

Using the OMOP Cohort Table to Link PRoMPT BOLUS Clinical Trial Participants to the PEDSnet Research Network

Co-Authors

Levon H. Utidjian MD MBI¹, Aqsa Khan MS BHI¹, Sahal Master BDS MPH¹, Atzael B. Campos MSDS¹, Ruchi Singh PhD¹, Aliyah Jones BA¹, Grace Park MPH¹, Sam Boss BS¹, Amy Goodwin Davies PhD¹, Fran Balamuth MD PhD¹, Scott L. Weiss MD MSCE², Michelle Denburg MD MSCE¹, Julie C. Fitzgerald MD PhD¹

¹: Children's Hospital of Philadelphia, Philadelphia, PA

²: Nemours Children's Hospital, Wilmington, DE

Background

The PRAgMatic Pediatric Trial of Balanced vs. nOrmaL Saline FLUId in Sepsis (PRoMPT BOLUS) is a large, multicenter, randomized clinical trial comparing two commonly used fluid resuscitation treatments for pediatric sepsis to determine if one is more effective and safer than the other.(1) The main trial assigns enrolled patients to different fluid exposure arms, and plans to follow patients for up to ninety days with manual collection of select clinical data points, such as specific lab, vital sign, treatment and diagnosis data. Manual chart review is a standard practice in clinical trials for data collection but is a time-consuming and expensive process that is difficult over long periods and is at risk of data entry errors. Linking clinical trial participants to their electronic health record (EHR) data for automated data extraction is known to more quickly and accurately capture clinical data, reduce the possibility of missed data, and allow for longer term outcome tracking.(2-4) However, scaling such EHR data extraction across multiple trial sites is challenging.

Eight of the PRoMPT BOLUS trial sites are members of the PEDSnet research network, which provides a central database storing multiple clinical data domains extracted and transformed from EHRs using a modified version of the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).(5) Access to a database like PEDSnet allows the study team to examine and analyze more longitudinal clinical data from before and after trial entry, providing more clinical data points than could be reasonably captured by chart review.

The OMOP CDM's cohort table can be used to store the trial enrollment details for participants identified by study recruitment teams. However, as these enrollments are done manually at the sites in real time rather than by a computable phenotype using rules or probabilistic cohort definitions, a key step in preparing the data set is to link clinical trial participants to their corresponding identifiers in the PEDSnet database. Here we

demonstrate a process for linking clinical trial participants to their PEDSnet data to increase the available EHR data for analysis of additional outcomes.

Methods

Clinical trial enrollment information, including participant identifiers and trial enrollment dates, is uploaded each quarter during the regular PEDSnet data refresh cycle to the cohort table. Each trial participant's identifier is mapped to the unique PEDSnet patient identifier for that patient, keeping patients deidentified to the PEDSnet data coordinating center (DCC). The most recent OMOP cohort table conventions are used in the PEDSnet version, with additional fields to handle potential dates for early study withdrawal and other study-specific participant identifiers. An EHR visit identifier is not available from the clinical trial enrollment data, so a combination of participant identifier and enrollment start date is used to determine the qualifying visit when the patient was enrolled. The PEDSnet DCC data scientists attempt to match each trial enrollment to the PEDSnet database algorithmically (Table 1).

Table 1. Initial Steps in Algorithm Linking Trial Enrollments to PEDSnet Database

Step	Rationale
(1) Confirm the trial enrollment participant is present in PEDSnet database	Ensures trial participant is correctly identified and present in current database
(2) Confirm trial enrollment date is in the data submission window	Ensures trial participant is not enrolled after data available in current database version
(3) Confirm the trial participant has a qualifying visit	Ensures trial participant has an emergency department and/or inpatient visit
(4) Confirm the qualifying visit is on the date of trial enrollment	Ensures that there is not an error in enrollment date entry

If participants cannot be matched by the initial algorithm, the DCC data scientists perform an internal review, looking for potential issues such as alternate visit types or visits starting the day prior to the date of enrollment. If the DCC cannot resolve the discrepancies, they are referred to the external sites for further review.

Results

This linkage process has been performed over 5 quarters to date and results are summarized in Table 2. The described algorithm can match 96.8-98.3% of all eligible enrollments for which there is data available in the PEDSnet database that quarter. Internal review by DCC data scientists further increases the match rate to 98.4-99.8% of eligible trial enrollments. The most common issue found on internal review is that participants have a qualifying visit within 24 hours of the reported enrollment date, usually reflecting

enrollments occurring after midnight on the next calendar day. External review has most often found manual errors in the entry of the enrollment date submitted. Over time, the number of enrollments needing external review has represented <1% of the current total of eligible enrollments.

Table 2. Clinical Trial Enrollment Matching by Quarter

Quarter Ending	Trial Enrollment	Matched Algorithmically (% of Eligible)	Matched after DCC Review (% of Eligible)	External Site Review Needed (% Eligible)
Mar 2024	1583	1438 (98.2%)	1439 (98.4%)	24 (1.6%)
Jun 2024	1930	1873 (98.3%)	1889 (99.1%)	17 (0.9%)
Sep 2024	2155	2033 (96.8%)	2092 (99.6%)	8 (0.4%)
Dec 2024	2327	2198 (97.3%)	2241 (99.2%)	16 (0.8%)
Mar 2025	2377	2291 (98.1%)	2330 (99.8%)	4 (0.2%)

*Only 7 sites contributed data in March 2024, all other quarters have 8 sites contributing.

Conclusions

We have demonstrated a process to link clinical PRoMPT BOLUS trial enrollments to the PEDSnet database with high rates of matching. The most commonly encountered issue on internal review, that qualifying visit dates and enrollment dates occur within 24 hours of each other, can be incorporated into the algorithmic approach to further automate the process. While we utilized standard OMOP cohort table elements in our matching process, inclusion of additional elements such as a separate initial visit identifier could be incorporated in future studies to further aid the linkage process. Lessons learned from our approach could be useful in assisting other clinical trials in linking to external research databases.

References

1. Weiss SL, Balamuth F, Long E, Thompson GC, Hayes KL, Katcoff H, et al. PRagMatic Pediatric Trial of Balanced vs nOrmaL Saline FLUId in Sepsis: study protocol for the PRoMPT BOLUS randomized interventional trial. *Trials*. 2021;22(1):776.
2. Miller TP, Li Y, Kavcic M, Troxel AB, Huang YS, Sung L, et al. Accuracy of Adverse Event Ascertainment in Clinical Trials for Pediatric Acute Myeloid Leukemia. *J Clin Oncol*. 2016;34(13):1537-43.
3. Miller TP, Li Y, Getz KD, Dudley J, Burrows E, Pennington J, et al. Using electronic medical record data to report laboratory adverse events. *Br J Haematol*. 2017;177(2):283-6.

4. Khan MS, Usman MS, Talha KM, Van Spall HGC, Greene SJ, Vaduganathan M, et al. Leveraging electronic health records to streamline the conduct of cardiovascular clinical trials. *Eur Heart J*. 2023;44(21):1890-909.
5. Forrest CB, Margolis PA, Bailey LC, Marsolo K, Del Beccaro MA, Finkelstein JA, et al. PEDSnet: a National Pediatric Learning Health System. *J Am Med Inform Assoc*. 2014;21(4):602-6.