Published in final edited form as:

Am J Epidemiol. 2024 September 03; 193(9): 1211–1214. doi:10.1093/aje/kwae060.

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

Elizabeth C. Ailes[*,1], **Martha M. Werler**[2], **Meredith M. Howley**[3], **Mary M. Jenkins**[1], **Jennita Reefhuis**[1]

[1]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

[2]Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, USA

[3]Birth Defects Registry, New York State Department of Health, Albany, New York, USA.

Abstract

Many examples of the use of real-world data in the area of pharmacoepidemiology include “big data,” such as insurance claims, medical records, or hospital discharge databases. However, “big” is not always better, particularly when studying outcomes with narrow windows of etiologic relevance. Birth defects are such an outcome, for which specificity of exposure timing is critical. Studies with primary data collection can be designed to query details about the timing of medication use, as well as type, dose, frequency, duration, and indication, that can better characterize the “real world.” Because birth defects are rare, etiologic studies are typically case-control in design, like the National Birth Defects Prevention Study, Birth Defects Study to Evaluate Pregnancy Exposures, and Slone Birth Defects Study. Recall bias can be a concern, but the ability to collect detailed information about both prescription and over-the-counter medication use and other exposures such as diet, family history, and sociodemographic factors is a distinct advantage over claims and medical record data sources. Case-control studies with primary data collection are essential to advancing the pharmacoepidemiology of birth defects.

Keywords

medications; birth defects; case-control studies; administrative data

Almost 90% of pregnant women take a medication, yet there is a paucity of data on the safety of many medications during pregnancy.[1,2] Among their 15 recommendations to

[*] Corresponding author: Elizabeth C. Ailes, Mailstop S106-3, Centers for Disease Control and Prevention, 4770 Buford Highway, Atlanta, GA 30341-3717 (EAiles@CDC.Gov).

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the New York State Department of Health.

Conflict of interest

M.M.W. serves as a diagnostic adjudicator for Novartis pregnancy registries. The other authors declare no conflicts.

Congress, the Task Force on Research Specific to Pregnant Women and Lactating Women recognized the need to "increase the quantity, quality, and timeliness of research on safety and efficacy of therapeutic products used by pregnant women and lactating women."[3] Many researchers have capitalized on the wealth of "big data" (ie, large secondary data sources, such as insurance claims or hospital discharge databases initially collected for purposes other than scientific research) to address this challenge.[4–8] In fact, many examples of "real-world" pharmacoepidemiology data use such data sources across a variety of subject domains. (Throughout this commentary, we use the term *real world* to indicate the actual experiences [ie, medication dispensation or use patterns] of women outside of a controlled clinical trial setting.) Despite their strengths, including prospective data collection, these analyses often rely on claims as a proxy for prescription medication exposure and billing codes to ascertain conditions of interest. Additionally, information on the use of over-the-counter (OTC) medications or supplements is limited to those few, if any, covered by insurance, which would still be a severe underestimate of actual use. Of relevance to birth defects research in particular are pharmacoepidemiologic studies based on insurance claims and medical records data sources, which often depend on algorithms to identify the pregnancies, outcomes, exposure, and, importantly, their timing. Another real-world approach to studying the safety of medication use in pregnancy complementary to analyses of administrative data are case-control studies, such as the population-based National Birth Defects Prevention Study (NBDPS), Birth Defects Study to Evaluate Pregnancy Exposures (BD-STEPS), and Slone Birth Defects Study (BDS).

Conducting pharmacoepidemiologic studies of birth defects is challenging for several reasons. It is critical to have accurate information not only on the timing of medication exposure but also on the beginning of pregnancy (conception). This is because the important window of exposure for many major structural birth defects is early in the first trimester of pregnancy, the period of organogenesis.[9] In many real-world secondary data sources, data on the timing of pregnancy are not available in any standardized fashion. Because health insurance claims data sources are for billing purposes, documentation of gestational timing is not typically essential. When there is no record of the date of conception, or even the identification of a pregnancy, its outcome, and the critical time periods in pregnancy (eg, first trimester), epidemiologists have to rely on algorithms based on diagnosis, procedure, and/or diagnosis-related group codes.[10–13] When algorithms have been compared with birth certificates or medical records, with more detailed pregnancy information, agreement between sources was not perfect, with variation by type of algorithm and gestational length.[14,15] Because of the rapid embryologic development, misestimation of the beginning of pregnancy by even a few weeks could result in biased estimates of the association between a medication and a birth defect, with bias more likely for algorithms that do not account for preterm births and/or focus on medications used episodically.[15] An advantage of real-world studies that use primary data collection is the ability to gather detailed information on the timing of medication exposures paired with clinical data on pregnancy dates. For example, NBDPS, BD-STEPS, and BDS ask women about their expected due date, which serves as an anchor for determining pregnancy timing. To determine whether a medication exposure occurred in the etiologically relevant window, determination of

gestational timing is important and gestational timing has been shown to be accurately reported retrospectively within 6 months of delivery.[16]

Although pharmacy records provide evidence of dispensed medications, there is no guarantee that a medication is taken at all or for the entire period prescribed, or that medication users continued use at the prescribed dose. By interacting with women, studies that use primary data collection may glean more accurate medication exposure information. Administrative data may underestimate use if prescription medications are paid for out of pocket or ascertained through pharmacies not affiliated with the underlying data source. In addition, medication sharing is common; in a 2008 survey, more than one-third of reproductive-aged women reported sharing or borrowing prescriptions.[17] Conversely, administrative data may overestimate medication use because of nonadherence if women do not fill all prescribed medications or do not take all dispensed medications. Given concerns about potential teratogenicity or other negative impacts during pregnancy, women may stop taking a prescribed medication after they find out they are pregnant or they may take a dose lower than prescribed.[18–20]

It is also important to understand the full spectrum of medication exposures. More than half of women report using acetaminophen in the first trimester, and OTC medications (e.g., analgesics, cough and cold medicines, gastrointestinal medications) are some of the most common medications used periconceptionally.[1,21] Given their frequency of use, it is vital to monitor the safety of OTC medications and to consider their concomitant use with other OTC and prescription medications.[22,23] For instance, Interrante et al[23] found associations with specific birth defects varied between women periconceptionally exposed to nonsteroidal anti-inflammatory drugs alone, opioids alone, and nonsteroidal anti-inflammatory drugs and opioids combined, compared with those who took acetaminophen. Herbal supplements are also worthy of consideration. Almost 6% of women reported taking an herbal supplement during the first trimester of pregnancy, and use increased slightly from 6.1% to 7.6% upon recognition of pregnancy.[24] The Dietary Supplement Health and Education Act of 1994 identified dietary supplements, including herbal supplements, as food.[25] Given the different regulatory pathways in the United States for prescription medications and herbal supplements, less information is available on the frequency and safety of use of herbal supplements during pregnancy. Without primary data, the safety of OTC and herbal supplement use, their concomitant use with other medications, and their confounding effects would be poorly understood.

Using primary data collection methods, information can also be obtained about the “why” of medication use—a vital consideration in pharmacoepidemiologic studies of birth defects when there is concern about the potential for confounding by indication. Indication for use of a medication can be difficult to ascertain in big data sources because prescriptions are generally not easily linked to a health care encounter and indication for use might not be clearly captured. For instance, in an analysis of Tennessee Medicaid data, Cooper et al[26] were only able to identify an indication for 173 of 391 women (44%) who filled a medication for a medication contraindicated during pregnancy. By querying women about acute and chronic conditions as well as medication use, for instance, NBDPS, BD-STEPS, and BDS have attempted to account for underlying conditions; for instance, assessing

the associations between antibiotics and birth defects only among women with reported urinary tract infections.[27] Still, these data are useful but certainly far from representing the complexities of underlying health status for those exposed to the disease or exposed to a given medication.

The comprehensive primary data collected in case-control studies also allow for collection of important information on potential confounders and effect modifiers such as diet, lifestyle factors, occupation, environmental exposures, sociodemographic factors, and genetics. Periconceptional smoking, alcohol consumption, and low intake of folic acid are strongly associated with a number of birth defects.[28–30] Information on these exposures is rarely available in administrative databases and, even when present, may be subject to misclassification if the information is not directly reported by study participants.[31] As a result of the availability of comprehensive risk factor information, for example, researchers identified that diet modified the association between nitrosatable drugs and preterm birth,[32] and folic acid intake modified the association between nonsteroidal anti-inflammatory drugs and spina bifida.[33] Such analyses are only possible with comprehensive primary data collection.

Beyond those noted above, advantages of case-control study design in assessing risk factors for birth defects include population-based ascertainment of participants, inclusion of pregnancy losses, accruing sufficient sample sizes of specific birth defects, and applying stringent and consistent case classification schema. However, these topics have been discussed extensively in the literature[34,35] and are outside the scope of this commentary.

Still, to better characterize potential differences in medication information ascertained through maternal interview as compared with medical, insurance, or other sources, validation studies are necessary. Several BD-STEPS sites are conducting pilot medication-validation projects, which will allow quantitative bias analyses in the future.[36] However, there is no true gold standard data source, because both dispensation data and interview data have limitations. Nevertheless, medication-validation studies provide useful data points for sensitivity analyses to establish a plausible range of exposure misclassification. Howley et al[37] compared maternally reported prescription information with information in medical records for medication used in early pregnancy among 184 BD-STEPS participants from New York. No significant or meaningful differences were noted in the concordance of prescription medications between the 2 data sources (maternally reported and medical records) by case and control status, supporting an argument that any resulting bias is more likely to be nondifferential in medication studies.[37] To be able to say a medication is causally associated with a birth defect requires consistency, 1 of the Bradford-Hill criteria of causality,[38] which, by definition, requires diverse populations and, ideally, diverse approaches—underlining the value of including both big data mining and primary data collection. Case-control studies are resource- and time-intensive, whereas big data are more readily available, but full access to electronic health data may come at a substantial cost. Although not limited to the examination of medication exposures, case-control studies can be useful in further examining teratogenic signals from big data analyses. For instance, Huybrechts et al[39] found an elevated association (relative risk = 3.70; 95% CI, 1.55–8.82) between hydroxychloroquine and oral clefts, based on a total of 25 exposed pregnancies, in

a large cohort study using MarketScan and Medicaid databases. In their subsequent analysis of NBDPS and BDS data, Howley et al[40] also noted an increase in clefts, based on a total of 6 exposed pregnancies, but notably did not find a discernable pattern (identifying 2 each of cleft palate alone, cleft lip alone, and cleft lip with or without cleft palate). Because of the specificity of outcome ascertainment, case-control studies may also detect signals that are not observed in big data when groups of specific birth defects are combined. Tinker et al[41] nicely described this phenomenon in a comparison of the Muanda et al cohort study[42] and NBDPS findings reported by Crider et al[43] Briefly, a possible explanation for the null finding between nitrofurantoin and cardiac defects in the cohort study could be lack of outcome specificity.[41,42] Although the NBDPS analysis found a similar estimate for nitrofurantoin and any cardiac malformation, because of the ability to look at specific defect types, an elevated association for 1 particular heart defect (hypoplastic left heart syndrome) was observed.[43]

Because pregnant women are often not included in the pre-marketing research, postmarketing surveillance and research of medication use during pregnancy will have to provide the data for the important question of whether medications are of concern during pregnancy. Along with big data, case-control studies play a pivotal role in providing the real-world evidence of medication exposures during pregnancy and the pharmacoepidemiology of birth defects. Future analyses will continue to monitor the safety of medications during pregnancy and build upon the comprehensive data collection of these studies to consider the likely multifactorial nature of birth defects.

Acknowledgments

We thank Dr. Suzanne Gilboa for her feedback on early drafts of this manuscript.

Funding

This project was supported through the Centers for Disease Control and Prevention (CDC) cooperative agreements under PA 96043, PA 02081, FOA DD09-001, FOA DD13-003, NOFO DD18-001, and NOFO DD23-001 to the Centers for Birth Defects Research and Prevention participating in the NBDPS and/or the BD-STEPS.

Data availability

The study questionnaires and process for accessing the data used in this study are described at https://www.cdc.gov/birth-defects/php/bd-steps-nbdps-data/index.html.

References

1. Thorpe PG, Gilboa SM, Hernandez-Diaz S, et al. Medications in the first trimester of pregnancy: most common exposures and critical gaps in understanding fetal risk. Pharmacoepidemiol Drug Saf. 2013;22(9):1013–1018. 10.1002/pds.3495 [PubMed: 23893932]
2. Mitchell AA, Gilboa SM, Werler MM, et al. Medication use during pregnancy, with particular focus on prescription drugs: 1976–2008. Am J Obstet Gynecol. 2011;205(1):51-e1–51 e8. 10.1016/j.ajog.2011.02.029
3. Task Force on Research Specific to Pregnant Women and Lactating Women. Report to Secretary, Health and Human Services, Congress. 2018. Accessed June 7, 2024. https://www.nichd.nih.gov/sites/default/files/2018-09/PRGLAC_Report.pdf

4. Bateman BT, Hernandez-Diaz S, Straub L, et al. Association of first trimester prescription opioid use with congenital malformations in the offspring: population based cohort study. BMJ. 2021;372:n102. 10.1136/bmj.n102 [PubMed: 33568363]

5. Challa AP, Niu X, Garrison EA, et al. Medication history-wide association studies for pharmacovigilance of pregnant patients. Commun Med (Lond). 2022;2:115. 10.1038/s43856-022-00181-w [PubMed: 36124058]

6. Hernandez-Diaz S, Huybrechts KF, Desai RJ, et al. Topiramate use early in pregnancy and the risk of oral clefts: a pregnancy cohort study. Neurology. 2018;90(4):e342–e351. 10.1212/WNL.0000000000004857 [PubMed: 29282333]

7. Huybrechts KF, Kulldorff M, Hernandez-Diaz S, et al. Active surveillance of the safety of medications used during pregnancy. Am J Epidemiol. 2021;190(6):1159–1168. 10.1093/aje/kwaa288 [PubMed: 33423046]

8. Garne E, Vinkel Hansen A, Morris J, et al. Risk of congenital anomalies after exposure to asthma medication in the first trimester of pregnancy - a cohort linkage study. BJOG. 2016; 123(10):1609–1618. 10.1111/1471-0528.14026 [PubMed: 27172856]

9. Carlson BM, ed. Developmental disorders: causes, mechanisms, and patterns. In: Human Embryology and Developmental Biology. 5th ed. Elsevier; 2014:136–153.

10. Bertoia ML, Phiri K, Clifford CR, et al. Identification of pregnancies and infants within a US commercial healthcare administrative claims database. Pharmacoepidemiol Drug Saf. 2022;31(8):863–874. 10.1002/pds.5483 [PubMed: 35622900]

11. Palmsten K, Huybrechts KF, Mogun H, et al. Harnessing the Medicaid analytic eXtract (MAX) to evaluate medications in pregnancy: design considerations. PloS One. 2013;8(6):e67405. 10.1371/journal.pone.0067405

12. Ailes EC, Zhu W, Clark EA, et al. Identification of pregnancies and their outcomes in healthcare claims data, 2008–2019: an algorithm. PloS One. 2023;18(4):e0284893. 10.1371/journal.pone.0284893

13. Margulis AV, Palmsten K, Andrade SE, et al. Beginning and duration of pregnancy in automated health care databases: review of estimation methods and validation results. Pharmacoepidemiol Drug Saf. 2015;24(4):335–342. 10.1002/pds.3743 [PubMed: 25627986]

14. Toh S, Mitchell AA, Werler MM, et al. Sensitivity and specificity of computerized algorithms to classify gestational periods in the absence of information on date of conception. Am J Epidemiol. 2008;167(6):633–640. 10.1093/aje/kwm367 [PubMed: 18194999]

15. Zhu Y, Hampp C, Wang X, et al. Validation of algorithms to estimate gestational age at birth in the Medicaid analytic eXtract-quantifying the misclassification of maternal drug exposure during pregnancy. Pharmacoepidemiol Drug Saf. 2020;29(11):1414–1422. 10.1002/pds.5126 [PubMed: 32909348]

16. Petersen JM, Mitchell AA, Van Bennekom C, et al. Validity of maternal recall of gestational age and weight at birth: comparison of structured interview and medical records. Pharmacoepidemiol Drug Saf. 2019;28(2):269–273. 10.1002/pds.4699 [PubMed: 30488571]

17. Petersen EE, Rasmussen SA, Daniel KL, et al. Prescription medication borrowing and sharing among women of reproductive age. J Womens Health (Larchmt). 2008;17(7):1073–1080. 10.1089/jwh.2007.0769 [PubMed: 18774892]

18. Einarson A, Selby P, Koren G. Abrupt discontinuation of psychotropic drugs during pregnancy: fear of teratogenic risk and impact of counselling. J Psychiatry Neurosci. 2001;26(1):44–48. [PubMed: 11212593]

19. Zetstra-van der Woude PA, Vroegop JS, Bos HJ, et al. A population analysis of prescriptions for asthma medications during pregnancy. J Allergy Clin Immunol. 2013;131(3):711–717. 10.1016/j.jaci.2012.08.027 [PubMed: 23063582]

20. Olesen C, Sondergaard C, Thrane N, et al. Do pregnant women report use of dispensed medications? Epidemiology. 2001;12(5):497–501. 10.1097/00001648-200109000-00006 [PubMed: 11505166]

21. Werler MM, Mitchell AA, Hernandez-Diaz S, et al. . Use of over-the-counter medications during pregnancy.[see comment]. American Journal of Obstetrics & Gynecology. 2005;193(3 Pt 1):771–777. 10.1016/j.ajog.2005.02.100 [PubMed: 16150273]

22. Cao Y, Rhoads A, Burns T, et al. Maternal use of cough medications during early pregnancy and selected birth defects: a US multisite, case-control study. BMJ Open. 2021;11(12):e053604. 10.1136/bmjopen-2021-053604

23. Interrante JD, Ailes EC, Lind JN, et al. Risk comparison for prenatal use of analgesics and selected birth defects, National Birth Defects Prevention Study 1997–2011. Ann Epidemiol. 2017; 27(10):645–653.e2. 10.1016/j.annepidem.2017.09.003 [PubMed: 28993061]

24. Broussard CS, Louik C, Honein MA, et al. Herbal use before and during pregnancy. Am J Obstet Gynecol. 2010;202(5):443 e1–443 e6. 10.1016/j.ajog.2009.10.865

25. Food US and Drug Administration. Milestones in U.S. Food and Drug Law. 2023 [updated January 30, 2023]. Accessed September 12, 2023. https://www.fda.gov/about-fda/fda-history/milestones-us-food-and-drug-law

26. Cooper WO, Hickson GB, Ray WA. Prescriptions for contraindicated category X drugs in pregnancy among women enrolled in TennCare. Paediatr Perinat Epidemiol. 2004;18(2):106–111. 10.1111/j.1365-3016.2003.00542.x [PubMed: 14996249]

27. Ailes EC, Gilboa SM, Gill SK, et al. Association between antibiotic use among pregnant women with urinary tract infections in the first trimester and birth defects, National Birth Defects Prevention Study 1997 to 2011. Birth Defects Res A Clin Mol Teratol. 2016;106(11):940–949. 10.1002/bdra.23570 [PubMed: 27891788]

28. US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, et al. Folic acid supplementation for the prevention of neural tube defects: US preventive services task force recommendation statement. JAMA. 2017;317(2):183–189. 10.1001/jama.2016.19438 [PubMed: 28097362]

29. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. In: The Health Consequences of Smoking-50 Years of Progress. A Report of the Surgeon General. Reports of the Surgeon General. Centers for Disease Control and Prevention; 2014.

30. O'Leary CM, Nassar N, Kurinczuk JJ, et al. Prenatal alcohol exposure and risk of birth defects. Pediatrics. 2010;126(4):e843–e850. 10.1542/peds.2010-0256 [PubMed: 20876169]

31. Mant J, Murphy M, Rose P, et al. The accuracy of general practitioner records of smoking and alcohol use: comparison with patient questionnaires. J Public Health Med. 2000;22(2):198–201. 10.1093/pubmed/22.2.198 [PubMed: 10912559]

32. Vuong AM, Shinde MU, Brender JD, et al. Prenatal exposure to Nitrosatable drugs, dietary intake of nitrites, and preterm birth. Am J Epidemiol. 2016;183(7):634–642. 10.1093/aje/kwv250 [PubMed: 26953287]

33. Esposito DB, Parker SE, Mitchell AA, et al. Periconceptional nonsteroidal anti-inflammatory drug use, folic acid intake, and the risk of spina bifida. Birth Defects Res. 2021;113(17):1257–1266. 10.1002/bdr2.1944 [PubMed: 34346174]

34. Tinker SC, Gilboa S, Reefhuis J, et al. Challenges in studying modifiable risk factors for birth defects. Curr Epidemiol Rep. 2015;2(1):23–30. 10.1007/s40471-014-0028-y [PubMed: 26236577]

35. Werler MM, Louik C, Mitchell AA. Case-control studies for identifying novel teratogens. Am J Med Genet C Semin Med Genet. 2011;157C(3):201–208. 10.1002/ajmg.c.30307 [PubMed: 21766434]

36. Lash TL, Fox MP, MacLehose RF, et al. Good practices for quantitative bias analysis. Int J Epidemiol. 2014;43(6):1969–1985. 10.1093/ije/dyu149 [PubMed: 25080530]

37. Howley MM, Fisher SC, Fuentes MA, et al. Agreement between maternal report and medical records on use of medications during early pregnancy in New York. Birth defects research. 2023;115(4):498–509. 10.1002/bdr2.2151 [PubMed: 36640121]

38. Fedak KM, Bernal A, Capshaw ZA, et al. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerg Themes Epidemiol. 2015;12:14. 10.1186/s12982-015-0037-4 [PubMed: 26425136]

39. Huybrechts KF, Bateman BT, Zhu Y, et al. Hydroxychloroquine early in pregnancy and risk of birth defects. Am J Obstet Gynecol. 2021;224(3):290.e1–290.e22. 10.1016/j.ajog.2020.09.007

40. Howley MM, Werler MM, Fisher SC, et al. Maternal exposure to hydroxychloroquine and birth defects. Birth defects research. 2021;113(17):1245–1256. 10.1002/bdr2.1943 [PubMed: 34296811]

41. Tinker SC, Crider KS, Ailes EC. Antibiotics and risk for birth defects. Br J Clin Pharmacol. 2018;84(7):1626–1627. 10.1111/bcp.13581 [PubMed: 29663497]

42. Muanda FT, Sheehy O, Berard A. Use of antibiotics during pregnancy and the risk of major congenital malformations: a population based cohort study. Br J Clin Pharmacol. 2017;83(11):2557–2571. 10.1111/bcp.13364 [PubMed: 28722171]

43. Crider KS, Cleves MA, Reefhuis J, et al. Antibacterial medication use during pregnancy and risk of birth defects: National Birth Defects Prevention Study. Arch Pediatr Adolesc Med. 2009;163(11):978–985. 10.1001/archpediatrics.2009.188 [PubMed: 19884587]