WebLog Journal of Obstetrics and Gynecology

Research Article
Published: 08 Aug, 2026

Testing the Limits: Can Federated EHR Networks Detect Teratogenic Signals?

Fady Tawfik[1], Justyna Sienkianiec[1], Mrinalini Deverapalli[2], Agasou Rameau[1], Oyinademi Ehikioya[1], Shola Taiwo[1], Mohammed Umair[2] and Miriam Michael[2,3]*

[1]Howard University College of Medicine, Washington, DC, USA

[2]Department of Internal Medicine, Howard University, Washington, DC, USA

[3]Department of Internal Medicine, University of Maryland School of Medicine, Baltimore, USA

Abstract

Background: First-trimester topiramate exposure carries an approximately 2-fold increased risk of Major Congenital Malformations (MCMs) compared with levetiracetam, which has among the lowest MCM rates of Anti-Seizure Medications (ASMs). Federated Electronic Health Record (EHR) networks are increasingly used for drug safety research; however, their ability to replicate known teratogenic signals has not been extensively evaluated.

Objective: To examine whether the TriNetX Global Collaborative Network can replicate the established differential teratogenic risk between topiramate and levetiracetam.

Methods: This retrospective cohort study used the TriNetX Global Collaborative Network (173 healthcare organizations). Women aged 15-45 with first-trimester gestational age codes prescribed topiramate without levetiracetam (n=40,990) or levetiracetam without topiramate (n=14,429) were identified. Propensity score matching (PSM; 1:1) on 28 covariates yielded 8,112 matched pairs. Outcomes ascertained from maternal records included overall congenital anomalies (ICD-10-CM Q00-QA0), congenital heart defects (CHDs; Q20-Q28), and clubfoot (Q66). Oral clefts, the most specific topiramate-associated anomaly, could not be reliably assessed because infant diagnoses were not linkable to maternal records. Hypospadias (Q54) was included as a negative control outcome.

Results: After matching, no significant differences were detected for overall congenital anomalies (4.7% *vs*. 5.1%; odds ratio [OR] 0.92, 95% confidence interval [CI] 0.79-1.07; *p*=0.265), CHDs (1.7% *vs*. 1.4%; OR 1.21, 95% CI 0.94–1.57; p=0.144), or clubfoot (0.2% *vs*.0.3%; OR 0.61, 95% CI 0.33-1.14; *p*=0.120). Hypospadias yielded zero events in both cohorts, confirming the absence of infant-level outcome capture. Kaplan-Meier analysis for CHDs paradoxically showed a higher event rate in the levetiracetam cohort (hazard ratio 1.41, 95% CI 1.10-1.83; *p*=0.008), likely reflecting differential diagnostic ascertainment rather than a true biological signal.

Conclusions: Despite 8,112 matched pairs, this analysis did not replicate the expected topiramate-associated teratogenic patterns. These findings highlight critical limitations of unlinked maternal EHR data for pregnancy medication safety research, including absent infant outcome capture, imprecise exposure timing, and insufficient power for rare anomalies. Federated EHR platforms without validated mother-infant linkage should not be used to generate evidence on medication teratogenicity.

Keywords: Topiramate; Levetiracetam; Teratogenicity; Congenital Malformations; Electronic Health Records; Federated Data Networks; Pregnancy Safety; Pharmacoepidemiology; TriNetX; Mother-Infant Linkage

OPEN ACCESS

Correspondence:
Fady Tawfik, Howard University College of Medicine, Washington, DC, USA, E-mail: fadytwfk@gmail.com

Received Date: *26 Jun 2026*
Accepted Date: *06 Aug 2026*
Published Date: *08 Aug 2026*

Citation:
Tawfik F, Sienkianiec J, Deverapalli M, Rameau A, Ehikioya O, Taiwo S, et al. Testing the Limits: Can Federated EHR Networks Detect Teratogenic Signals?. WebLog J Obstet Gynecol. wjog.2026.h0811. https://doi.org/10.5281/zenodo.21860752

Introduction

Medication safety in pregnancy remains a critical public health concern, as pregnant women are typically excluded from clinical trials and post-marketing evidence must inform clinical decision-making [1, 2]. Topiramate is a well-characterized teratogen; data from pregnancy registries and population-based cohort studies consistently demonstrate an approximately 2-fold increased risk of MCMs compared with lamotrigine or levetiracetam [3, 4]. The North American Antiepileptic Drug (NAAED) Pregnancy Registry estimated a topiramate MCM risk of 5.1%, with a Relative Risk (RR) of 2.4 (95% CI 1.5-3.8) compared with lamotrigine [3]. A large Nordic population-based study

reported an adjusted RR of 1.81 (95% CI 1.26-2.60) for topiramate versus lamotrigine [4]. The most specific topiramate-associated anomaly is oral clefts, with a risk of 4.1 per 1,000 live births and a pooled RR of 5.27 (95% CI 2.88-9.65) across seven studies [5]. The US Food and Drug Administration (FDA) label for topiramate warns of increased risk of MCMs, including oral clefts and small-for-gestational-age births [6]. In contrast, levetiracetam has a MCM prevalence of approximately 2.6% (95% CI 1.6-4.4%) in pooled cohort data, with no evidence of increased risk compared with untreated controls [7].

Federated EHR networks, which enable distributed analyses across multiple healthcare organizations without centralizing patient-level data, are increasingly used for drug safety research [8]. These platforms offer advantages including large sample sizes, rapid accrual, and diverse populations. However, their fitness-for-purpose for pregnancy medication safety research, particularly for outcomes requiring infant-level data, has not been systematically evaluated [9, 10]. The TriNetX Global Collaborative Network is one such federated platform, encompassing 173 healthcare organizations.

The objective of this study was to determine whether the TriNetX platform could replicate the well-established differential teratogenic risk between topiramate and levetiracetam. We hypothesized that if the platform were fit-for-purpose, the known signal would be detectable. Failure to replicate this signal would suggest important limitations for teratogenicity research.

Methods

Study design and data source

This was a retrospective cohort study conducted using the TriNetX Global Collaborative Network, a federated research network providing access to de-identified EHR data (diagnoses, procedures, medications, laboratory values, genomic information) from 173 healthcare organizations. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [11].

Study population

Two mutually exclusive cohorts were defined:

- Cohort 1 (Levetiracetam): Women aged 15-45 years with a first-trimester gestational age code (ICD-10-CM Z3A.0 [unspecified or <10 weeks], Z3A.10-Z3A.13 [10-13 weeks]) and a prescription for levetiracetam (RxNorm: 114477), without any prescription for topiramate (RxNorm: 38404). This cohort included 14,429 patients from 73 responding healthcare organizations.

- Cohort 2 (Topiramate): Women aged 15-45 years with a first-trimester gestational age code (same codes as above) and a prescription for topiramate (RxNorm: 38404), without any prescription for levetiracetam (RxNorm: 114477). This cohort included 40,990 patients from 70 responding healthcare organizations.

Index event and time window

The index event was defined as the first co-occurrence of the gestational age code and the respective medication prescription. Outcomes were assessed starting 1 day after the index event, with no specified end date, capturing all subsequent events.

Propensity score matching

One-to-one PSM was performed on 28 covariates across three domains: demographics (age at index, sex, race [7 categories], ethnicity [3 categories]); diagnoses (type 1 diabetes, type 2 diabetes, alcohol-related disorders, multiple gestation, epilepsy and recurrent seizures, migraine, overweight and obesity, hypertensive diseases, depressive episode, major depressive disorder [recurrent], other anxiety disorders, bipolar disorder, nicotine dependence); and medication (folic acid use). PSM yielded 8,112 matched pairs.

Outcomes

All outcomes were included from maternal EHR records. Patients with the outcome documented prior to the time window were excluded from the respective analysis. This includes overall congenital anomalies, Congenital Heart Defects (CHDs), clubfoot, and hypospadias (negative control). Hypospadias was included as a pre-specified negative control outcome. As a male-only diagnosis that would never appear in a maternal record, zero events in both cohorts would confirm the absence of infant-level outcome capture within the platform [12, 13]. Oral clefts, the most specific and well-established topiramate-associated anomaly, could not be assessed because infant diagnoses were not linkable to maternal records within this platform [5, 14].

Statistical analysis

Measures of association included risk difference, risk ratio, and odds ratio with 95% CIs. Kaplan-Meier survival analysis with log-rank testing and Cox proportional hazards regression were performed for time-to-event analyses. Covariate balance after PSM was assessed using standardized differences, with values <0.1 considered well-balanced.

Results

Baseline characteristics

Before matching, the two cohorts differed substantially in clinical indication profiles. The levetiracetam cohort had a significantly higher prevalence of epilepsy (47.8% vs. 3.6%; standardized difference [Std Diff] 1.175), while the topiramate cohort had higher rates of migraine (55.5% *vs*. 18.4%; Std Diff 0.832) and overweight/obesity (52.1% *vs*. 23.2%; Std Diff 0.625). The levetiracetam cohort was younger (mean age 28.5 *vs*. 30.4 years; Std Diff 0.317), had a higher proportion of Black or African American patients (30.0% *vs*. 20.7%; Std Diff 0.214), and greater folic acid use (35.2% *vs*. 23.0%; Std Diff 0.271).

After PSM, 8,112 matched pairs were obtained. All standardized differences were reduced to <0.05, indicating adequate balance. Residual statistically significant but clinically small differences remained for migraine (23.4% *vs*. 22.0%; p=0.030) and obesity (28.5% *vs*. 27.0%; p=0.034). Epilepsy prevalence was balanced at 16.4% *vs*. 15.4% (p=0.059; Std Diff 0.030). Mean follow-up was 972 days (SD 882) in the levetiracetam cohort and 1,187 days (SD 939) in the topiramate cohort.

Primary outcomes

Results are summarized in Table 1.

No significant differences were detected for any outcome. The point estimate for CHDs paradoxically favored higher risk in the levetiracetam cohort (OR 1.21), opposite to the expected direction based on established literature [3, 4]. The clubfoot point estimate (OR 0.61) was directionally consistent with a protective effect of levetiracetam, but the confidence interval was wide and included the null. Additional results are shown in Figures 1 and 2.

Forest plot of matched odds ratios with 95% confidence intervals

Table 1: Outcomes after propensity score matching (8, 112 Matched Pairs).

Outcome	Levetiracetam n/N (%)	Topiramate n/N (%)	OR (95% CI)	*p*-value
Overall congenital anomalies	332/7, 097 (4.7%)	367/7, 225 (5.1%)	0.92 (0.79-1.07)	0.265
Congenital heart defects	129/7, 738 (1.7%)	109/7, 894 (1.4%)	1.21 (0.94-1.57)	0.144
Clubfoot	16/8,085 (0.2%)	26/8,056 (0.3%)	0.61 (0.33-1.14)	0.120
Hypospadias (negative control)	0/8,112 (0%)	0/8,112 (0%)	-	-

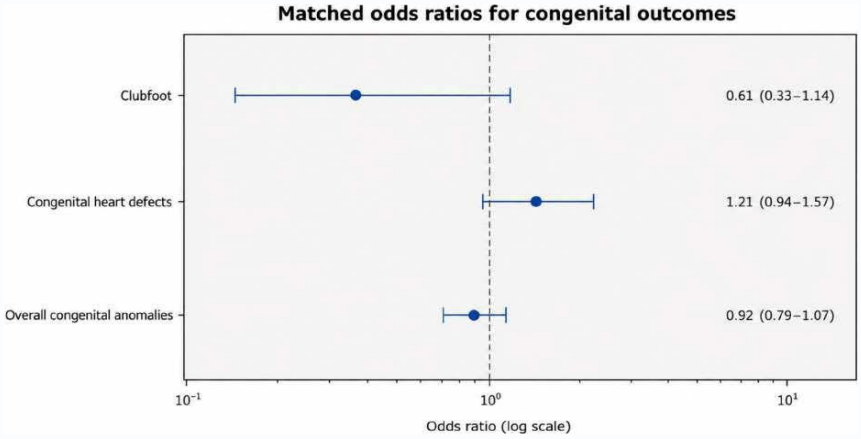

Figure 1: Matched odds ratios for congenital outcomes after first-trimester exposure.

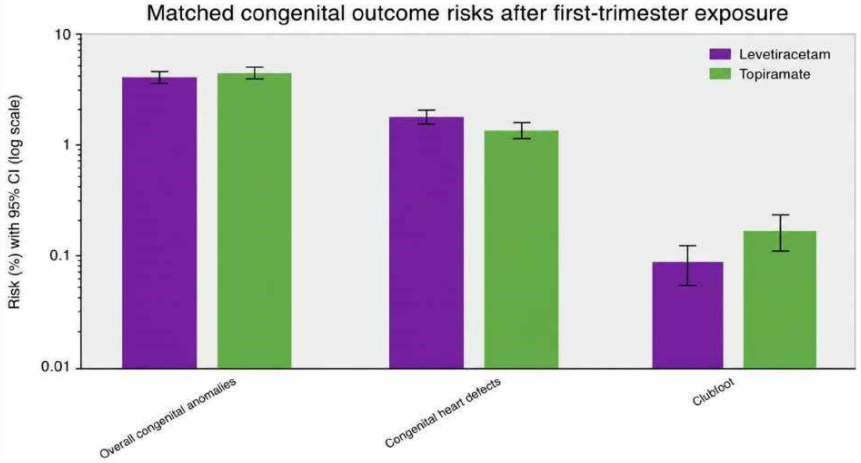

Figure 2: Matched congenital outcome risks after first-trimester exposure.

for overall congenital anomalies, congenital heart defects, and clubfoot after first-trimester exposure. The dashed line represents the null value of 1.0; all confidence intervals crossed 1.0.

Bar graph comparing absolute risk percentages with 95% confidence intervals for congenital outcomes in the levetiracetam and topiramate groups. Risks were similar across groups, with overlapping confidence intervals.

Kaplan-Meier survival analysis

Survival analysis for overall congenital anomalies showed no significant difference (log-rank p=0.274; hazard ratio [HR] 1.09, 95% CI 0.94-1.26). For CHDs, the Kaplan-Meier analysis paradoxically demonstrated a statistically significant higher event rate in the levetiracetam cohort (log-rank p=0.008; HR 1.41, 95% CI 1.10-1.83; proportionality test p=0.036). This finding is inconsistent with the known safety profile of levetiracetam and likely reflects differential diagnostic ascertainment: 374 levetiracetam patients (4.6%) versus 218 topiramate patients (2.7%) were excluded for pre-existing CHD diagnoses, suggesting more complete cardiac diagnostic coding in the levetiracetam cohort, possibly related to the higher baseline epilepsy prevalence and associated clinical surveillance. Clubfoot survival analysis showed no significant difference (log-rank p=0.371; HR 0.75, 95% CI 0.40-1.41).

Negative control outcome

Hypospadias yielded zero events in both cohorts (0/8,112 in each). As a male-only congenital anomaly that would be diagnosed exclusively in the infant's record, this result confirms that the TriNetX platform does not capture infant-level diagnoses through maternal records. This finding validates the hypothesis that the platform's outcome ascertainment is limited to diagnoses coded in the maternal chart.

Pre-Existing outcome exclusions

The number of patients excluded for pre-existing congenital anomaly diagnoses was notably high: 1,015 (12.5%) in the levetiracetam cohort and 887 (10.9%) in the topiramate cohort. These likely represent maternal congenital diagnoses (e.g., congenital heart disease in the mother herself) rather than infant outcomes, further confirming that Q-code outcomes in this platform are ascertained from maternal records.

Discussion

This study demonstrates that the TriNetX Global Collaborative Network, a large federated EHR platform encompassing 173 healthcare organizations, was unable to replicate the well-established differential teratogenic risk between topiramate and levetiracetam. Despite a matched sample of 8,112 pairs, no significant differences were detected for overall congenital anomalies, CHDs, or clubfoot. These null findings, in the context of robust evidence from pregnancy registries and a male-only diagnosis that cannot appear in a maternal record, provide a useful negative control confirming this gap. Oral clefts, the most specific and well-replicated topiramate-associated anomaly (pooled RR 5.27 across seven studies; OR 6.8 in a French population-based study), could not be assessed at all [5, 14]. Recent reviews have emphasized that accurate mother-infant linkage is essential for pregnancy pharmacoepidemiology, yet linkage methods remain heterogeneous and are absent from many federated platforms [9, 10]. Without validated linkage algorithms, whether through shared insurance identifiers, tokenization, or probabilistic matching, infant outcomes cannot be reliably ascertained [10].

Paradoxical CHD findings

The significant Kaplan-Meier finding for congenital heart defects (HR 1.41, p=0.008) was opposite to the expected direction, with a higher observed risk in the levetiracetam group. This finding is more likely explained by differences in diagnostic ascertainment than by a true biological effect. The levetiracetam cohort had nearly twice as many patients excluded for pre-existing CHD diagnoses (374 *vs*. 218), suggesting more thorough cardiac diagnostic coding, possibly driven by the higher baseline epilepsy prevalence and associated clinical surveillance in that group. This paradoxical finding underscores the risk of generating artifactual signals when using unlinked maternal EHR data for congenital outcome research.

Confounding by indication

Before matching, the two treatment groups differed substantially in their baseline characteristics: the levetiracetam cohort was predominantly composed of women with epilepsy (47.8%), while the topiramate cohort was predominantly prescribed for migraine (55.5%) and obesity (52.1%). Although PSM reduced standardized differences to <0.05 for all covariates, the matched cohort (with epilepsy prevalence of ~16% in both groups) represents a selected subgroup that is not representative of either typical levetiracetam or topiramate users. This substantial sample reduction (from 52,503 to 16,224 patients) raises concerns about generalizability. Furthermore, the inability to match on dose, a critical determinant of topiramate teratogenicity, with adjusted RRs for oral clefts of 1.64 at doses ≤100 mg/day versus 5.16 at doses >100 mg/day, represents an important unmeasured confounder [5]. Confounding by indication, while present, would be expected to bias results toward the null for topiramate-associated teratogenicity, given the lower doses typically used for migraine and obesity indications. The complete absence of a detectable signal, rather than merely an attenuated one, combined with zero hypospadias events and paradoxical CHD findings, points to a more fundamental problem with outcome ascertainment rather than residual confounding.

Exposure timing uncertainty

The gestational age codes used to define first-trimester exposure (Z3A.0 through Z3A.13) include Z3A.0 ("unspecified or less than 10 weeks"), which is a broad category. The study design required only that a medication prescription and a first-trimester gestational age code co-existed in the record, without confirming that the medication was actively taken during the etiologically relevant window for organogenesis (weeks 3-8 post-conception). This imprecision in exposure timing is a recognized limitation of EHR-based pregnancy research [2, 15]. Primary data collection studies and claims-based studies with validated pregnancy algorithms can better characterize the timing, dose, and duration of medication exposure [15, 16].

Power considerations

Although the matched cohort included 8,112 pairs, it remained underpowered to detect rare congenital anomalies. Clubfoot, with only 16 and 26 events in the respective cohorts, yielded a wide confidence interval (OR 0.61, 95% CI 0.33-1.14). Previous studies have shown that evaluating specific malformation subtypes requires very large pregnancy cohorts. For example, the Medicaid study that identified the association between topiramate and oral clefts included more than 1.3 million pregnancies [5]. The rarity of specific congenital anomalies, combined with the outcome ascertainment limitations described above, renders this platform inadequate for detecting even well-established teratogenic signals.

Implications for Federated EHR research

These findings have important implications for the growing use of federated EHR networks in drug safety research. While such platforms have demonstrated value for many pharmacoepidemiologic questions, our findings suggest that these platforms should not be used for studies of congenital outcomes unless validated mother-infant linkage is available [8, 17]. This aligns with recent commentary emphasizing that "big data" is not always better for birth defects research, where specificity of exposure timing and outcome ascertainment are paramount [15]. Platforms with validated linkage -such as Nordic national registries with universal personal identifiers, US Medicaid data with mother-infant linkage through shared enrolment, or the FDA Sentinel System with established linkage algorithms - remain the appropriate data sources for teratogenicity research [4, 5, 16].

Strengths and Limitations

Strengths of this study include the large initial sample size, the use of a well-characterized positive control (the topiramate-levetiracetam teratogenic differential), the inclusion of a negative control outcome (hypospadias), and the propensity score matching across 28 demographic and clinical variables. The study was designed as a validation exercise, and the null findings are informative precisely because the expected signal is well-established. Limitations include the absence of mother-infant linkage, inability to assess oral clefts, imprecise exposure timing, lack of dose information, inability to distinguish monotherapy from sequential therapy, and the inability to confirm that medication exposure occurred during the relevant gestational window. Additionally, the platform does not distinguish between live births, stillbirths, and elective terminations, which may

differentially affect outcome ascertainment.

Conclusions

Despite a matched cohort of 8,112 pairs drawn from 173 healthcare organizations, the TriNetX Global Collaborative Network failed to replicate the well-established teratogenic signal for topiramate relative to levetiracetam. The absence of mother-infant record linkage, confirmed by zero hypospadias events in both cohorts, represents a fundamental barrier to infant outcome ascertainment. Paradoxical findings for CHDs further demonstrate the potential for artifactual signals in unlinked maternal data. Federated EHR platforms without validated mother–infant linkage should not be used as standalone data sources for studies of congenital outcomes Future efforts should prioritize the development and validation of mother-infant linkage algorithms within federated networks to enable reliable teratogenicity surveillance.

Statements and Declarations

Ethical Publication and Informed Consent Statement: This retrospective analysis used only de-identified data from the TriNetX Global Collaborative Network. Since no identifiable private information was accessed, this work did not constitute human subjects research and institutional review board approval and informed consent were not required.

Declaration of conflicting interest: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article

Funding statement: The authors received no financial support for the research, authorship, and/or publication of this article.

Data availability statement: The data used in this study were obtained from the TriNetX database, a global federated research network. Access to the database is available to subscribing institutions, and data are not publicly available due to licensing agreements. However, aggregated results generated during this study are available from the corresponding author upon reasonable request.

Author contributions: All authors contributed to study conception and design, data curation within the TriNetX environment, interpretation of results, and drafting or critical revision of the manuscript. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Acknowledgments: None.

Prior presentations: Poster Presentation, Society for Birth Defects Research and Prevention Annual Meeting, Washington, DC. June 2026.

References

1. Wood ME, Andrade SE, Toh S. Safe Expectations: Current State and Future Directions for Medication Safety in Pregnancy Research. Clin Ther. 2019; 41(12): 2467-2476.
2. Huybrechts KF, Bateman BT, Hernández-Díaz S. Modern Evidence Generation on Medication Effectiveness and Safety During Pregnancy: Study Design Considerations. Clin Pharmacol Ther. 2025; 117(4): 895-909.
3. Hernandez-Diaz S, Quinn M, Conant S, Lyons A, Paik H, Ward J, et al. Use of Antiseizure Medications Early in Pregnancy and the Risk of Major Malformations in the Newborn. Neurology. 2025; 105(3): e213786.
4. Cohen JM, Alvestad S, Cesta CE, Bjørk M, Leinonen MK, Nørgaard M, et al. Comparative Safety of Antiseizure Medication Monotherapy for Major Malformations. Ann Neurol. 2023; 93(3): 551-562.
5. Hernandez-Diaz S, Huybrechts KF, Desai RJ, Cohen JM, Mogun H, Pennell PB, et al. Topiramate Use Early in Pregnancy and the Risk of Oral Clefts: A Pregnancy Cohort Study. Neurology. 2018; 90(4): e342-e351.
6. Topamax. Food and Drug Administration. 2026.
7. Bromley R, Adab N, Bluett-Duncan M, Clayton-Smith J, Christensen J, Edwards K, et al. Monotherapy Treatment of Epilepsy in Pregnancy: Congenital Malformation Outcomes in the Child. The Cochrane Database of Systematic Reviews. 2023; 8: CD010224.
8. Norén GN, Donegan K, Muñoz MA, Alshammari TM, Pratt N, Candore G, et al. Recommendations to Enable Broader Use of Real-World Evidence to Inform Decision-Making Throughout Pharmacovigilance Signal Management. Pharmacoepidemiol Drug Safety. 2025; 34(10): e70231.
9. Jacobson MH, Krumme AA, Gifkins DM, Margulis AV, Weaver J, Davis KJ. Review of Mother–Infant Linkage Algorithms in Health Care Data Sources. Pharmacoepidemiol Drug Saf. 2025; 34(8): e70189.
10. Kaplan S, Ayele HT, Bennett D, Grandi SM, Schachterle SE, Simeone JC, et al. Unlocking Insights From Mother-Infant Linked Data in Pharmacoepidemiology: Opportunities, Challenges, and Future Directions. Curr Med Res Opin. 2025; 41(9): 1611-1618.
11. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. Ann Intern Med. 2007; 147(8): 573-577.
12. Levintow SN, Nielson CM, Hernandez RK, Breskin A, Pritchard D, Lash TL, et al. Pragmatic considerations for negative control outcome studies to guide non-randomized comparative analyses: A narrative review. Pharmacoepidemiol Drug Saf. 2023; 32(6): 599-606.
13. Muntner P, Hernandez RK, Kent ST, Browning JE, Gilbertson DT, Hurwitz KE, et al. Staging and clean room: Constructs designed to facilitate transparency and reduce bias in comparative analyses of real-world data. Pharmacoepidemiol Drug Saf. 2024; 33(3): e5770.
14. Blotière PO, Raguideau F, Weill A, Elefant E, Perthus I, Goulet V, et al. Risks of 23 Specific Malformations Associated With Prenatal Exposure to 10 Antiepileptic Drugs. Neurology. 2019; 93(2): e167-e180.
15. Ailes EC, Werler MM, Howley MM, Jenkins MM, Reefhuis J. Real-World Data Are Not Always Big Data: The Case for Primary Data Collection on Medication Use in Pregnancy in the Context of Birth Defects Research. Am J Epidemiol. 2024; 193(9): 1211-1214.
16. Ailes EC, Zhu W, Clark EA, Huang YA, Lampe MA, Kourtis AP, et al. Identification of Pregnancies and Their Outcomes in Healthcare Claims Data, 2008-2019: An Algorithm. PloS One. 2023; 18(4): e0284893.
17. Coste A, Wong A, Bokern M, Bate A, Douglas IJ. Methods for drug safety signal detection using routinely collected observational electronic health care data: A systematic review. Pharmacoepidemiol Drug Saf. 2023; 32(1): 28-43.