. Precision, Recall, and F1 Scores of Endometriosis Cohort Definitions

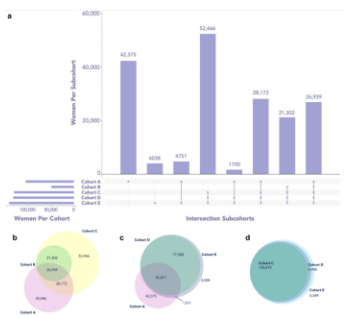


Figure 3a-d. Patient Set Overlap Among Study Cohorts in Optum® EHR Dataset. a) UpSet plot illustrating intersections across all cohorts. b) Venn diagram of Cohorts A, B, and C. c) Venn diagram of Cohorts A, D, and E. d) Venn diagram of Cohorts C, D, and E.

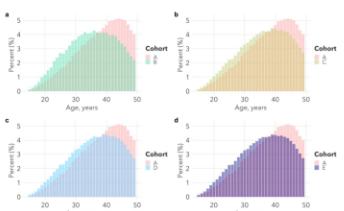


Figure 4a-d. Age at Diagnosis of Endometriosis in the Optum® EHR Dataset. On average, women in Cohort A were diagnosed with endometriosis at an older age than women in Cohorts B-E (mean age difference Cohort A vs Cohorts C-E = 2 years; p<0.001 for all comparisons).

Table 2. Healthcare Utilization (Optum® EHR Dataset)

Encounter Type	Characteristic	Cohort A (N = 100,697)	Cohort B (N=48,241)	Cohort C (N = 128,879)	Cohort D (N = 133,819)	Cohort E (N = 139,368)
Outpatient	n (%)	100,362 (100)	47,755 (99)	126,060 (98)	130,889 (98)	136,191 (98)
	mean (SD)	17.4 (16.7)	20.5 (18.8)	17.1 (17.6)	16.8 (17.4)	16.4 (17.3)
	median (IQR)	13 (6-23)	15 (8-27)	12 (5-23)	12 (5-22)	11 (5-22)
ER Visits	n (%)	30,560 (30)	19,738 (41)	43,170 (34)	44,118 (33)	44,949 (32)
	mean (SD)	0.8 (3.2)	1.4 (4.3)	1 (3.9)	1 (3.8)	0.9 (3.8)
	median (IQR)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)
Hospitalizations	n (%)	45,459 (45)	12,287 (26)	32,564 (25)	33,599 (25)	34,830 (25)
	mean (SD)	0.6 (1.1)	0.4 (1.8)	0.4 (1.5)	0.4 (1.5)	0.4 (1.4)
	median (IQR)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)

Abbreviations: ER, emergency room. IQR, interquartile range. SD, standard deviation.

Symptoms and Treatment

- Pain was the most common symptom (91% of A, 100% of B-E). Localized pain of abdomen and pelvis (69%-70% in A, 83%-90% in B-E) were documented without mention of guideline-based symptoms in 2-5% of cases (14,795 women).
- Among guideline-recommended treatments, women in B-E were more likely to receive hormones and related agents (44% in A, 61% in B, 50-51% in C-E). Opioid use was very common (96% in A, 82% in B, 77-78% in C-E).

Age at Diagnosis and Date of Cohort Entry

- Compared to women with diagnoses based on surgical confirmation, those diagnosed by imaging and/or symptoms were younger (Fig. 4; mean age = 38 years [SD = 8] in A, 35 years [SD = 9] in B; 36 [SD = 8] in C-E; p<0.001 for all).
- We found little-to-no difference in date of cohort entry among patients identified by multiple definitions (median difference compared to A = 0 days for B-E, IQR = 0-45 days for B, 0-15 days for C, 0-14 days for D, and 0-13 days for E).

Healthcare Utilization

- Women diagnosed based on imaging and symptoms made more ER visits (Table 2; 41% of women in B vs 30% in A, p<0.001; mean visits = 1.4, SD = 4.3 in B vs mean visits = 0.8, SD=3.2 in A; p<0.001).
- While nearly half of women (45%) in A had at least one hospitalization prior to diagnosis (mean = 0.6, SD=1.1), this was the case for only one-fourth (26%) in B (mean = 0.4, SD=1.8). Similar trends held when comparing C-E to A.

Discussion

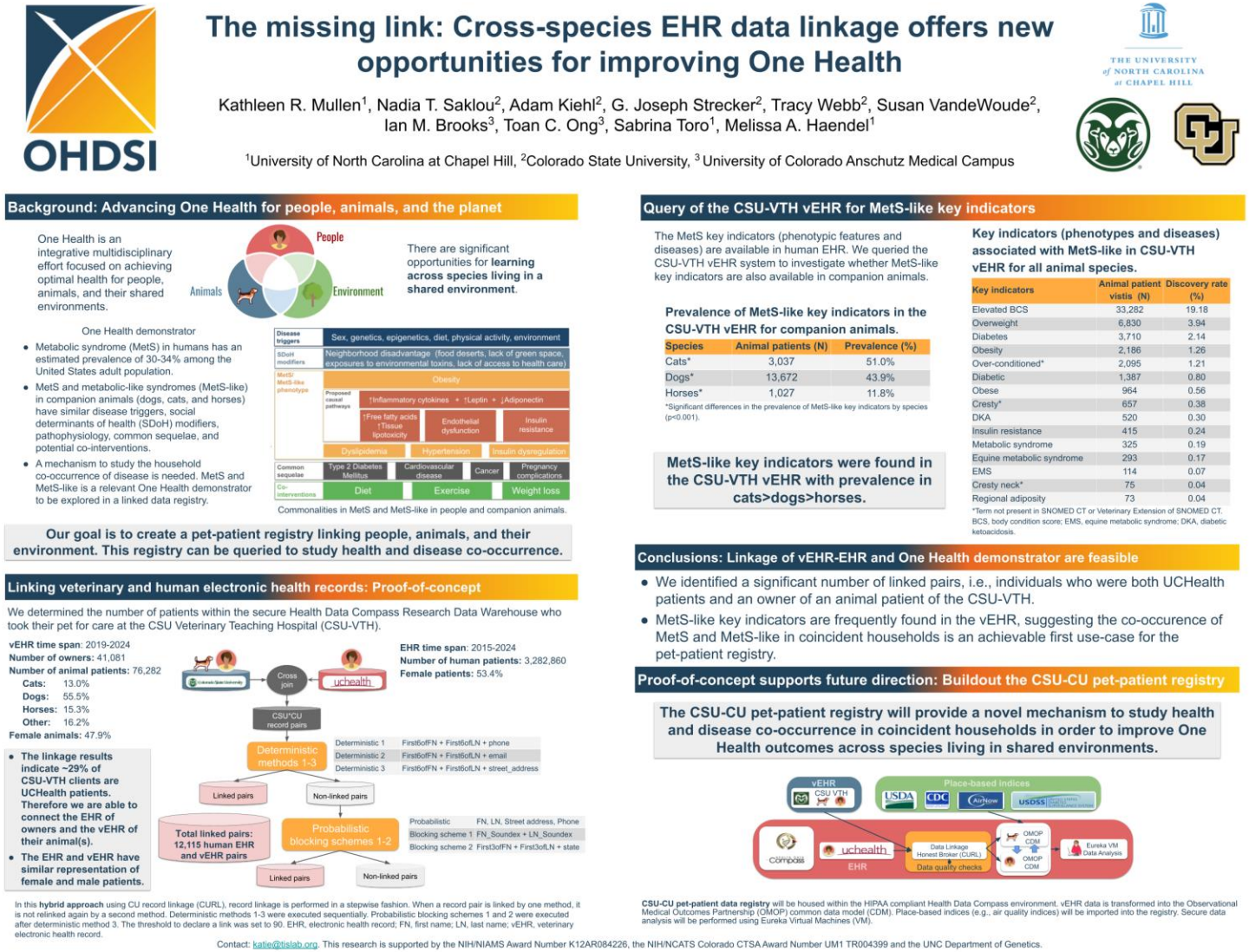
- Similar precision among all five cohort definitions despite poor overlap among women identified illustrates both the heterogeneous presentation of the disease and importance of expanding diagnostic criteria. For example:
  - Cohorts derived from updated guidelines identified younger patients at time of diagnosis.
  - Women diagnosed based on imaging had higher rates of ER visits while patients diagnosed via laparoscopy had a larger number of hospitalizations.
- A sizable percentage of women presented with only pelvic and/or abdominal pain and none of the guideline-recognized symptoms, underscoring continued need for improved access to timely and appropriate care, particularly among those with non-classical symptoms, different care-seeking patterns, or lack of available surgical intervention.
- Limitations associated with secondary use of claims and EHR data should be considered, including reliance on codes that do not necessarily capture all diagnostic guideline criteria.

#OHDSISocialShowcase This Week

Friday

The missing link: Cross-species EHR data linkage offers new opportunities for improving One Health

(Kathleen R. Mullen, Nadia T. Saklou, Adam Kiehl, G. Joseph Strecker, Tracy Webb, Susan VandeWoude, Ian M. Brooks, Toan Ong, Sabrina Toro, Melissa Haendel)



The missing link: Cross-species EHR data linkage offers new opportunities for improving One Health

Kathleen R. Mullen[1], Nadia T. Saklou[2], Adam Kiehl[2], G. Joseph Strecker[2], Tracy Webb[2], Susan VandeWoude[2], Ian M. Brooks[3], Toan C. Ong[3], Sabrina Toro[1], Melissa A. Haendel[1]

[1]University of North Carolina at Chapel Hill, [2]Colorado State University, [3] University of Colorado Anschutz Medical Campus

Background: Advancing One Health for people, animals, and the planet

One Health is an integrative multidisciplinary effort focused on achieving optimal health for people, animals, and their shared environments.

People, Animals, Environment

There are significant opportunities for learning across species living in a shared environment.

One Health demonstrator

- Metabolic syndrome (MetS) in humans has an estimated prevalence of 30-34% among the United States adult population.
- MetS and metabolic-like syndromes (MetS-like) in companion animals (dogs, cats, and horses) have similar disease triggers, social determinants of health (SDoH) modifiers, pathophysiology, common sequelae, and potential co-interventions.
- A mechanism to study the household co-occurrence of disease is needed. MetS and MetS-like is a relevant One Health demonstrator to be explored in a linked data registry.

Commonalities in MetS and MetS-like in people and companion animals.

Our goal is to create a pet-patient registry linking people, animals, and their environment. This registry can be queried to study health and disease co-occurrence.

Linking veterinary and human electronic health records: Proof-of-concept

We determined the number of patients within the secure Health Data Compass Research Data Warehouse who took their pet for care at the CSU Veterinary Teaching Hospital (CSU-VTH).

vEHR time span: 2019-2024
Number of owners: 41,081
Number of animal patients: 76,282
Cats: 13.0%
Dogs: 55.5%
Horses: 15.3%
Other: 16.2%
Female animals: 47.9%

EHR time span: 2015-2024
Number of human patients: 3,282,860
Female patients: 53.4%

- The linkage results indicate ~29% of CSU-VTH clients are UCHealth patients. Therefore we are able to connect the EHR of owners and the vEHR of their animal(s).
- The EHR and vEHR have similar representation of female and male patients.

Total linked pairs: 12,115 human EHR and vEHR pairs

In this hybrid approach using CU record linkage (CURL), record linkage is performed in a stepwise fashion. When a record pair is linked by one method, it is not relinked again by a second method. Deterministic methods 1-3 were executed sequentially. Probabilistic blocking schemes 1 and 2 were executed after deterministic method 3. The threshold to declare a link was set to 90. EHR, electronic health record; FN, first name; LN, last name; vEHR, veterinary electronic health record.

Query of the CSU-VTH vEHR for MetS-like key indicators

The MetS key indicators (phenotypic features and diseases) are available in human EHR. We queried the CSU-VTH vEHR system to investigate whether MetS-like key indicators are also available in companion animals.

Prevalence of MetS-like key indicators in the CSU-VTH vEHR for companion animals.

Species	Animal patients (N)	Prevalence (%)
Cats*	3,037	51.0%
Dogs*	13,672	43.9%
Horses*	1,027	11.8%

*Significant differences in the prevalence of MetS-like key indicators by species ($p<0.001$).

MetS-like key indicators were found in the CSU-VTH vEHR with prevalence in cats>dogs>horses.

Key indicators (phenotypes and diseases) associated with MetS-like in CSU-VTH vEHR for all animal species.

Key indicators	Animal patient vistis (N)	Discovery rate (%)
Elevated BCS	33,282	19.18
Overweight	6,830	3.94
Diabetes	3,710	2.14
Obesity	2,186	1.26
Over-conditioned*	2,095	1.21
Diabetic	1,387	0.80
Obese	964	0.56
Cresty*	657	0.38
DKA	520	0.30
Insulin resistance	415	0.24
Metabolic syndrome	325	0.19
Equine metabolic syndrome	293	0.17
EMS	114	0.07
Cresty neck*	75	0.04
Regional adiposity	73	0.04

*Term not present in SNOMED CT or Veterinary Extension of SNOMED CT. BCS, body condition score; EMS, equine metabolic syndrome; DKA, diabetic ketoacidosis.

Conclusions: Linkage of vEHR-EHR and One Health demonstrator are feasible

- We identified a significant number of linked pairs, i.e., individuals who were both UCHealth patients and an owner of an animal patient of the CSU-VTH.
- MetS-like key indicators are frequently found in the vEHR, suggesting the co-occurence of MetS and MetS-like in coincident households is an achievable first use-case for the pet-patient registry.

Proof-of-concept supports future direction: Buildout the CSU-CU pet-patient registry

The CSU-CU pet-patient registry will provide a novel mechanism to study health and disease co-occurrence in coincident households in order to improve One Health outcomes across species living in shared environments.

CSU-CU pet-patient data registry will be housed within the HIPAA compliant Health Data Compass environment. vEHR data is transformed into the Observational Medical Outcomes Partnership (OMOP) common data model (CDM). Place-based indices (e.g., air quality indices) will be imported into the registry. Secure data analysis will be performed using Eureka Virtual Machines (VM).

Contact: katie@tislab.org. This research is supported by the NIH/NIAMS Award Number K12AR084226, the NIH/NCATS Colorado CTSA Award Number UM1 TR004399 and the UNC Department of Genetics.

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?

May 6: Evidence Synthesis

Martijn Schuemie

Research Fellow, Global Epidemiology Organization
Johnson & Johnson

Yong Chen

Professor of Biostatistics
University of Pennsylvania

Evidence Synthesis

An R package for combining causal effect estimates without sharing individual person data

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