



# New Adopters/Community Members

OHDSI Community Call  
July 12, 2022 • 11 am ET



# July OHDSI Community Calls

Date	Topic
July 12	New Adopter Introductions and Q&A
July 19	Workgroup Updates
July 26	CDM Update Process



# Three Stages of The Journey

**Where Have We Been?**

**Where Are We Now?**

**Where Are We Going?**





# OHDSI Shoutouts!



Congratulations to the team of **Erica Voss, Saberi Rana Ali, Arun Singh, Peter Rijnbeek, Martijn Schuemie, and Daniel Fife** on the publication of **Hip Fracture Risk After Treatment with Tramadol or Codeine: An Observational Study in Drug Safety.**

Drug Safety  
<https://doi.org/10.1007/s40264-022-01198-9>

## ORIGINAL RESEARCH ARTICLE



### Hip Fracture Risk After Treatment with Tramadol or Codeine: An Observational Study

Erica A. Voss<sup>1,2,3</sup> · Saberi Rana Ali<sup>1</sup> · Arun Singh<sup>1</sup> · Peter R. Rijnbeek<sup>2,3</sup> · Martijn J. Schuemie<sup>1,2</sup> · Daniel Fife<sup>1</sup>

Accepted: 29 May 2022  
© The Author(s) 2022

#### Abstract

**Introduction** Hip fractures among older people are a major public health issue, which can impact quality of life and increase mortality within the year after they occur. A recent observational study found an increased risk of hip fracture in subjects who were new users of tramadol compared with codeine. These drugs have somewhat different indications. Tramadol is indicated for moderate to severe pain and can be used for an extended period; codeine is indicated for mild to moderate pain and cough suppression.

**Objective** In this observational study, we compared the risk of hip fracture in new users of tramadol or codeine, using multiple databases and analytical methods.

**Methods** Using data from the Clinical Practice Research Datalink and three US claims databases, we compared the risk of hip fracture after exposure to tramadol or codeine in subjects aged 50–89 years. To ensure comparability, large-scale propensity scores were used to adjust for confounding.

**Results** We observed a calibrated hazard ratio of 1.10 (95% calibrated confidence interval 0.99–1.21) in the Clinical Practice Research Datalink database, and a pooled estimate across the US databases yielded a calibrated hazard ratio of 1.06 (95% calibrated confidence interval 0.97–1.16).

**Conclusions** Our results did not demonstrate a statistically significant difference between subjects treated for pain with tramadol compared with codeine for the outcome of hip fracture risk.

#### Key Points

When subjects treated with cough prior to the index date were excluded from the analyses of the US claims databases, there was a decrease in the number of codeine-treated versus tramadol-treated subjects, suggesting that codeine was often used for cough instead of pain.

This finding highlights the importance of accounting for differences in indications, when comparing data from subjects treated with tramadol versus codeine.

#### 1 Introduction

Hip fractures are a major public health issue, particularly for older persons [1]. These fractures of the upper portion of the femur, are classified per anatomical location: femoral-neck, intertrochanteric, or subtrochanteric [2]. Hip fractures are associated with a 25% reduction in life expectancy and approximately 17% of patients who experience fractures spend their remaining life in a nursing facility [3]. Globally, hip fractures affect 18% of women and 6% of men and rank among the top ten causes of disability [4, 5]. Measures that reduce the risk of hip fracture are therefore important to patient welfare.





# OHDSI Shoutouts!



Congratulations to the team of **Ines Reinecke, Mirko Gruhl, Martin Pinnau, Fatma Betül Altun, Michael Folz, Michéle Zoch, Franziska Bathelt, and Martin Sedlmayr** on the publication of **An OHDSI ATLAS Extension to Support Feasibility Requests in a Research Network** in Volume 295 of *Studies in Health Technology and Informatics*.

*Advances in Informatics, Management and Technology in Healthcare*  
J. Mantas et al. (Eds.)

© 2022 The authors and IOS Press.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

doi:10.3233/SHIT220778

515

## An OHDSI ATLAS Extension to Support Feasibility Requests in a Research Network

Ines REINECKE<sup>a,1</sup>, Mirko GRUHL<sup>a</sup>, Martin PINNAU<sup>b</sup>, Fatma Betül ALTUN<sup>b</sup>, Michael FOLZ<sup>b</sup>, Michéle ZOCH<sup>a</sup>, Franziska BATHELT<sup>a</sup> and Martin SEDLMAYR<sup>a</sup>

<sup>a</sup>Carl Gustav Carus Faculty of Medicine, Center for Medical Informatics, Institute for Medical Informatics and Biometry, Technische Universität Dresden, Dresden, Germany

<sup>b</sup>Institute of Medical Informatics, Goethe University Frankfurt, University Hospital Frankfurt, Frankfurt am Main, Germany

**Abstract.** Checking the feasibility of real-world data to answer a certain research question is crucial especially in a multi-site research network. In this work we present an extension of the ATLAS user interface for the OMOP common data model that integrates into an existing national feasibility network and thus foster capabilities for future participation in international research studies.

**Keywords.** OHDSI, OMOP, feasibility requests, interoperability

### 1. Introduction

Before running studies in a research network on real world data (RWD) spread across different sites, it is crucial to evaluate whether the number of available medical records that fits specific criteria is sufficient to answer a research question. Those collaborative efforts are currently driven in Germany by the Medical Informatics Initiative (MI-I) [1]. The Observational Health Data Science and Informatics (OHDSI) [2] software ATLAS is a user interface for research analytics that can be used by a single site and connects against data stored in the Observational Medical Outcomes Partnership (OMOP) common data model (CDM). The importance of the OMOP CDM in research is continuously increasing over the past years [3]. Thus, this paper presents an enriched ATLAS functionality to support cross-site feasibility requests in the MI-I consortium Medical Informatics in Research and Care in University Medicine (MIRACUM) [4].



# OHDSI Shoutouts!



Congratulations to the team of **Anthony Molinaro and Frank DeFalco** on the publication of **Empirical assessment of alternative methods for identifying seasonality in observational healthcare data** in **BMC Medical Research Methodology**.

Molinaro and DeFalco  
*BMC Medical Research Methodology* (2022) 22:182  
<https://doi.org/10.1186/s12874-022-01652-3>

BMC Medical Research  
Methodology

RESEARCH

Open Access

## Empirical assessment of alternative methods for identifying seasonality in observational healthcare data



Anthony Molinaro\* and Frank DeFalco

### Abstract

**Background:** Seasonality classification is a well-known and important part of time series analysis. Understanding the seasonality of a biological event can contribute to an improved understanding of its causes and help guide appropriate responses. Observational data, however, are not comprised of biological events, but timestamped diagnosis codes the combination of which (along with additional requirements) are used as proxies for biological events. As there exist different methods for determining the seasonality of a time series, it is necessary to know if these methods exhibit concordance. In this study we seek to determine the concordance of these methods by applying them to time series derived from diagnosis codes in observational data residing in databases that vary in size, type, and provenance.

**Methods:** We compared 8 methods for determining the seasonality of a time series at three levels of significance (0.01, 0.05, and 0.1), against 10 observational health databases. We evaluated 61,467 time series at each level of significance, totaling 184,401 evaluations.

**Results:** Across all databases and levels of significance, concordance ranged from 20.2 to 40.2%. Across all databases and levels of significance, the proportion of time series classified seasonal ranged from 4.9 to 88.3%. For each database and level of significance, we computed the difference between the maximum and minimum proportion of time series classified seasonal by all methods. The median within-database difference was 54.8, 34.7, and 39.8%, for  $p < 0.01$ , 0.05, and 0.1, respectively.

**Conclusion:** Methods of binary seasonality classification when applied to time series derived from diagnosis codes in observational health data produce inconsistent results. The methods exhibit considerable discord within all databases, implying that the discord is a result of the difference between the methods themselves and not due to the choice of database. The results indicate that researchers relying on automated methods to assess the seasonality of time series derived from diagnosis codes in observational data should be aware that the methods are not interchangeable and thus the choice of method can affect the generalizability of their work. Seasonality determination is highly dependent on the method chosen.

**Keywords:** ACHILLES, ARIMA, CASTOR, Classification, Common data model, Cyclical, Observational data, OHDSI, OMOP CDM, R, Seasonality, Time series



# OHDSI Shoutouts!



Congratulations to the team of **Gayathri Delanerolle, Robert Williams, Ana Stipancic, Rachel Byford, Anna Forbes, Sneha Anand, Declan Bradley, Ruby Tsang, Siobhán Murphy, Ashley Akbari, Stuart Bedston, Ronan Lyons, Rhiannon Owen, Jillian Beggs, Antony Chuter, Dominique Balharry, Mark Joy, Aziz Sheikh, F.D. Richard Hobbs, and Simon de Lusignan** on the publication of **Methodological issues for using a common data model (CDM) of COVID-19 vaccine uptake and important adverse events of interest (AEIs): the Data and Connectivity COVID-19 Vaccines Pharmacovigilance (DaC-VaP) United Kingdom feasibility study** in JMIR Formative Research.

Currently accepted at: [JMIR Formative Research](#)

Date Submitted: Mar 8, 2022

Date Accepted: May 17, 2022

Date Submitted to PubMed: Jul 5, 2022



This paper has been accepted and is currently in production.

It will appear shortly on [10.2196/37821](#)

The final accepted version (not copyedited yet) is in [this tab](#).

An "ahead-of-print" version has been submitted to Pubmed, see PMID: [35786634](#)

Preprint

Accepted  
Manuscript

Methodological issues for using a common data model (CDM) of COVID-19 vaccine uptake and important adverse events of interest (AEIs): the Data and Connectivity COVID-19 Vaccines Pharmacovigilance (DaC-VaP) United Kingdom feasibility study.

Gayathri Delanerolle; Robert Williams; Ana Stipancic; Rachel Byford; Anna Forbes; Sneha Anand; Declan Bradley;

Ruby Tsang; Siobhán Murphy; Ashley Akbari; Stuart Bedston; Ronan Lyons; Rhiannon Owen; Jillian Beggs;

Antony Chuter; Dominique Balharry; Mark Joy; Aziz Sheikh; F.D. Richard Hobbs; Simon de Lusignan

## ABSTRACT

### Background:

The Data and Connectivity COVID-19 Vaccines Pharmacovigilance (DaC-VaP) UK-wide collaboration was created to monitor vaccine uptake and effectiveness and provide pharmacovigilance using routine clinical and administrative data. To monitor these, pooled analyses may be needed. However, variation in terminologies present a barrier as, England uses the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), while the rest of the UK uses the Readv2 terminology in primary care. The availability of data sources is not uniform across the UK.



# OHDSI Shoutouts!



Congratulations to the team of **Yae Won Tak, Seng Chan You, Jeong Hyun Han, Soon-Seok Kim, Gi-Tae Kim and Yura Lee** on the publication of **Perceived Risk of Re-Identification in OMOP-CDM Database: A Cross-Sectional Survey** in the Journal of Korean Medical Science.

**JKMS**  
JOURNAL OF KOREAN MEDICAL SCIENCE

Open Access, Peer-reviewed, Weekly

About

View Full-text

For Contributors

Archive > v.37(26); Jul 2022 > 10.3346/jkms.2022.37.e205

Original Article

Open Access

ABSTRACT | **ARTICLE** | PDF | PUBREADER | ePUB | FIGURES+TABLES | REFERENCES | SUPPL. MATERIALS

Download Citation

J Korean Med Sci. 2022 Jul 04;37(26):e205. English.

Published online Jun 20, 2022. <https://doi.org/10.3346/jkms.2022.37.e205>

© 2022 The Korean Academy of Medical Sciences.

## Perceived Risk of Re-Identification in OMOP-CDM Database: A Cross-Sectional Survey

Yae Won Tak ,<sup>1,\*</sup> Seng Chan You ,<sup>2,\*</sup> Jeong Hyun Han ,<sup>1</sup> Soon-Seok Kim ,<sup>3</sup> Gi-Tae Kim ,<sup>4</sup> and Yura Lee

<sup>1</sup>Department of Information Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea.

<sup>2</sup>Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Korea.

<sup>3</sup>Department of Big Data Science, Halla University, Wonju, Korea.

<sup>4</sup>UPS Data Corporation, Seoul, Korea.

Address for Correspondence: Yura Lee, MD, PhD. Department of Information Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea. Email: [haepary@naver.com](mailto:haepary@naver.com); Email: [haepary@amc.seoul.kr](mailto:haepary@amc.seoul.kr)

\*Yae Won Tak and Seng Chan You contributed equally to this work.





# OHDSI Shoutouts!



Congratulations to the team of **Emily Pfaff, Andrew Girvin, Tellen Bennett, Abhishek Bhatia, Ian Brooks, Rachel Deer, Jonathan Dekermanjian, Sarah Elizabeth Jolley, Michael Kahn, Kristin Kostka, Julie McMurry, Richard Moffitt, Anita Walden, Christopher Chute, Melissa A Haendel, and the N3C Consortium** on the publication of **Identifying who has long COVID in the USA: a machine learning approach using N3C data** The Lancet Digital Health.

## Identifying who has long COVID in the USA: a machine learning approach using N3C data

Emily R Pfaff\*, Andrew T Girvin\*, Tellen D Bennett, Abhishek Bhatia, Ian M Brooks, Rachel R Deer, Jonathan P Dekermanjian, Sarah Elizabeth Jolley, Michael G Kahn, Kristin Kostka, Julie A McMurry, Richard Moffitt, Anita Walden, Christopher G Chute, Melissa A Haendel, The N3C Consortium†

### Summary

**Background** Post-acute sequelae of SARS-CoV-2 infection, known as long COVID, have severely affected recovery from the COVID-19 pandemic for patients and society alike. Long COVID is characterised by evolving, heterogeneous symptoms, making it challenging to derive an unambiguous definition. Studies of electronic health records are a crucial element of the US National Institutes of Health's RECOVER Initiative, which is addressing the urgent need to understand long COVID, identify treatments, and accurately identify who has it—the latter is the aim of this study.

**Methods** Using the National COVID Cohort Collaborative's (N3C) electronic health record repository, we developed XGBoost machine learning models to identify potential patients with long COVID. We defined our base population (n=1793 604) as any non-deceased adult patient (age ≥18 years) with either an International Classification of Diseases-10-Clinical Modification COVID-19 diagnosis code (U07.1) from an inpatient or emergency visit, or a positive SARS-CoV-2 PCR or antigen test, and for whom at least 90 days have passed since COVID-19 index date. We examined demographics, health-care utilisation, diagnoses, and medications for 97 995 adults with COVID-19. We used data on these features and 597 patients from a long COVID clinic to train three machine learning models to identify potential long COVID among all patients with COVID-19, patients hospitalised with COVID-19, and patients who had COVID-19 but were not hospitalised. Feature importance was determined via Shapley values. We further validated the models on data from a fourth site.

**Findings** Our models identified, with high accuracy, patients who potentially have long COVID, achieving areas under the receiver operator characteristic curve of 0.92 (all patients), 0.90 (hospitalised), and 0.85 (non-hospitalised). Important features, as defined by Shapley values, include rate of health-care utilisation, patient age, dyspnoea, and other diagnosis and medication information available within the electronic health record.

**Interpretation** Patients identified by our models as potentially having long COVID can be interpreted as patients warranting care at a specialty clinic for long COVID, which is an essential proxy for long COVID diagnosis as its definition continues to evolve. We also achieve the urgent goal of identifying potential long COVID in patients for clinical trials. As more data sources are identified, our models can be retrained and tuned based on the needs of individual studies.



Lancet Digit Health 2022;  
4: e532-41

Published Online  
May 16, 2022  
[https://doi.org/10.1016/S2589-7500\(22\)00048-6](https://doi.org/10.1016/S2589-7500(22)00048-6)

\*Co-first authors

†Members are listed at the end of the Article

Department of Medicine, UNC Chapel Hill School of Medicine, Chapel Hill, NC, USA (E R Pfaff PhD); Palantir Technologies, Denver, CO, USA (A T Girvin PhD); Section of Informatics and Data Science, Department of Pediatrics (T D Bennett MD, M G Khan MD) and Section of Critical Care Medicine, Department of Pediatrics (T D Bennett), Colorado Center for Personalised Medicine, Division of Biomedical Informatics & Personalized Medicine, Department of Medicine (I M Brooks PhD), Department of Biostatistics and Informatics, Colorado School of Public Health (J P Dekermanjian MS), Division of Pulmonary and Critical Care Medicine, Department of



@OHDSI

[www.ohdsi.org](http://www.ohdsi.org)

#JoinTheJourney



ohdsi



# 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](mailto: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

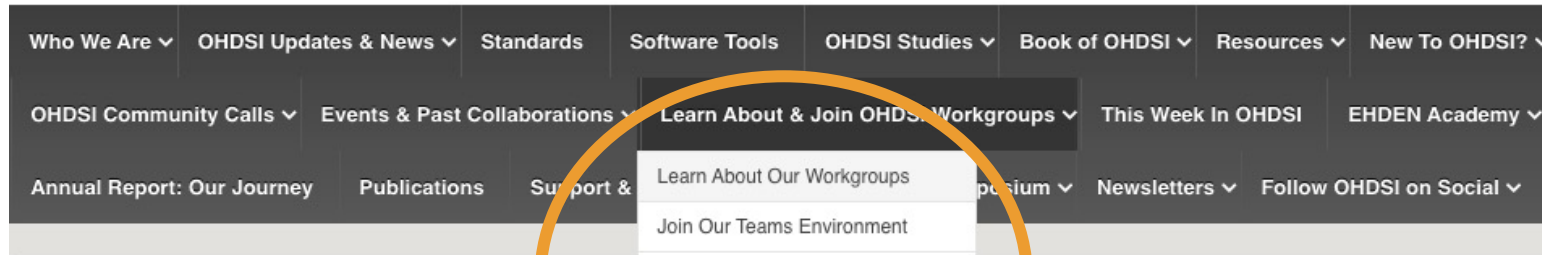


Date	Time (ET)	Meeting
Tuesday	12 pm	Common Data Model Vocabulary Subgroup
Tuesday	3 pm	OMOP CDM Oncology Outreach/Research Subgroup
Wednesday	7 am	Medical Imaging
Wednesday	11 am	Open-Source Community
Wednesday	12 pm	FHIR and OMOP Terminologies Subgroup (ZOOM)
Wednesday	2 pm	Natural Language Processing
Thursday	10 am	Data Quality Dashboard
Thursday	12 pm	FHIR and OMOP Oncology Subgroup
Thursday	1 pm	OMOP CDM Oncology Vocabulary/Development Subgroup
Friday	9 am	GIS – Geographic Information Systems
Friday	9 am	Education
Monday	10 am	Healthcare Special Interest Group

[www.ohdsi.org/upcoming-working-group-calls](http://www.ohdsi.org/upcoming-working-group-calls)



# Join OHDSI Workgroups



## OHDSI Workgroups

OHDSI's central mission is to improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care. We work towards that goal in the areas of data standards, methodological research, open-source analytics development, and clinical applications.

Our workgroups present opportunities for all community members to find a home for their talents and passions, and make meaningful contributions. We are always looking for new collaborators. Learn more about these workgroups by checking out this page. Any workgroup that provided a community call update is highlighted in the top section.

**See an area where you want to contribute? Please Join The Journey!**

### Join Our Workgroup Efforts!

[Form To Join Workgroups In MSTeams](#)

[Weekly Workgroup Meeting Schedule](#)

[www.ohdsi.org/ohdsi-workgroups/](http://www.ohdsi.org/ohdsi-workgroups/)



# OHDSI European Symposium Videos

## The Main Conference

### Session 1

#### INTRODUCTION

3:08 – Welcome to the European OHDSI Journey (Peter Rijnbeek, Chair, Department of Medical Informatics Erasmus MC)  
13:00 – Journey of OHDSI: Where Have We Been? (George Hripcsak, Vivian Beaumont Allen Professor and Chair, Biomedical Informatics, Columbia University Medical Center)

#### 34:45 – A CRUISE AROUND THE OHDSI EUROPE COMMUNITY (moderated by Nigel Hughes, Janssen Research and Development)

37:00 – Estonia: Conversion of Estonian health data into the OMOP CDM (Marek Oja, Institute of Computer Science, University of Tartu)  
42:59 – Finland: The Finnish OMOP data network (FinOMOP) (Javier Gracia-Tabuenca, Tampere University of Technology)  
49:33 – Denmark: Transforming Danish Registries to the OMOP Common Data Model: use case on the Danish Colorectal Cancer Group (DCCG) Database (Adamantia Tsochnika, Center for Surgical Science, Zealand University Hospital)  
57:04 – Norway: Norwegian registries onto OMOP Common Data Model: mapping challenges and opportunities for pregnancy studies (Eimir Hurley, University of Oslo)  
1:04:25 – Germany: OHDSI Germany: A recap after one year (Michele Zoch, Technische Universität Dresden)  
1:12:43 – Italy: The Italian national node of OHDSI Europe (Lucia Sacchi, University of Pavia)  
1:17:45 – Greece: An update from the Greek National Node (Pantelis Natsiavas, Centre for Research & Technology Hellas)  
1:23:07 – Ukraine: Integration prospects of the Ukrainian healthcare system with OMOP CDM (Mariia Kolesnyk, SciForce)  
1:29:40 – Israel: The journey from isolated EHR's to unified CDM network (Guy Livne, Israel Ministry of Health)  
1:34:30 – France: Semantic harmonization of the French National healthcare database (SNDS) (Lorien Benda, Health Data Hub)  
1:40:40 – Panel discussion including all European collaborators listed above.



### Session 2

#### COLLABORATOR SHOWCASE

1:33 – Collaborator Showcase Intro (Katia Verhamme, MD, Associate Prof Observational Data Analysis, Department of Medical Informatics, Erasmus MC, Rotterdam)  
2:48 – FeederNet (Federated E-Health Big Data for Evidence Renovation Network) platform in Korea (Chungsoo Kim, Ajou University)  
8:04 – OMOP Genomic mapping capacities in conversion of comprehensive genomic profiling results (Maria Rogozhina, Odysseus)  
12:59 – OMOP Mapping of Real-World Data From Brazil & Pakistan Towards Management of COVID-19 in the Global South (Sara Khalid, University of Oxford)  
19:23 – Impact of random oversampling and random undersampling on the development and validation of prediction models using observational health data (Cynthia Yang, Erasmus MC)  
24:23 – Real-world evidence is in demand: a summary of 'live' requests for RWE studies published by a European health technology assessment (HTA) agency (Jamie Elvidge, National Institute for Health and Care Excellence (NICE))  
31:48 – Why predicting risk can't identify 'risk factors': empirical assessment of model stability in machine learning across observational health databases (Aniek Markus, Erasmus MC)  
38:15 – TrajectoryViz: Interactive visualization of treatment trajectories (Maarja Pajusalu, Institute of Computer Science, University of Tartu)  
44:47 – Assessing treatment effect heterogeneity using the RiskStratifiedEstimation R-package (Alexandros Rekkas, Erasmus MC)  
49:45 – Defining the valid analytic space for quantitative bias analysis in pharmacoepidemiology (James Weaver, Janssen R&D)  
58:03 – A pilot study to evaluate the feasibility of using Observational Health Data Sciences and Informatics analytics tools for supporting the validation of safety signals (Ceyda Pekmez Kristiansen, Novo Nordisk)  
1:03:32 – Findable, standardized data at scale through the EHDEN Database Catalogue (Julia Kurps, The Hyve)



### Session 3

0:52 – Characterizing Adverse Events in COVID-19 infected patients across the OHDSI network (Erica A. Voss, MPH, Janssen Research and Development, Erasmus MC)  
28:10 – Data Analysis and Real World Interrogation Network (DARWIN EU®) (Peter R. Rijnbeek, PhD, Chair, Department of Medical Informatics, Erasmus MC)  
42:45 – Reaction panel with key stakeholders  
**Moderator**  
Dani Prieto-Alhambra, MD, PhD Professor of Pharmacology and Device Epidemiology University of Oxford, Professor of Real World Evidence and Methods Research, Erasmus MC  
**Panelists**  
Catherine Cohet, European Medicines Agency  
Filip Maljković, Heliant, Serbia  
Daniel Morales, Dundee University, UK  
Dalia Dawoud, NICE, UK  
Patrick Ryan, Janssen Research and Development, USA  
1:29:45 – Closure: Peter Rijnbeek



[ohdsi.org/2022-european-symposium/](https://ohdsi.org/2022-european-symposium/)





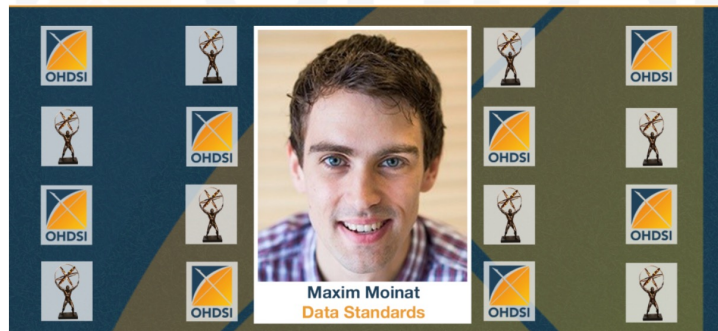
# Titan Awards Nominations Are Open

**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.

## 2021 OHDSI Titan Awards



**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 Ostropelets](#), [Columbia Univ.](#); [Dmitry Dymshyts](#), [Odysseus Data Services](#))

## 2021 OHDSI Titan Awards



## 2021 OHDSI Titan Awards



## 2021 OHDSI Titan Awards



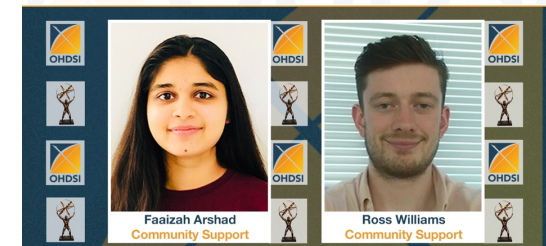
## 2021 OHDSI Titan Awards



## 2021 OHDSI Titan Awards



## 2021 OHDSI Titan Awards



[ohdsi.org/titan-awards](https://ohdsi.org/titan-awards)

# Latest OHDSI Newsletter Is Available

## Community Updates

### Where Have We Been?

- The 2022 OHDSI European Symposium brought together more than 350 collaborators on the Steam Ship Rotterdam for our first in-person event since the start of the COVID pandemic. Learn more about the symposium and some of its outputs later in this newsletter.
- The OHDSI community and SNOMED International formalized their long-time relationship with a five-year collaborative agreement that will benefit both of their user communities. The collaboration provides OHDSI and its user community with comprehensive ontologies on specific healthcare domains and content such as devices, social determinants of health, disease severity scores and modifiers of cancers, as well as better concept definitions and resolutions of composite concepts in large-scale observational research.

### Where Are We Now?

- A new tool to track OHDSI publications, citations, new authors and more has been developed by Paul Nagy and his team. [This tool is available](#) on the front page of the OHDSI web site.
- OHDSI had a record total of 139 submissions for the upcoming OHDSI 2022 Collaborator Showcase. The scientific review committee will go over each submission in July and notify accepted authors by August 3. Submissions came in the form of posters, software demos, and oral presentations. Thank you to everybody who submitted brief reports for our October global symposium.
- The #OHDSISocialShowcase has returned to highlight the Collaborator Showcase research presented at the European Symposium. Please follow our [Twitter](#) and [LinkedIn](#) feeds to learn more about the exciting work happening within our community.

## June Publications

Shoabi, A., Rao, G.A., Voss, E.A. *et al.* [Phenotype Algorithms for the Identification and Characterization of Vaccine-Induced Thrombotic Thrombocytopenia in Real World Data: A Multinational Network Cohort Study](#). *Drug Saf* 45, 685–698 (2022). doi: 10.1007/s40264-022-01187-y

Khera R, Schuemie MJ, Lu Y, et al. [Large-scale evidence generation and evaluation across a network of databases for type 2 diabetes mellitus \(LEGEND-T2DM\): a protocol for a series of multinational, real-world comparative cardiovascular effectiveness and safety studies](#). *BMJ Open* 2022;12:e057977. doi: 10.1136/bmjopen-2021-057977



## The Journey Newsletter (July 2022)

Our community gathered together for the first time since the COVID pandemic for the 2022 European Symposium, while leaders in our open-source community provided tutorials on four tools that can aid global research. OHDSI and SNOMED formalized an important agreement that will aid collaboration opportunities around the world, and our community publications and presentations from June are linked below. All that, as well as community updates and plenty more, are available in our latest newsletter.

[#JoinTheJourney](#)

## European Symposium Recap



The 2022 OHDSI European Symposium was held June 24-26 on the SS Rotterdam in the Netherlands. More than 350 collaborators gathered together for the community's first in-person symposium since the COVID pandemic to connect, share research, and learn from each other.

Among the topics during the symposium was the use of the OMOP-CDM, tool development, and future research. The first day included a collaborator showcase which featured both posters and podium presentations to highlight OHDSI's research achievements, and interactive demonstrations of OHDSI's open-source software tools.

## OHDSI, SNOMED International Formalize 5-Year Agreement To Open New Research Opportunities For Research Communities

**SNOMED, OHDSI Finalize Five-Year Collaboration Agreement To Open New Opportunities For Research Communities**



The OHDSI community and SNOMED International have formalized their long-time relationship with a five-year collaborative agreement that will benefit both of their user communities.

## Collaborator Spotlight: Nicole Pratt

### Spotlight: Nicole Pratt



The work that has been generated in **LEGEND** and **EUMAEUS** is important clinically. It can help to update clinical guidelines and provides robust evidence for medicine regulators — but for me these landmark studies have also provided critical insights into which methodologies are appropriate under which conditions — especially the value of empirical calibration!



Nicole Pratt, a longtime collaborator with the OHDSI community who was recently named [one of eight new ISPE Fellows for 2022](#), is the Deputy Director of the Quality Use of Medicines and Pharmacy Research Centre at the University of South Australia. She is a member of the Drug Utilisation Subcommittee (DUSC) of the Australian Department of Health Pharmaceutical Benefits Advisory Committee (PBAC).



@OHDSI

[www.ohdsi.org](http://www.ohdsi.org)

[#JoinTheJourney](#)

[in](#) [ohdsi](#)





# Latest OHDSI Newsletter Is Available

The screenshot displays the OHDSI website header and navigation menu. The OHDSI logo, consisting of a stylized orange and blue square with a white 'X' and the text 'OHDSI OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS', is at the top. Below it is a dark navigation bar with various links. The 'Newsletters' link is highlighted with an orange circle, and its dropdown menu is open, showing options from 'Subscribe' to 'Full Archive'. The 'July 2022' option is the most prominent in the list.

**Who We Are** ▾ **OHDSI Updates & News** ▾ **Standards** **Software Tools** **OHDSI Studies** ▾ **Book of OHDSI** ▾ **Resources** ▾ **New To OHDSI?** ▾

**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** ▾

**Welcome to OHDSI!**

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions

**2022 OHDSI Symposium**

The 2022 OHDSI Symposium will be held from October 14-16 at the Bethesda Conference Center. We will hold the symposium in person, Oct. 14-16, and a virtual version will be held Oct. 17-18.

**Newsletters**

- Subscribe
- July 2022
- June 2022
- May 2022
- April 2022
- March 2022
- February 2022
- 2021 In Review
- Full Archive



# Job Openings

Professor **Dani Prieto-Alhambra** and his team at the University of Oxford will be hiring two Research Assistants in Health Data Sciences.

The application deadline is August 8, 2022.

The link and more information will be available on the community calls page.

**UNIVERSITY OF OXFORD**

UK date and time: 11-July-2022 16:46

### Applicant Options

- New Search
- Login
- Job Details
- Help
- Terms of Use & Privacy Policy



### Job Details

#### Research Assistant in Health Data Sciences (2 posts)

**Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, Windmill Road, Oxford**

We have an exciting opportunity for an enthusiastic and dedicated Research Assistants in Health Data Sciences to join the Pharmaco- and Device epidemiology research group lead by Professor Daniel Prieto-Alhambra at the Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford.

As a Research Assistant in Health Data Sciences you will support the programming of analytical pipelines for the analysis of routinely collected data mapped to the OMOP Common Data Model. You will prepare analytical packages to run a number of pre-specified analyses, contribute to wider project planning, including ideas for new research projects and manage own research and administrative activities, within guidelines provided by senior colleagues.

You will hold a relevant post-graduate degree in Mathematics, Engineering, Health Data Sciences or Biostatistics. You will have an experience in biostatistics as well as experience in analysis of OMOP-mapped data. Knowledge of medical statistics and expertise in handling large patient level datasets, good knowledge of programming in R packages for statistical analyses and ability to communicate results effectively with colleagues in any discipline are essential. Expertise in pharmaco and/or vaccine epidemiology, experience working with electronic medical records/routinely collected data and experience of working within an academic environment are desirable.

This is a full-time fixed-term appointment for 2 years.

The closing date for this position is 12 noon on Monday 08/08/2022. You will be required to upload a CV and supporting statement as part of your online application.

Contact Person :	HR Team, NDORMS	Vacancy ID :	159236
Contact Phone :		Closing Date & Time :	08-Aug-2022 12:00
Pay Scale :	STANDARD GRADE 6	Contact Email :	<a href="mailto:hr@ndorms.ox.ac.uk">hr@ndorms.ox.ac.uk</a>
Salary (£) :	£29,614 to £36,326 p. a.		

Click on the link(s) below to view documents	Filesize
<a href="#">159236_JD</a>	472

[Return to Search Results](#)[Apply Now](#)





# Job Openings

Assistant professor **Brianne Oliveri-Mui** announced an opening for an Postdoctoral Fellow to work at the Roux Institute at Northeastern University.

If you are interested, please reach out to Dr. Mui at [b.mui@northeastern.edu](mailto:b.mui@northeastern.edu).

The link and more information will be available on the community calls page.

## Observational Health Data Sciences and Informatics Postdoctoral Fellow

Apply

📍 Portland, ME

🕒 Full time

🕒 Posted 30+ Days Ago

📄 R105484

### About the Opportunity

The Roux Institute at Northeastern University has one opening for a Postdoctoral Research Fellow beginning on or about September 1, 2022. The fellow will have an opportunity to conduct observational and administrative database research (e.g., analysis of existing datasets) on health outcomes for older adults with HIV or LGBT older adults, under the supervision of the PI. The fellow will devote most of their time to independent research aligned with the PI's interests and across federated and local research models.

Position offers exceptional opportunity for collaboration at the OHDSI center on major projects in the U.S. and overseas. This research will directly improve our ability to use real world data to characterize under-represented and marginalized patient populations, construct population level estimates relating exposures to health outcomes, and to enhance clinical decision making through improved patient-level predictions. The term of fellowship appointment will be for two years, contingent on continued funding. Stipend will be commensurate with experience, based on levels mandated by NIH.

The main research areas specific to older people with HIV or in the LGBTQ+ communities are as follows:

- Measurement of comorbidities, care quality, health outcomes and healthcare utilization patterns
- Risk assessment of multimorbidity, healthcare and prescription access

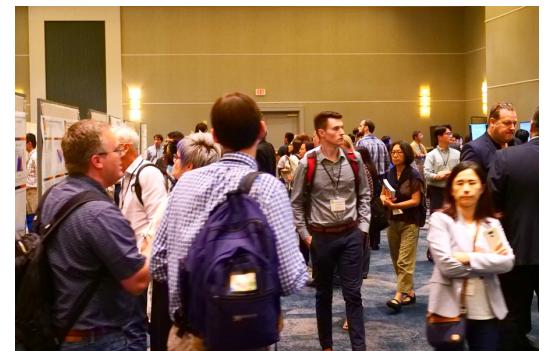


# 2022 OHDSI Symposium

Registration is OPEN for  
**#OHDSI2022!**

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

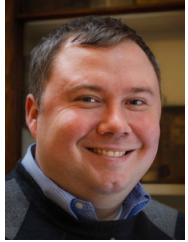
[www.ohdsi.org/ohdsi2022symposium](http://www.ohdsi.org/ohdsi2022symposium)





# An Introductory Journey From Data To Evidence

OHDSI2022 Tutorial • Saturday, Oct. 15 • Bethesda, Md.



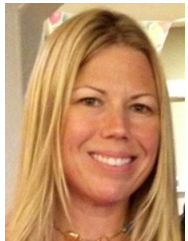
**The OHDSI Journey:  
Where Are We Going?**

**Patrick Ryan**



**OMOP Common Data  
Model and Vocabulary**

**Clair Blacketer**



**ETL – A Source Database  
Into OMOP CDM**

**Melanie Philofsky**



**Creating Cohort  
Definitions**

**Asieh Golozar**



**Phenotype Evaluations**

**Gowtham Rao**



**Characterization**

**Kristin Kostka**



**Estimation**

**Martijn Schuemie**



**Prediction**

**Jenna Reps**



**The OHDSI Journey: Where  
Do We Go From Here?**

**George Hripcsak**





# Workgroup Activities

Saturday, Oct. 15, and Sunday, Oct. 16

Saturday, Oct 15					
Start Time (ET)	End Time (ET)				
800	900	Tutorial	HADES Hack-a-thon: Part 1	Oncology WG	FHIR-OMOP: Terminologies Subgroup, Part 1
900	1000				FHIR-OMOP: Increasing the Value of Data Through a Rich Set of Attributes
1000	1100		Lunch	Lunch	Lunch
1100	1200				
1200	1300				
1300	1400				
1400	1500				
1500	1600		Methods Research (PLE/PLP)	Oncology WG (continued)	FHIR-OMOP: Data Model Harmonization Subgroup
1600	1700			Natural Language Processing	FHIR-OMOP: Oncology Subgroup
1700	1800				FHIR-OMOP: Terminologies Subgroup, Part 2
1800	1900				
Sunday, Oct 16					
800	900	All-Hands Workgroup Meeting			
900	1000				
1000	1100				
1100	1200	Lunch		Lunch	Lunch
1200	1300				
1300	1400	Phenotype Evaluation	HADES Hack-a-thon: Part 2	Education	CDM and Data Quality
1400	1500			Health Equity	
1500	1600				
1600	1700				



# #OHDSISocialShowcase This Week

Mapping concepts from the Netherlands Cancer Registry to the OMOP-CDM - experiences and challenges

PRESENTER: Chiara Attanasio

## INTRODUCTION

The Netherlands Cancer Registry (NCR) is a population-based structured cancer registry with nationwide coverage since 1989 and 3 million patients total.

We commenced conversion of the NCR data to the OMOP-CDM in 2020. Here we describe our experiences and challenges with the mapping work still ongoing as part of an EHDEN type two grant.

## METHODS

We focus on two tables from the NCR.

- The "Event" table, which contains clinical events such as diagnostic tests and primary treatment.
- The "Event Detail" table, which contains details of entries in the "Event" table and associated values.

## Mapping workflow:

1. Selection of events and details to map,
2. Pre-processing of the source concepts,
3. Multiple mapping rounds and reviews,
4. Post-processing steps,
5. Final review by domain experts,
6. Implementation.

## RESULTS

The first batch of source concepts from the NCR that we processed within the EHDEN grant were related to the most frequently occurring diagnostic events. After the pre-processing step around 350 pre-coordinated source concepts needed to be mapped, accounting for around 10% of all NCR diagnostic concepts.

Events in the NCR rely heavily on post-coordination. This is not supported in the OMOP-CDM.

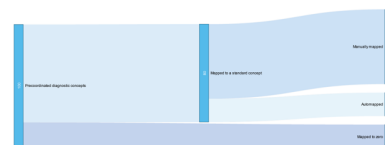


Figure 1: Mapping results for the diagnostic concepts. Auto mapped means mapped according to the automatic suggestion from the edenceReviewer of edenceHealth.

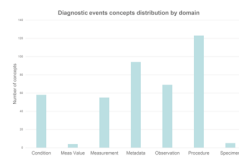


Figure 2: Distribution of the NCR diagnostic events over the OMOP domains.

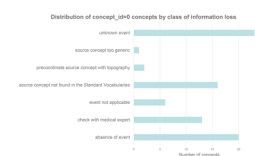


Figure 3: Overview of the reasons why diagnostic events are mapped to zero.

Apart from a few small mapping challenges, the main roadblock we faced was linked to the structure of the "Event" and "Event Detail" information in the NCR, which did not match the OMOP-CDM. An "Event", "Event Detail", "Value" combination can be mapped in many different ways, and an event may contain more than one detail attached to it.

A big effort went into identifying all occurring 'mapping' situations firstly to standardize our approach during the mapping, and secondly to feed this knowledge into the ETL design and implementation step.

## CONCLUSIONS

So far, prior the medical review step, we have mapped around 10% of the diagnostic events from the NCR. If we consider only source concepts that we wanted to add to the CDM, but for which we could not find a suitable standard concept, then only 5% of these were mapped to 0.

Within the EHDEN grant effort, we will not be able to map all variables in the NCR, but we aim to have an interesting data set to participate in international studies, starting by adding primary treatment events.

We have already done so for a PIONEER study-a-thon in 2021 and we are currently participating in the HANA project on colon cancer treatment effects with South Korea.

Make sure to check the other two posters from IKNL!

Chiara Attanasio ([c.attanasio@iknl.nl](mailto:c.attanasio@iknl.nl)),  
Floor Klijn ([f.klijn@iknl.nl](mailto:f.klijn@iknl.nl)),  
Jennifer Caffarel ([j.caffarel@iknl.nl](mailto:j.caffarel@iknl.nl)),  
Peter Prinsen ([p.prinsen@iknl.nl](mailto:p.prinsen@iknl.nl))  
IKNL

edence Health

EHDEN

IKNL  
Integraal  
Kankercentrum  
Nederland

OHDSI

## MONDAY Mapping concepts from the Netherlands Cancer Registry to the OMOP-CDM - experiences and challenges

Lead: Chiara Attanasio



# #OHDSISocialShowcase This Week

The journey from  
central operational data-lake  
to Medical Centers CDM network

by Guy Livne

#### INTRO:

- Directorate of government medical centers in Israel established a CDM network combine EHR data from 11 medical centers, dealing with regulation, privacy and technology issues

#### Main Challenges:

1. Translate local nonstandard Israeli terminologies to OMOP standard.
2. Define cloud security regulations as the first cloud-based solution for medical records in the Israel.
3. Combine data from different sites to one data-lake, preserving patient privacy with consistency across sites.

#### Achieves:

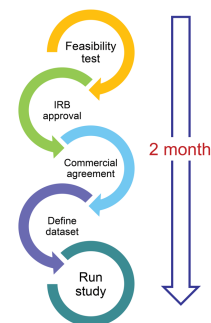
- ✓ CDM network of 3 medical centers and plan for additional 3 by the end of year 2022.
- ✓ Unify person\_id solution.
- ✓ Encryption & anonymization implemented in the ETL process.
- ✓ Use of ML models & NLP for terminology translation.
- ✓ 2-month flow from feasibility check to study.



Anonymize EHR data from  
country-wide Hospitals network  
combined to one CDM  
15 years of history records



#### Fast track for study



by Guy Livne, Nadav Rappoport,  
Nir Makover, Hadas Eshel-Geva,  
Hadar Kapach, Tomer Hadad, Yarin  
Alon, Naama Perry-Cohen



TUESDAY

The journey from central operational data-lake to Medical Centers CDM network

Lead: Guy Livne



# #OHDSISocialShowcase This Week

*Informativeness of clinical lymph node metastasis staging for patients undergoing curative intended surgery for colorectal cancer:*  
A national multi-register study  
PRESENTER: **Andreas Weinberger Rosen**

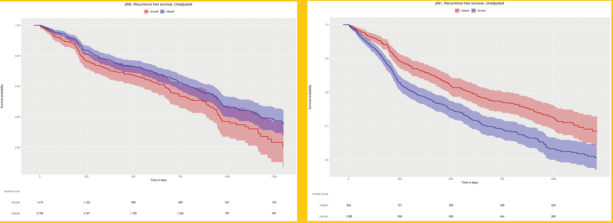
**INTRO:**  
Lymph node involvement is a driving factor for long-term oncological outcomes for patients undergoing surgery for colorectal cancer. Multidetector computed tomography is used for clinical staging of size of lymph nodes is used as a surrogate of lymph node involvement in clinical staging. Increased lymph nodes sized is also associated with better survival, possible being a proxy for a strong antitumoral immune response.

**METHODS**  
Four national register were used to identify all patients undergoing curative intended, surgery for colorectal cancer in an elective setting, with available clinical and pathological lymph node staging and assessment of mismatch repair proteins.  
Patients were divided by clinical N category into N0 or N1+ and compared to each other by the pathological N category. Numbers of covariates with a standardized difference of mean  $\geq 0.1$  were recorded. Recurrence, survival and recurrence-free-survival were investigated with incidence rates at 3 years and with Cox proportional Hazards and Kaplan Meier for 5 years.

Outcome	HR (95%CI)	p-value
cN1pN0 vs cN0pN0		
Overall survival	1.07 (0.794-1.43)	0.05
Recurrence free survival	1.25 (1.027-1.516)	0.025
Recurrence	1.37 (1.077-1.735)	0.01
cN0pN1 vs cN1pN1		
Overall survival	0.688 (0.522-0.948)	0.0011
Recurrence free survival	0.69 (0.581-0.816)	0.00002
Recurrence	0.676 (0.505-0.922)	0.00008
Hazard ratios of the various outcomes		
	cN1pN0 vs cN0pN0	cN0pN1 vs cN1pN1
Demographics	0/2	0/2
Condition occurrence	26/513	20/513
Procedure occurrence	21/576	21/576
Measurement occurrence	11/276	10/276
Observation occurrence	7/222	1/222
Drug exposure	0/667	0/667

Number of covariates with a standardized difference  $\geq 0.1$

Divergence between clinical and pathological N category is associated with long-term oncological outcomes – but not in the way we would suspect



	Persons	Deaths	Proportion per 1k persons	Time at risk (years)	Rate per 1k years
Clinically correct staged N0 disease	2596	103	39.57	6294	16.30
Observed	414	21	50.72	839	25.03
Expected	2394	77	32.16	5242	14.69
Clinically misstaged N0 disease	1516	50	32.98	3215	15.95
Observed	338	17	50.30	680	25.00
Expected	1179	32	28.34	2438	13.13
Clinically correct staged N1+ disease	1440	178	123.61	2972	59.89
Observed	200	21	105.00	394	53.30
Expected	1196	153	127.63	2499	61.22
Clinically misstaged N1+ disease	1051	83	78.97	2140	36.79
Observed	101	12	118.81	203	58.11
Expected	927	70	75.51	1993	36.98

3 year incidence rates of the various outcomes for patients with pN0 disease, stratified by MM status

	Persons	Deaths	Proportion per 1k persons	Time at risk (years)	Rate per 1k years
Clinically correct staged N0 disease	2596	103	39.57	6294	16.30
Observed	414	21	50.72	839	25.03
Expected	2394	77	32.16	5242	14.69
Clinically misstaged N0 disease	1516	50	32.98	3215	15.95
Observed	338	17	50.30	680	25.00
Expected	1179	32	28.34	2438	13.13
Clinically correct staged N1+ disease	1440	178	123.61	2972	59.89
Observed	200	21	105.00	394	53.30
Expected	1196	153	127.63	2499	61.22
Clinically misstaged N1+ disease	1051	83	78.97	2140	36.79
Observed	101	12	118.81	203	58.11
Expected	927	70	75.51	1993	36.98

3 year incidence rates of the various outcomes for patients with pN1 disease, stratified by MM status

Andreas Weinberger Rosen, Ilze Ose, Andi Tsuchnik, Ismail Gögenur



WEDNESDAY

Informativeness of clinical lymph node metastasis staging for patients undergoing curative intended surgery for colorectal cancer: A national multi-register study

Lead: **Andreas Weinberger Rosen**



# #OHDSISocialShowcase This Week

## Concept extraction from Dutch clinical text

PRESENTER: **Tom Seinen**

### INTRO:

- EHR databases contain vast amounts of unstructured text data.
- Free-text cannot be directly used for analysis.
- Named-Entity-Recognition (NER) is the task of extracting clinical concept from the free-text.
- No open-source NER tools exist for concept extraction from Dutch clinical text.
- We created and evaluated an open-source extraction tool for the extraction of concepts from Dutch clinical text by converting an existing framework, MedSpaCy.

### METHODS

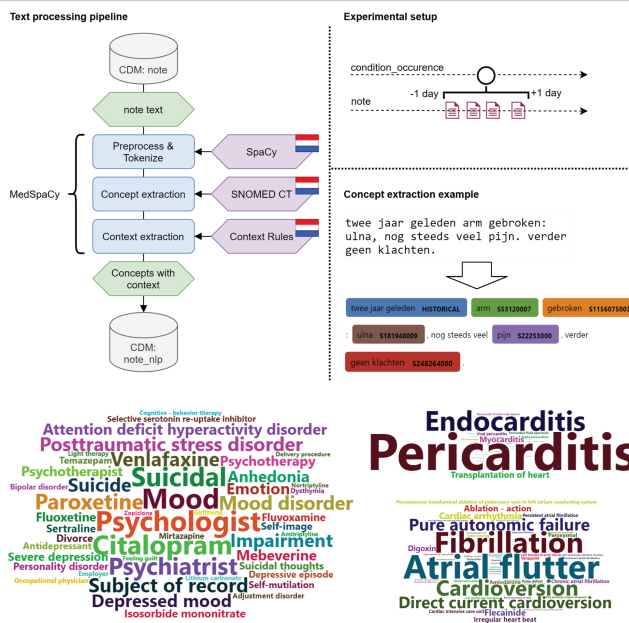
**Dataset** - Dutch GP database with 2.8 million patients (IPI) from 1992 to 2022, converted to the OMOP CDM.  
**Text preprocessing** - Only keep alphanumeric characters, tokenize with Dutch SpaCy model.  
**Concept extraction** - MedSpaCy's quickUMLS using the Dutch SNOMED CT ontology.  
**Context extraction** - MedSpaCy's context extraction using Dutch target rules.  
**Exploratory setup** - Framework was applied to notes surrounding the occurrences of 6 specific coded conditions. A window was of 1 day before and 1 day after the code occurrence. The most important concepts were identified for each code using the TFIDF value.

### RESULTS

	Alzheimer's disease	Depressive disorder	Infection disease of cardiovascular system
# code occurrences	1,228,047	5,581,178	122,482
# codes per patient	45.4	31.2	15.5
Mean # notes per code occurrence	4.7	4.0	4.8
Median; mean # words per note	7; 19.0	6; 15.0	4; 15.5
# concepts	4,216,892	14,702,150	452,448
# unique concepts	10,560	15,713	3,991
% negated concepts	16.1%	13.1%	14.8%
% historical concepts	3.5%	5.0%	5.5%
Mean # extracted concepts	26.2	13.6	27.5
per code occurrence			
Mean # extracted concepts per note	7.5	5.6	7.9
Ratio extracted concepts / note size	0.38	0.37	0.51

Summarizing statistics over 3 condition codes

## Extracting concepts from Dutch clinical text



### Top 5 most important concepts (by TFIDF) per condition

Alzheimer's disease	Depressive disorder	Infection disease of cardiovascular system	Fracture ulna/radius
1. Alzheimer's disease	1. Mental health care	1. Pericarditis	1. Bone structure of distal radius
2. Dementia	2. Psychologist	2. Endocarditis	2. Entire radius
3. Dementia	3. Mood	3. Transplantation of heart	3. Entire ulna
4. Alzheimer's disease	4. Case manager	4. Myocarditis	4. Fracture
5. Alzheimer's disease	5. Type 2 diabetes mellitus	5. Viral pericarditis	5. Reduction - action

### CONCLUSION

- We analyzed concepts around 6 coded condition occurrences in a Dutch OMOP CDM.
- The found concepts are descriptive and informative of the coded conditions.
- The extracted concepts show the ambiguity of several ICDPC codes.
- The detailed information extracted from the free-text can be used in further research or to improve the ETL to OMOP.

### FUTURE STEPS

- Quantitative evaluation and validation of the concept extraction framework
- Use of data in:
  - Patient level prediction
  - Diagnostic classification
  - ETL to OMOP CDM
- Effects of spelling correction on concept extraction
- Compare extracted concepts with the structured data in the OMOP CDM.

Tom M. Seinen<sup>1</sup>, Jan A. Kors<sup>1</sup>, Erik M. van Mulligen<sup>1</sup>, Peter R. Rijnbeek<sup>1</sup>, Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands



## Concept extraction from Dutch clinical text

THURSDAY

Lead: Tom Seinen



# #OHDSISocialShowcase This Week

**OHDSI Italia:**  
the Italian national node  
of OHDSI Europe

👤 **Lucia Sacchi**

**INTRO:**

- In Italy, as in Europe, the interest revolving around the OHDSI program and the OMOP Common Data Model has been growing in the last years. Academic research groups, data partners (e.g. hospitals and registries) and SMEs, thanks also to the drive provided by the EHDEN project, are, each from their own point of view, paying increasing attention to the different aspects that characterize the heterogeneous community of OHDSI.
- The "OHDSI Italia" node aims at becoming a point of reference and meeting place for all these Italian realities, to exploit each other's experiences in approaching the international community as well as to address typically national issues that would find little space / interest in other OHDSI working groups.

**OBJECTIVES:**

- Promote OMOP/OHDSI
  - through dissemination events
  - by adding new members and data partners to the national node
- Promote national projects
  - ICT
  - Observational studies
- Coordinate dialogue with
  - Local Regions
  - Ministry of research and Ministry of health
  - Existing projects
- Contribute to the OHDSI community
  - Mapping Italian terminologies and codes on OMOP
  - National codes: e.g. AIC codes for drugs (Federfarma)
  - Regional codes
- Define common administrative procedures
  - DPO approval
  - EC approval
  - AGID guidelines (for public entities)
  - Internal SOP / IO



## OHDSI Italia



**Members:**



**30+ people**  
**20 institutions**  
**(14 data partners)**

...and it's just the beginning!

**Kick-off: June 15<sup>th</sup> 2022**

**1<sup>st</sup> Goal:**  
OHDSI Italia paper  
in 2023



- **Lucia Sacchi** (University of Pavia, SIBIM (Italian Society for Biomedical Informatics))
- **Riccardo Bellazzi** (University of Pavia, SIBIM)
- **Matteo Gabetta** (Biomeris)
- **Mauro Bucalo** (Biomeris)
- **Eleonora Ferretti** (AUSL-IRCCS di Reggio Emilia)
- **Gianluigi Galli** (Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico)
- **Paolo Balli** (Fondazione Istituto Nazionale dei Tumori)
- **Annalisa Trama** (Fondazione Istituto Nazionale dei Tumori)
- **Sara Boveri** (IRCCS Policlinico San Donato)
- **Emanuele Girani** (IRCCS Policlinico San Donato)
- **Irene Tramacere** (Fondazione IRCCS Istituto Neurologico Carlo Besta)
- **Nicola Gentili** (IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) 'Dino Amadori')
- **Valentina Tibollo** (Istituto Maugeri IRCCS)
- **Matteo Puntoni** (Azienda Ospedaliera Universitaria di Parma)
- **Luigia Scudeller** (Policlinico S. Orsola-Malpighi, Bologna)
- **Giuseppe Caruana** (IRCCS ISMETT)
- **Catherine Kiersy** (Fondazione IRCCS Policlinico San Matteo, Pavia)
- **Riccardo Spizzo** (IRCCS Centro di Riferimento Oncologico, Aviano)
- **Mirko Orsini** (DataRiver)
- **Enrico Calanchi** (DataRiver)
- **Dario Montermini** (PGMD)
- **Matteo Spezia** (PGMD)
- **Stefano Dalmiani** (Fondazione Toscana Gabriele Monasterio per la Ricerca Medica e di Sanità Pubblica)
- **Mario Cannataro** (Università di Catanzaro)
- **Andres Vitali** (Casa di Cura Privata del Policlinico, Milano)
- **Massimo Caprino** (Casa di Cura Privata del Policlinico, Milano)
- **Massimo Corbo** (Casa di Cura Privata del Policlinico, Milano)



FRIDAY

OHDSI Italia: the Italian national node of OHDSI Europe

Lead: Lucia Sacchi



# 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?**

