

Adapting, Extending, and Reimagining the OHDSI Ecosystem for Pediatric Research: The PEDSnet Experience

Hanieh Razzaghi PhD MPH^{1,2}, Charles Bailey MD PhD^{1,2}

¹Children's Hospital of Philadelphia, ²Perelman School of Medicine University of Pennsylvania

Background

PEDSnet is the largest multi-institutional pediatric learning health network in the United States, spanning a dozen children's hospital systems. While the OHDSI ecosystem provides a powerful foundation for observational research, participation in a pediatric research network introduces requirements that are not fully addressed by default implementations. These include examples such as developmental metrics, specialized clinical workflows, rare disease populations, and the need for rigorous cross-site comparability. To serve a diverse portfolio of research use cases, PEDSnet has systematically adapted the OMOP Common Data Model (CDM), extended existing OHDSI tools, and developed novel data quality infrastructure. This abstract describes that evolution and the lessons learned.

Methods

PEDSnet's adaptations span three interconnected areas. First, we extended the OMOP CDM to capture pediatric-specific clinical information not accommodated by the standard model. This included adding gestational age to the Person table, care site specialty to the Care Site table, and introducing two new tables — `measurement_organism` and `immunization` — to represent microbiology results and vaccination data. We additionally precomputed commonly used pediatric metrics, such as age- and sex-adjusted height and weight z-scores, and made targeted deviations from standard conventions (e.g., in `condition_source_value` and `visit_detail`) to better preserve source data fidelity.

Second, we extended existing OHDSI tools to meet network-specific operational needs. ATLAS was augmented to support site or institutional stratification of results, enabling detection and characterization of site-level heterogeneity that is critical given the regional variation across children's hospitals. We also extended the underlying architecture to allow ATLAS to query data organized in the PCORnet common data model, enabling cross-network research. To improve performance and reduce infrastructure cost, several HADES packages were re-implemented to run on Trino, an open-source distributed query engine.

Third, recognizing that the OHDSI Data Quality Dashboard (DQD), while powerful, was designed for general-purpose use, we developed two complementary data quality packages tailored to our network's needs: one focused on longitudinal tracking of data quality metrics over time to monitor network evolution, and a second enabling study-specific quality checks designed to surface heterogeneity and anomalies in highly specialized cohorts, such as rare disease populations. Together, these tools provide a systematic framework for operating on a data resource that changes as clinical practices, source systems, and network membership evolve.

Results

These adaptations have materially expanded PEDSnet's capacity to support high-quality pediatric research. CDM extensions have enabled studies requiring developmental context (e.g., prematurity-adjusted analyses) and infection-specific endpoints that would otherwise require ad hoc workarounds. Site-stratified ATLAS functionality has enabled network feasibility assessments and site selection. Trino-based query execution has allowed for addressing larger datasets, reducing query runtimes and making more complex queries feasible. The longitudinal and study-specific data quality frameworks have identified data anomalies that standard DQD checks missed, improving the reliability of network-wide analyses and cohort definitions in rare disease studies.

Conclusion

PEDSnet's experience illustrates both the necessity and the tractability of adapting the OHDSI ecosystem for specialized research contexts. Many of the extensions described here — pediatric-specific CDM fields, site-level stratification in ATLAS, open-source execution environments, and longitudinal data quality tracking — may have broader relevance beyond pediatrics. We discuss how these contributions can inform OHDSI community standards, where divergence from defaults creates sustainability risk, and how networks like PEDSnet can better contribute adaptations back to the ecosystem to benefit the broader community.