Examining the Differences in Baseline Characteristics of One-code and Two-code Phenotype Algorithms

Presenters: Pranav Bhimani, Raechel Davis

BACKGROUND

- The guidance and implications regarding broad and narrow phenotype algorithm (PA) use remain unclear.
- Broad PAs requiring one diagnostic code identify a greater number of subjects, often producing higher sensitivities, albeit with lower positive predictive values (PPVs).
- Narrow PAs that require a second diagnostic code during some timeframe after the first diagnostic code, are often accompanied by lower sensitivities but produce higher PPVs.
- The objective of this study was to compare the similarity of baseline characteristics for phenotype algorithms requiring one and two diagnostic codes for health outcomes in therapeutic areas of neurology, immunology, oncology, and cardiology using six real world databases.

METHODS

- A network of six US observational databases that were transformed to the Observational Medical Outcomes Partnership (OMOP) Common Data Model version 5.3.1 were used [1].

Database	Years	Number of Persons (millions)	Median Follow-up (years)
CCAE	2000-2021	162	1.56
MDCD	2006-2020	33	1.52
MDCR	2000-2021	10	2.46
Optum DOD	2007-2021	92	1.48
Optum EHR	2007-2021	105	2.63
Pharmetrics	2013-2021	162	3.25

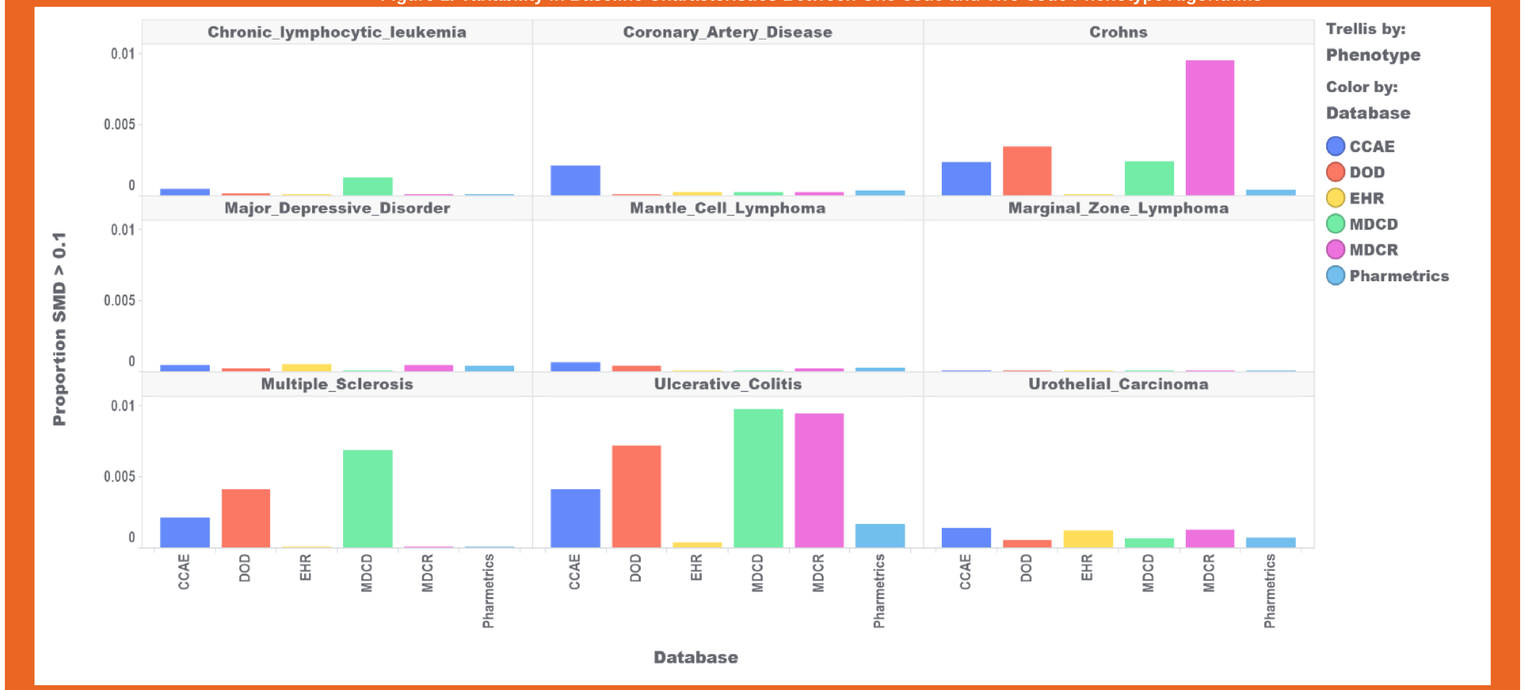
Table 1. US Databases with inpatient and outpatient visit types. Data type is Insurance Claims for: IBM MarketScan® Commercial Claims and Encounters (CCAE), IBM MarketScan® Multi-State Medicaid Database (MDCD), IBM MarketScan® Multi-State Medicare Database (MDCR), IQVIA® Adjudicated Health Plan Claims Data (PharMetrics), Optum® Clinformatics® Data Mart (DOD). Data type is Electronic Health Records for Optum® Clinformatics® Electronic Health Records (Optum EHR). The use of Optum and CCAE was reviewed by the New England Institutional Review Board and was determined to be exempt from broad Institutional Review Board approval as this project did not involve human subject research.

- PAs for associated outcomes within each therapeutic area were analyzed.

Therapeutic Area	Condition
Neurology	Multiple Sclerosis
	Major Depressive Disorder
Immunology	Ulcerative Colitis
	Crohn's Disease
Oncology	Mantle Cell Lymphoma
	Chronic Lymphocytic Leukemia
	Marginal Zone Lymphoma
	Urothelial Carcinoma
Cardiology	Coronary Artery Disease

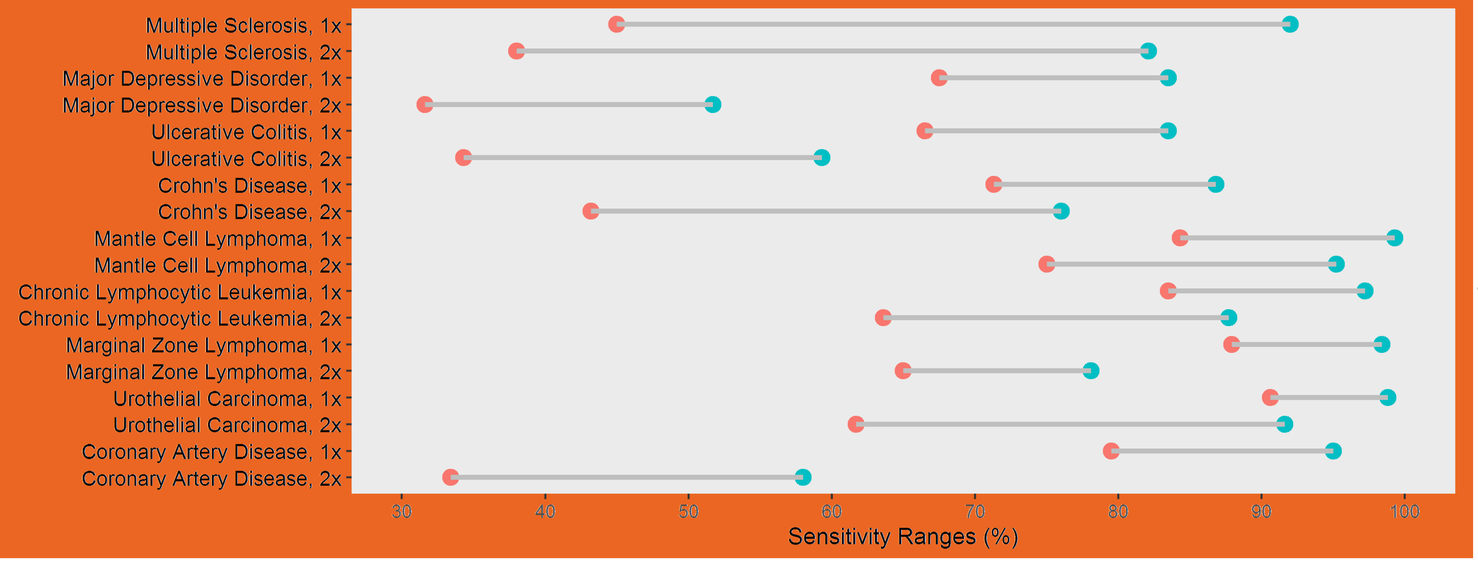
- The ATLAS tool was used to create PAs and generate cohorts [2].
- The PheValuator method provided performance characteristics, i.e., sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), associated with each PA [3].
- The Cohort Diagnostics tool allowed for evaluation of PAs at a cohort-level, providing a comparison of baseline covariates between the one-code and two-code algorithms [4].
- SMD > 0.1 is used as an ad-hoc heuristic for what constituted a substantial difference in compared covariates [5].
- Baseline characteristics were compared by generating a proportion. The numerator is the number of covariates with an absolute standardized mean difference (SMD) > 0.1 and the denominator is the total number of covariates compared between phenotypes.

Figure 1. Variability in Baseline Characteristics Between One-code and Two-code Phenotype Algorithms



Comparing baseline covariates between broad and narrow phenotype algorithms provides a more complete understanding of algorithm differences.

Figure 2. Ranges of Sensitivities Across all Databases for Broad (One-code) and Narrow (Two-code) Phenotype Algorithms, red = lowest; blue = highest



RESULTS

- In six of the nine outcomes there was minimal variability in the comparison of baseline covariates.
- Comparisons between the Ulcerative Colitis and Crohn's Disease one-code and two-code PAs showed the greatest variability, while the Multiple Sclerosis comparison showed a moderate level of variability (Figure 1).
- Sensitivities of one-code algorithms were observed to be higher and less variable across databases than two-code algorithms (Figure 2).

View Results Interactively:

https://data.ohdsi.org/PhenotypeComparisons/

CONCLUSIONS

- Comparisons of baseline characteristics in 1-code and 2-code PAs in most (6 of the 9) outcomes showed minimal variability.
- For outcomes in specific therapeutic areas such as immunology, greater variability in baseline covariates may be present when comparing 1 and 2 code algorithms.
- Comparison of the similarity of baseline covariates between phenotype algorithms provides a more complete understanding of algorithm differences.

REFERENCES

1. Voss EA MR, Matcho A, Ma Q, Knoll C, Schuemie M, et al. . Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. . J Am Med Inform Assoc 2015;22(3):553-64.
2. Schuemie, M. and DeFalco F. (2019). OHDSI Analytics Tools. In Observational Health Data Sciences and Informatics The Book of OHDSI (pp. 109-123). Independently published.
3. Swerdel JN, Hripcsak G, Ryan PB. PheValuator: Development and evaluation of a phenotype algorithm evaluator. Journal of Biomedical Informatics. 2019;97:103258.
4. Rao G, Schuemie M, Ryan P, Weaver J, Gilbert J (2022). CohortDiagnostics: Diagnostics for OHDSI Cohorts. https://ohdsi.github.io/CohortDiagnostics, https://github.com/OHDSI/CohortDiagnostics.
5. Austin PC, Steyerberg EW. The number of subjects per variable required in linear regression analyses. J Clin Epidemiol. 2015;68(6):627-36.

Jill Hardin[1,2], Pranav Bhimani[1,2,3], Raechel Davis[1,2,4], Joel N. Swerdel[1,2]
[1]Janssen Research and Development, Titusville, NJ, USA; [2]Observational Health Data Sciences and Informatics (OHDSI), New York, NY; [3]Weill Cornell Graduate School, New York, NY; [4]Yale School of Public Health, New Haven, CT