

Democratizing Evidence Generation through Vibe Coding with OHDSI Tools: A Use Case in Regional and Seasonal Diseases in Kids*

Author: Patrick Ryan^{1,2}

Acknowledged Contributors: Katherine Ryan^{3*}, Kaia Schuemie^{4*}, Renske Los⁵

1. Johnson & Johnson, Raritan, NJ, USA
2. Columbia University, New York, NY, USA
3. Baldwin School, Bryn Mawr, PA, USA
4. Markland College, Zevenbergen, NL
5. Erasmus MC, Rotterdam, NL

Background

The OHDSI community is committed to generating reliable, reproducible evidence at scale. However, the technical threshold required to translate clinical questions into standardized observational research, including SQL, R, and HADES package fluency, remains a barrier for many domain-expert researchers. Vibe coding, an iterative natural language-driven development paradigm facilitated by large language models (LLMs), presents an opportunity to lower this barrier by translating clinical intent into executable study code. We hypothesized that this workflow could be used to design and implement a complex OHDSI-compliant analysis conceived by early-stage researchers (KR, KS) without requiring manual coding from scratch and enabling appropriate interpretation of results without advanced statistical training.

Methods

We used a human-in-the-loop vibe-coding workflow in which a lead scientist (PR) provided prompts to describe the clinical question specified by the early-stage researchers and technical constraints, and an LLM iteratively drafted and refined R scripts in VS Code using Claude Sonnet 4.6. The final pipeline used OHDSI HADES tools (Capr, CohortGenerator, CohortIncidence, DatabaseConnector) on a large US claims database standardized to OMOP CDM deployed in Databricks. All code was manually quality controlled and independently confirmed.

The use case evaluated seasonal and geographic heterogeneity in 20 pediatric outcomes among children aged 5-15 years. The diseases (allergic rhinitis, atopic dermatitis, burn of skin, cough, diarrhea, dislocation of shoulder joint, fever, fracture of bone, fracture of wrist, fracture of finger, general effects of low temperature (like frostbite), headache, influenza, louse infestation, Lyme disease, nausea, otitis externa ('swimmer's ear'), pedal cycle accident, sprain of ankle, and sprain of knee) were selected based on hypotheses by the early-stage researchers on potential seasonal or regional effects. Target cohorts were defined as state-by-calendar-month groups pooled across 2022-2024 using custom SQL, requiring at least 90 days of prior observation and observation through month end. Outcome cohorts were generated as all qualifying condition occurrence or observation records per concept, with 30-day era collapse.

Incidence analysis used one time-at-risk window from target cohort start to target cohort end (full month). For each target-outcome pair, we estimated incidence proportion (per 1000 persons), with a 30-day clean window for recurrent outcomes. We created U.S. state-month trellis maps in ggplot2/usmap using relative incidence (state-month incidence proportion divided by each outcome's grand mean across all states and months), classified into seven fixed bins. As a contextual comparator, we added a temperature panel based on NOAA 1991-2020 monthly state normals. Condition-level pattern summaries were generated using Kruskal-Wallis tests for variation by month and by a prespecified 4-region grouping, and Spearman correlation between incidence proportion and state-month temperature.

Results

The vibe-coding workflow produced a reproducible, HADES-based analysis pipeline, showing that complex observational study logic can be implemented through iterative natural language prompting rather than de novo manual programming.

We studied more than 1.5 million children aged 5–15 years from January 2022 through December 2024. Figure 1 shows the incidence of 12 diseases (columns) by state and month (rows). Table 1 shows the statistical tests for patterns for all 20 diseases.

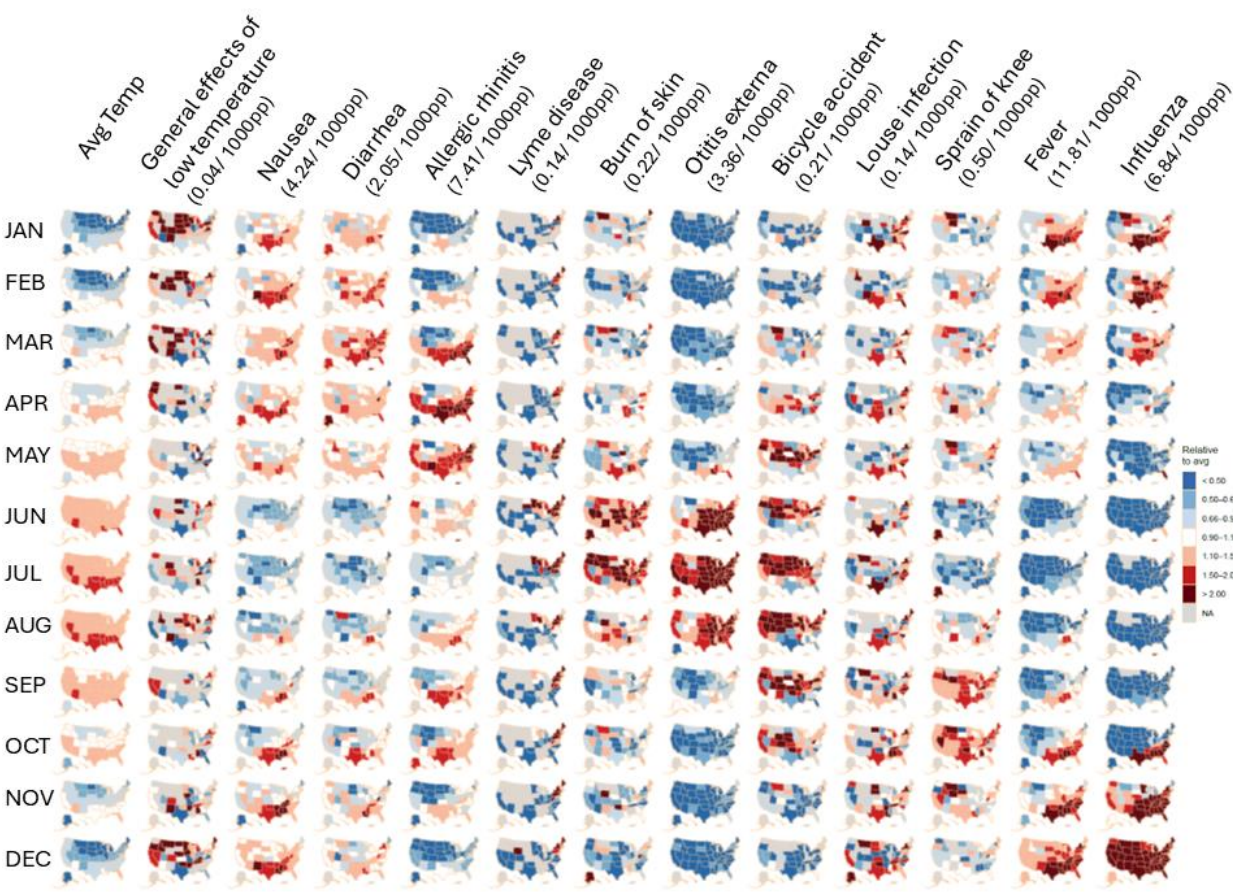


Figure 1. Incidence by month and state

Table 1. Seasonality, regionality, and correlation with temperature for all diseases

Condition	Varies by Month	Varies by Region	Correlated with Temperature
Allergic rhinitis	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = 0.49 (p = < 0.001) *
Atopic dermatitis	Yes (p = 0.008)	Yes (p = < 0.001)	rho = 0.09 (p = 0.022) *
Burn of skin	Yes (p = < 0.001)	No (p = 0.112)	rho = 0.39 (p = < 0.001) *
Cough	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.43 (p = < 0.001) *
Diarrhea	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.28 (p = < 0.001) *
Dislocation of shoulder joint	Yes (p = < 0.001)	No (p = 0.096)	rho = 0.10 (p = 0.016) *
Fever	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.41 (p = < 0.001) *
Fracture of bone	Yes (p = < 0.001)	Yes (p = 0.025)	rho = 0.37 (p = < 0.001) *
Fracture of bone of wrist region	Yes (p = < 0.001)	No (p = 0.882)	rho = 0.37 (p = < 0.001) *
Fracture of phalanx of finger	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = 0.05 (p = 0.251)
General effects of low temperature	Yes (p = < 0.001)	No (p = 0.178)	rho = -0.18 (p = < 0.001) *
Headache	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.27 (p = < 0.001) *
Influenza	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.65 (p = < 0.001) *
Louse infestation	No (p = 0.270)	Yes (p = < 0.001)	rho = 0.13 (p = 0.001) *
Lyme disease	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = 0.17 (p = < 0.001) *
Nausea	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.30 (p = < 0.001) *
Otitis externa	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = 0.73 (p = < 0.001) *
Pedal cycle accident	Yes (p = < 0.001)	Yes (p = 0.034)	rho = 0.51 (p = < 0.001) *
Sprain of ankle	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = 0.06 (p = 0.151)
Sprain of knee	Yes (p = < 0.001)	Yes (p = < 0.001)	rho = -0.01 (p = 0.845)

Conclusion

This project demonstrates that vibe coding can broaden access to rigorous OHDSI-based evidence generation while maintaining methodological transparency and reproducibility. By combining AI-assisted code generation with expert oversight and standardized HADES tooling, nontraditional researchers can execute complex, auditable observational workflows. Continued development of low-code wrappers and quality-assured AI-assisted OHDSI workflows may help expand participation and accelerate reproducible real-world evidence generation. This specific use case reinforces that calendar month and geography are important sources of variation in pediatric observational analyses.