1]Jacob S. Zelko 1, 2]Malina Hy 1, 2]Varshini Chinta [1]Georgia Tech Research Institute [2]Georgia Institute of Technology

My JuliaCon proceeding

[

ABSTRACT

2 - 3 sentences

Keywords

Observational Health, OMOP Common Data Model, Database Management, Characterization

1. Introduction

Paragraph 1: Background on observational health research
Observational health data captures a comprehensive view of patient cohort populations, and they are often used to assess incidence rates, prevalence rates, risk factors, and prognosis associated with physical and mental health conditions and treatment interventions. By analyzing healthcare records and clinical events, observational health research can perform cohort characterization, population-estimation, and patient-level predictions to evaluate population health, target interventions and prevention measures, and ultimately transform healthcare policy.

However, current observational health and epidemiological surveys are limited by their data collection and administration methods, sample size representativeness, cohort phenotype definitions, and statistical/algorithmic methods to combine outside, disparate data sources. There is a need for timely, unbiased, accurate, and quality data regarding the mental illness and the ability detect and characterize trends in prevalence, severity, and the association with other chronic illnesses/morbidities.

This effort could help leading health officials implement better informed interventions that will increase equitable access to mental healthcare across vulnerable populations, support efforts to decrease acute mental health-related morbidities, and improve efficacy of health spending.

Paragraph 2: What is OHDSI + HADES R ecosystem
Observational Health Data Sciences and Informatics (OHDSI) is a global, collaborative, interdisciplinary, and open-science network of researchers, healthcare professionals, students, patients, other individuals and organizations committed to large-scale health data analytics and generating reliable evidence, with the goal of promoting better health decisions and better care.

Since current observational health data sources are aggregated at different levels and cannot be generalized to the overall population, the data is often incomplete, inaccurate, and prone to issues of bias, error handling, and multidimensional complexity. The goal of OHDSI is to produce reproducible, robust, and reliable evidence by developing standardized data formats and research protocol designs to enable large-scale collaborative research and improve the transparency and scalability of observational health research.

The development of the Observational Medical Outcomes Partnership (OMOP) was created in 2009 to address concerns related to data type, study design, and data privacy concerns. The OMOP Common Data Model (CDM) is a "person-centered" information model with standardized relationships and structures with other event tables to create a longitudinal view of all clinical, healthcare-related events for each person.

OHDSI also develops open-source analytical and software tools, such as HADES and ATLAS, which features a comprehensive set of functions needed to conduct network studies from assessing demographic information to generating summary statistics and interactive dashboards. The best practices set forth by OHDSI has improved the quality and reliability of evidence in observational studies, and it improved network research study collaboration, supported large-scale analytics, and standardized data structures and medical vocabulary.

Health Analytics Data-to-Evidence Suite (HADES) is a collection of 20 R packages, developed by the OHDSI team, that interact directly with the OMOP CDM to support large-scale analytics, including cohort construction (CirceR, CohortGenerator), population-level estimation (CohortMethod), patient-level prediction (PatientLevelPrediction), and return statistics, figures, and tables.

Paragraph 3: Why tooling was needed

2. Methods

2.1 Tooling for Observational Health Research

Paragraph 1: Introduction of novel packages
Paragraph 2: Development philosophy and algorithmic design

2.2 Example Use Cases

2.2.1 Minimal Characterization Study. Paragraph 1: Condensed version of tutorial: https://juliahealth.org/OMOPCDMCohortCreator.jl/dev/tutorials/
Paragraph 2: Example of using results (base off this paragraph): From here, the potential to calculate valuable metrics disease prevalence across niche populations is rich and opens up the possibility to do even more with these subpopulations one could generate (e.g. cross walking with survey panel data, area estimation, etc.). Furthermore, I have tested this but have not had an opportunity to write a tutorial on the matter, but these functionalities work with Distributed.jl to enable non-blocking asynchronous code execution (still working out exact approach/functionalities) to fully utilize Julia's performance when working and transforming large amounts of data that is often found in observational studies.

2.2.2 Composability within Ecosystem. Paragraph 1 - 2: Distributed computing + Distributed.jl (maybe with Memoization) + Light benchmarking (?)
Paragraph 3: Support for other databases + FunSQL.jl + SQL output

3. Discussion

Paragraph 1: How it is being used

Paragraph 2: Next steps and Future - base off of this: As I was recently funded another year to continue my health disparities research (yay!), I have great ambitions to build out observational health research capacity in JuliaHealth through avenues such as: Finalize a submission to JuliaCon Proceedings (need to finish writing tests across this portfolio and adding documentation in the coming weeks) Partnering on potentially running studies within the Julia community using these tools Scoping out better support for survey and panel data Building additional tooling around observational health research (such as determining fairness metrics, easily deployable interactive results explorers, interop with OHDSI tools, etc.) Lowering barrier to entry of prototyping large scale studies for researchers

4. Conclusion

Paragraph 1: Final thoughts

5. Acknowledgements

6. References

7. References