

Target Trial Emulation of Intravenous versus Oral Acetaminophen in Febrile Pediatric Inpatients: An OMOP-CDM Feasibility Study

Hong Kang Lee¹, Suhyun Kim¹, Kwangsoo Kim¹, Hee-Soo Kim^{2,†}

¹ Department of Transdisciplinary Medicine, Seoul National University Hospital, Seoul 03122, Republic of Korea

² Department of Anesthesiology and Pain Medicine, Seoul National University Hospital, Seoul 03080, Republic of Korea

† Corresponding author: Hee-Soo Kim (dami0605@snu.ac.kr)

Background

Generating pediatric real-world evidence within the OHDSI ecosystem remains difficult for medication-route decisions in very young inpatients [4]. Intravenous (IV) and oral acetaminophen (APAP) are selected for clinically different patients, creating confounding by indication, while randomized route-comparison evidence in this age group is scarce. OMOP-CDM extract-transform-load processes may also fail to preserve sub-daily medication and temperature timestamps, limiting granular assessment of antipyretic response. We evaluated the feasibility of a prespecified target-trial emulation [1] comparing initial IV versus oral APAP among febrile inpatients younger than 24 months and documented the data-model requirements for such analyses. This report presents the feasibility phase of an ongoing study.

Methods

We conducted a retrospective cohort study using 2019–2024 EHR data mapped to the OMOP-CDM at a tertiary care hospital. Cohort-entry logic was first expressed as an ATLAS candidate cohort. Because that expression used broader age, route, indication, and temporal windows than the protocol, downstream restrictions generated the locked analytic cohort; differences between the candidate definition and the final analytic cohort were documented.

The submitted target-trial specification required inpatients younger than 24 months, at least 30 days of prior observation, maximum recorded temperature ≥ 38.0 °C within 12 hours before time zero, and no APAP exposure during the preceding 48 hours. Treatment strategies were initial IV versus initial oral APAP at the first eligible administration, defining time zero.

Because sub-daily post-index timestamps were incompletely preserved, effectiveness and treatment-burden outcomes were assessed on the next calendar day. Day-level analyses do not themselves validate the protocol's exact hour-based eligibility windows. Any next-day afebrile status required at least one recorded temperature < 38.0 °C; sustained next-day afebrile status required all recorded temperatures < 38.0 °C. Exploratory safety was defined as recorded AST or ALT elevation $\geq 2\times$ the upper limit of normal within 7 days among patients with normal baseline values and follow-up measurements. Process outcomes were any next-day APAP continuation, at least three administrations, and mean administration count.

Propensity scores were estimated by logistic regression using age in months, sex, prior observation ≥ 180 days, and documented fever on the calendar day before index. Proxies for oral intolerance, fasting status, peri-operative care, and illness acuity were not reliably time-aligned. Overlap weighting [2] targeted the clinical-equipoise population without discarding patients. Risk differences for binary outcomes and a mean difference for count were reported with 95% confidence intervals from 300 bootstrap replications; weighted outcome-regression models checked the two effectiveness outcomes. E-values [3] were risk-ratio-scale point-estimate sensitivity descriptors, not corrections for residual confounding. The Institutional Review Board of Seoul National University College of Medicine / Seoul National University Hospital granted exemption from review (E-2510-018-1681).

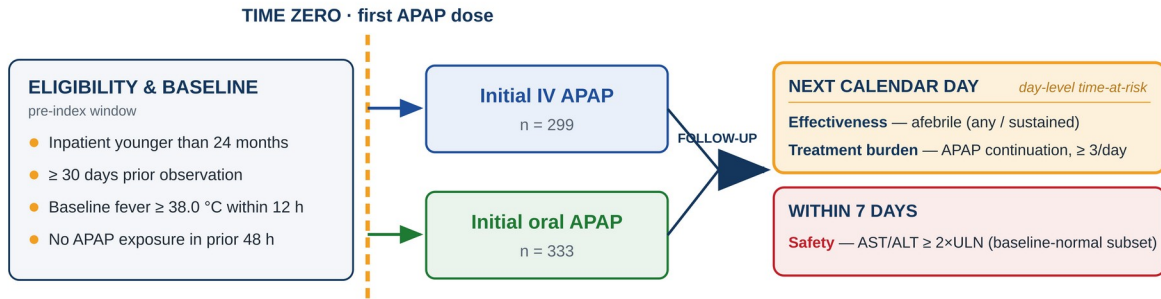


Figure 1. Submitted target-trial specification and outcome windows. Time zero is specified as the first eligible APAP administration after a 48-hour APAP-free period. Reported effectiveness and treatment-burden outcomes use the next calendar day; exploratory safety is assessed within 7 days.

Results

The final cohort comprised 632 patients: 299 receiving initial IV APAP and 333 receiving initial oral APAP. Before weighting, prior-day fever showed the largest imbalance (4.3% versus 32.4%; standardized mean difference 0.778). After overlap weighting, all modeled propensity-score covariates had absolute standardized mean differences <0.03 . Next-day temperature was available for all included patients (IV 299/299; oral 333/333). These counts describe the included cohort only and do not establish representativeness of episodes without next-day temperature records.

In overlap-weighted analyses, any recorded next-day afebrile status was near-universal (IV 100.0% versus oral 99.3%; risk difference $+0.7$ percentage points, 95% CI 0.0 to 2.2), limiting discrimination between routes. Sustained next-day afebrile status was 59.9% versus 61.5% (risk difference -1.6 percentage points, 95% CI -10.3 to 7.4), remaining compatible with clinically meaningful differences in either direction.

In the reported baseline-normal safety subset, overlap-weighted risk of recorded AST or ALT elevation $\geq 2 \times$ the upper limit of normal within 7 days was 1.7% with IV and 3.5% with oral APAP (risk difference -1.8 percentage points, 95% CI -6.3 to 3.0). The interval crossed the null; no comparative hepatic-safety conclusion was drawn.

Exploratory treatment-burden outcomes were lower after initial IV APAP: any next-day continuation (40.9% versus 68.7%; risk difference -27.8 percentage points, 95% CI -36.2 to -19.5), ≥ 3 next-day administrations (35.8% versus 54.8%; risk difference -19.0 percentage points, 95% CI -27.3 to -10.0), and mean administration count (difference -0.63 administrations, 95% CI -0.96 to -0.25). These process outcomes may reflect symptom persistence, scheduled dosing, order defaults, or clinician workflow rather than antipyretic efficacy.

Weighted outcome-regression estimates were materially unchanged ($+0.2$ and -1.7 percentage points). E-values were 1.09 and 1.19 for effectiveness, 3.50 for safety, and 2.75 and 2.43 for treatment burden. These point-estimate sensitivity measures do not resolve residual confounding; the safety confidence interval already includes the null.

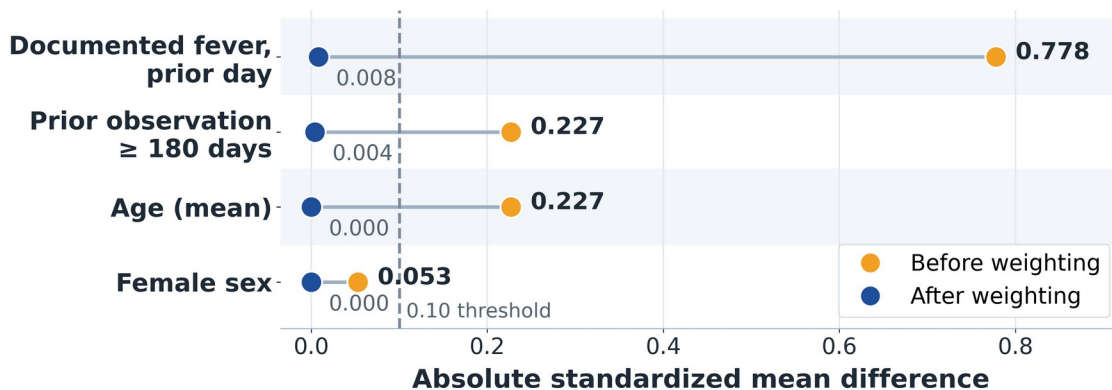


Figure 2. Absolute standardized mean differences before and after overlap weighting. All modeled covariates had absolute SMDs <0.03 after weighting; no patients were discarded.

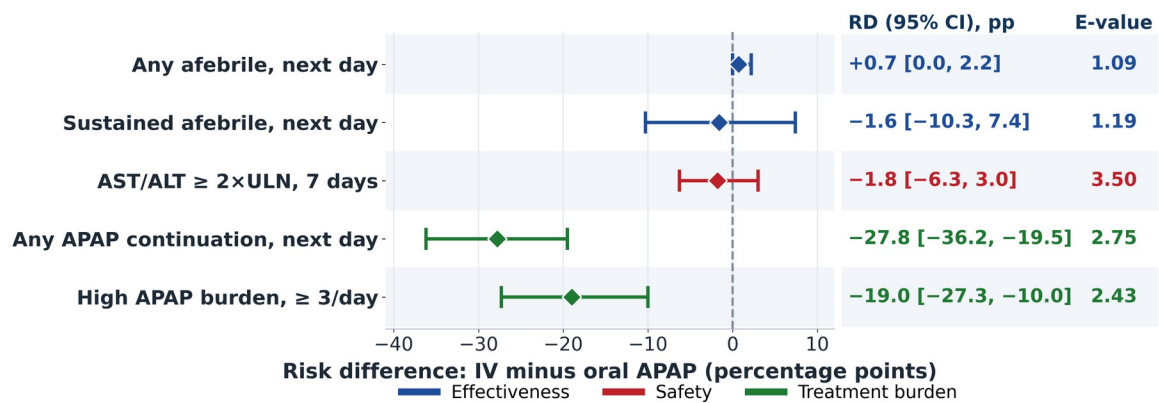


Figure 3. Overlap-weighted risk differences (IV minus oral APAP, percentage points) with 95% confidence intervals and E-values on the risk-ratio scale. Effectiveness estimates were near-null; the largest differences were treatment-burden process outcomes.

Conclusions

This pilot implemented a pediatric target-trial framework for cohort construction and day-level outcome analysis in OMOP-CDM, while identifying requirements for stronger inference. Three implementation lessons emerged. First, incomplete sub-daily post-index timing supported calendar-day follow-up rather than hour-level response assessment. Second, the ATLAS indication window spanned -1 to +1 day and could admit post-index information; eligibility must use pre-index data only. Third, candidate concept sets required product-level review because the enteral set contained a rectal suppository, the IV set contained a fixed-combination injection, and candidate age limits differed by route.

Interpretation is limited by the single-center design, sparse propensity-score model, and clinically driven laboratory ascertainment. Arm-specific effective sample sizes and safety-subset counts are not reported, limiting assessment of weighted precision and the exploratory laboratory comparison. Large-scale propensity-score (LSPS) analyses, negative-control calibration, and cross-site validation are not included in these pilot results. Follow-up research is addressing the limitations identified here, with priorities including source-timestamp verification for exact eligibility windows, validated route-selection covariates, finer temperature-response endpoints, and multi-site evaluation [5]. Updated analyses are outside this feasibility report; these findings motivate further development rather than establish definitive comparative effectiveness or hepatic safety.

Acknowledgements

Supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (RS-2025-02223731).

References

1. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–64.
2. Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. *Am J Epidemiol*. 2019;188(1):250–7.
3. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. 2017;167(4):268–74.
4. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform*. 2015;216:574–8.
5. Schuemie MJ, Ryan PB, Pratt N, et al. Principles of large-scale evidence generation and evaluation across a network of databases (LEGEND). *J Am Med Inform Assoc*. 2020;27(8):1331–7.