

Phenotype Development for Neonatal Hypoxic Ischemic Encephalopathy Using Electronic Health Record and Claims Datasets

Star Liu¹, Tony Widenor², Danielle Boyce¹, Gowtham Rao³, Evan Minty^{*4} and Khyzer Aziz^{*1}

¹Johns Hopkins School of Medicine, Baltimore, MD, USA

²Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

³Janssen Research and Development, Titusville, NJ, USA

⁴O'Brien Institute for Public Health, University of Calgary, Canada

*These authors contributed equally

Background

Hypoxic Ischemic Encephalopathy (HIE) is a type of brain injury among newborns. The incidence of HIE is 1.5-2.5 per 1000 live births in developed countries, and, among infants with HIE, 40-60% die by 2 years of age or develop severe disabilities.¹⁻⁴ HIE occurs primarily due to acute perinatal asphyxia – a state where impaired cerebral blood flow and oxygen delivery to the brain result in decreased energy production, leading to cell death.^{5,6} Neonatologists and pediatricians are further challenged by the differential diagnosis of neonatal encephalopathy (2-6 per 1000 term births), of which HIE accounts for a portion of cases.⁷ Better clinical characterization, clinical phenotyping, and identification of subgroups most responsive to existing therapies for patients with HIE could potentially lower the risk of adverse neurodevelopmental outcomes. The objective was to develop a phenotype for HIE and assess the algorithm using Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) datasets and tools.

Methods

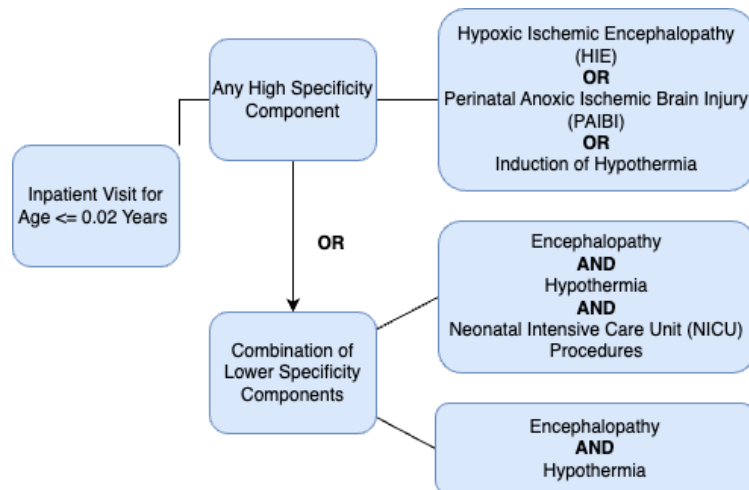


Figure 1. HIE Phenotype algorithm pathways.

Phenotype. The baseline candidate for the HIE cohort was defined as a patient with age < 0.02 years, an inpatient visit, and a single condition occurrence of the HIE, Perinatal Anoxic Ischemic Brain Injury, or the Induction of Hypothermia (Figure 1). We compared the cohort distribution of the individual components of the baseline candidate across 5 OMOP datasets (3 claims datasets and 2 EHR datasets). We assessed the phenotype algorithm after increasing the complexity of diagnostic and therapy criteria. We examined the inclusion of Encephalopathy + Hypothermia, and Encephalopathy + Hypothermia + Neonatal Intensive Care Unit (NICU) procedures, where NICU procedures included a broader definition of echocardiography and Hypothermia included condition (Dx), medication (Rx) and multimodal monitoring and treatment

(MMT). We created ‘marginal’ cohorts that examined the marginal contribution of these combination cohorts that would not have otherwise been included by virtue of having a higher specificity code.

Concept sets were iterated using PPhenotype Observed Entity Baseline Endorsement (PHOEBE) (<https://data.ohdsi.org/PHOEBE/>). The HIE cohort was built using ATLAS (<https://github.com/OHDSI/Atlas>). Cohort Diagnostics (<https://github.com/OHDSI/CohortDiagnostics>) was used to evaluate the phenotype algorithm for cohort overlap and covariate alignment across data sources. Sample covariate alignments were presented using the Optum EHR dataset.

Results

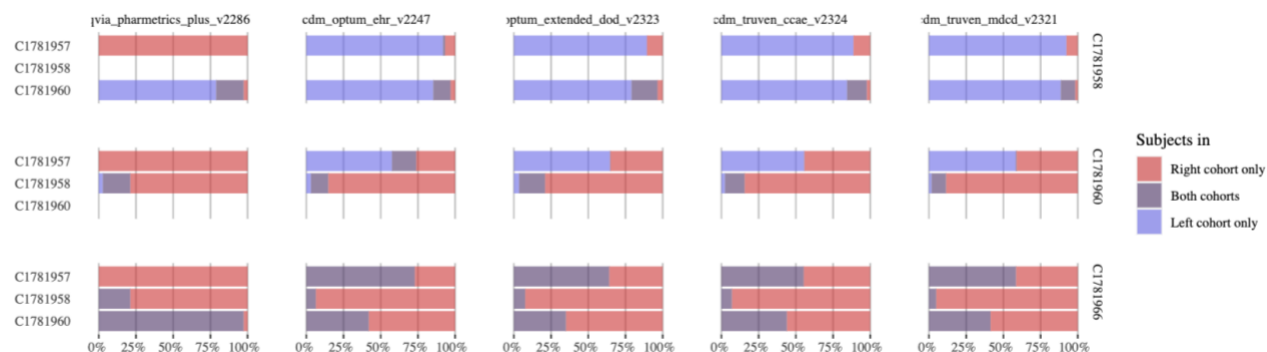


Figure 2. Cohort overlap among 5 datasets (IQVIA Pharmetrics, Optum EHR, Optum EHR Extended, IBM CCAE, and IBM MDCCD). C1781957 is the concept ID for Perinatal Anoxic Ischemic Brain Injury (PAIBI) diagnostic code (Dx). C1781958 is the concept ID for the Induction of Hypothermia. C1781960 is the concept ID for HIE Dx. C1781966 is the concept ID for the union of HIE Dx, PAIBI Dx, and the Induction of Hypothermia. The top group compares concepts to C1781958. The second group compares concepts to C1781960. The bottom group compares concepts to the union concept ID C1781966 (right cohort).

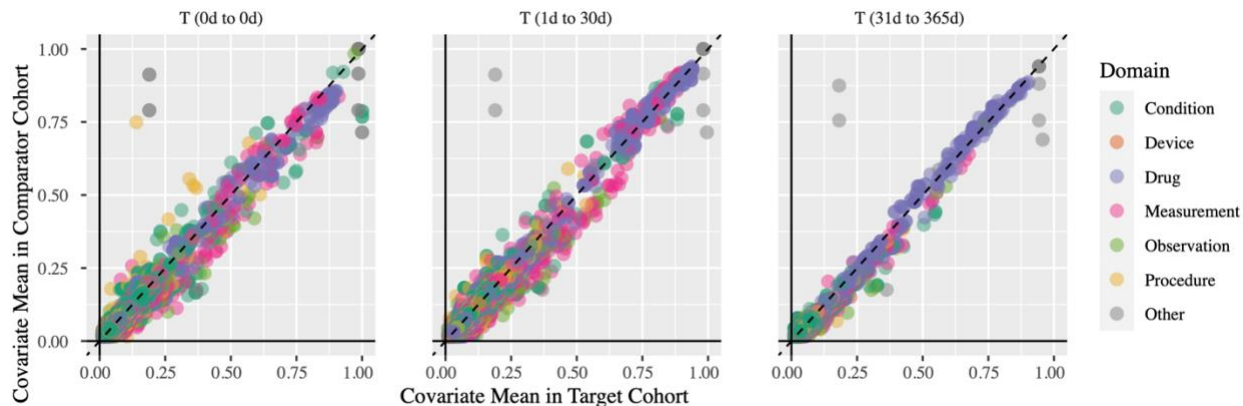


Figure 3. Covariate alignment comparing the target cohort (2 occurrences of HIE) and the comparator cohort (Induction of Hypothermia), using the Optum EHR dataset.

Across the 5 OMOP data sources, the Induction of Hypothermia (C1781958) contributed the least to the HIE cohort (Figure 2). HIE (C1781960) and PAIBI (C1781957) demonstrated greater overlap with the union concept (C1781966) with 35-44% and 55-73% overlap respectively. This suggests that both HIE and PAIBI made greater contribution to the HIE cohort compared to the Induction of Hypothermia. Thus, depending on the data source, both HIE and PAIBI are possible inclusion criteria in a phenotype algorithm in a network study. Similarly, the covariate plot comparing 2 occurrences of HIE and the Induction of Hypothermia shows alignment among the covariates (Figure 3). We also examined the marginal contribution of more complex definitions. Inclusion of concept sets representing encephalopathy and hypothermia increased the cohort size by 8-10%. Inclusion of concept sets representing encephalopathy, hypothermia, and NICU procedures increased the cohort size by 4-6%.

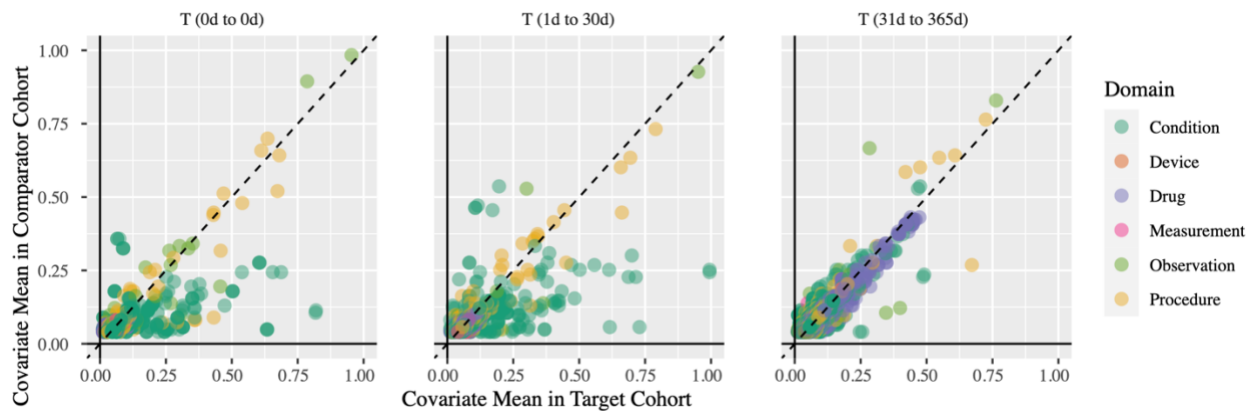


Figure 4. Covariate alignment comparing the target cohort (2 occurrences of HIE or PAIBI) and the comparator cohort (encephalopathy + hypothermia Dx, Rx, or MMT + without HIE or the induction of hypothermia, using the IQVIA Pharmetrics dataset).

We compared the algorithm of having 2 occurrences of HIE or PAIBI (highly specific) with the algorithm of encephalopathy + hypothermia (Dx, Rx, or MMT) + without HIE or the induction of hypothermia (i.e. the marginal contribution of that combination cohort). While covariate alignment was observed in some datasets (i.e., the Optum EHR dataset), it was not consistent across datasets. In the IQVIA Pharmetrics dataset, condition variables' (green dots) means did not align (Figure 4). Thus, the combination component was not added to the baseline candidate algorithm.

We also compared the algorithm of having 2 occurrences of HIE or PAIBI with the algorithm of encephalopathy + hypothermia (Dx, Rx, or MMT) + NICU procedures (echocardiography or echoencephalography) + without HIE, PAIBI, or the induction of hypothermia. When NICU procedures were included, we did not observe the level of covariate alignment observed in previous cohort comparisons. Possible explanations for the misalignment are critical care trajectories and trajectories that matched procedure criteria. The lack of understanding of the misalignment, and the small marginal contribution to the person count, prompted us to exclude it from the baseline algorithm. The final candidate phenotype for HIE was the union of HIE, PAIBI, or the Induction of Hypothermia (C1781966).

Conclusion

We established an entry definition for the HIE cohort using OMOP tools and datasets. Each individual component of the baseline candidate exhibited significant contribution to the HIE cohort. As such, satisfying any of the individual components qualifies for cohort entry. Several limitations exist. Claims and EHR databases' inherent workflows and data granularities potentially lend themselves to observed inconsistencies. The inclusion of both clinical diagnosis and disease severity in the phenotype warrants further assessment of its impact on the phenotype performance. Future work will further refine the phenotypes by exploring other combinatoric outcomes and inclusion of procedural criteria. Additionally, we will assess pheValuator results to understand the true negative and false negative contributions across OMOP datasets. Evaluating the baseline candidate and the impact of increasing cohort complexity support the development of highly specific phenotype for early identification of HIE.

References

1. Shankaran S, Laptook AR, Ehrenkranz RA, Tyson JE, McDonald SA, Donovan EF, et al. Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy. *N Engl J Med* 2005 -10-13;353(15):1574-1584.
2. Lai M, Yang S. Perinatal hypoxic-ischemic encephalopathy. *J Biomed Biotechnol* 2011;2011:609813.

3. Hypoxic ischaemic encephalopathy - ScienceDirect.
4. Chalak LF, Sánchez PJ, Adams-Huet B, Lupton AR, Heyne RJ, Rosenfeld CR. Biomarkers for severity of neonatal hypoxic-ischemic encephalopathy and outcomes in newborns receiving hypothermia therapy. *J Pediatr* 2014 -03;164(3):468-474.e1.
5. Graham EM, Ruis KA, Hartman AL, Northington FJ, Fox HE. A systematic review of the role of intrapartum hypoxia-ischemia in the causation of neonatal encephalopathy. *Am J Obstet Gynecol* 2008 -12;199(6):587-595.
6. Northington FJ, Chavez-Valdez R, Martin LJ. Neuronal cell death in neonatal hypoxia-ischemia. *Ann Neurol* 2011 -05;69(5):743-758.
7. Russ JB, Simmons R, Glass HC. Neonatal Encephalopathy: Beyond Hypoxic-Ischemic Encephalopathy. *Neoreviews* 2021 -03;22(3):e148-e162.