

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/182622>

Please be advised that this information was generated on 2021-10-26 and may be subject to change.

Administrative Claims Data Versus Augmented Pregnancy Data for the Study of Pharmaceutical Treatments in Pregnancy

Susan E. Andrade¹ · Anick Bérard² · Hedvig M.E. Nordeng^{3,4} · Mollie E. Wood³ · Marleen M.H.J. van Gelder^{5,6} · Sengwee Toh⁷

Published online: 18 April 2017
© Springer International Publishing AG 2017

Abstract

Purpose of Review Administrative claims databases, which collect reimbursement-related information generated from healthcare encounters, are increasingly used to evaluate medication safety in pregnancy. We reviewed the strengths and limitations of claims-only databases and how other data sources may be used to improve the accuracy and completeness of information critical for studying medication safety in pregnancy.

Recent Findings Research on medication safety in pregnancy requires information on pregnancy episodes, mother-infant linkage, medication exposure, gestational age, maternal and birth outcomes, confounding factors, and (in some studies) long-term follow-up data. Claims data reliably identifies live births and possibly other pregnancies. It allows mother-infant linkage and has prospectively collected prescription medication information. Its diagnosis and procedure information allows estimation of gestational age. It captures maternal med-

ical conditions but generally has incomplete data on reproductive and lifestyle factors. It has information on certain, typically short-term maternal and infant outcomes that may require chart review confirmation. Other data sources including electronic health records and birth registries can augment claims data or be analyzed alone. Interviews, surveys, or biological samples provide additional information. Nationwide and regional birth and pregnancy registries, such as those in several European and North American countries, generally contain more complete information essential for pregnancy research compared to claims-only databases.

Summary Claims data offers several advantages in medication safety in pregnancy research. Its limitations can be partially addressed by linking it with other data sources or supplementing with primary data collection. Rigorous assessment of data quality and completeness is recommended regardless of data sources.

This article is part of the Topical Collection on *Reproductive and Perinatal Epidemiology*

✉ Sengwee Toh
darren_toh@harvardpilgrim.org

¹ Meyers Primary Care Institute, Fallon Community Health Plan, Reliant Medical Group, University of Massachusetts Medical School, 425 North Lake Avenue, Worcester, MA 01605, USA

² Faculty of Pharmacy, and CHU Ste-Justine Research Center, University of Montreal, 3175 Côte-Ste-Catherine, Montreal, QC H3T 1C5, Canada

³ Pharmacoepidemiology and Drug Safety Research Group, School of Pharmacy, PharmaTox Strategic Research Initiative, Faculty of Mathematics and Natural Sciences, University of Oslo, P.O. Box 1068, Blindern, 0316 Oslo, Norway

⁴ Department of Child Health, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, 0403 Oslo, Norway

⁵ Department for Health Evidence, Radboud Institute for Health Sciences, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

⁶ Radboud REshape Innovation Center, Radboud University Medical Center, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

⁷ Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 401 Park Drive, Suite 401 East, Boston, MA 02215, USA

Keywords Administrative claims data · Birth registry · Electronic health records · Pharmacoepidemiology · Pregnancy

Introduction

Due to the ethical concerns of conducting randomized controlled trials of medications in pregnant women, well-designed observational studies have become the primary way of generating evidence on the benefits and risks of medication use in pregnant women and their offspring. Administrative claims databases are increasingly being used in observational pregnancy research. These databases offer several strengths over other more traditional epidemiological data sources, such as interview- or questionnaire-based data. On the other hand, they are created for administrative and billing purposes, so they have a number of limitations not typically seen in data collected specifically for pregnancy research or other types of electronic data sources (e.g., birth registries).

In recent years, researchers have attempted to address these limitations by linking claims databases with other data sources or supplementing it with information from primary data collection. In this article, we first briefly review data issues critical for studying medication safety in pregnancy. We then describe how adequately claims-only data sources address these issues and how “augmented pregnancy data” may be used to fill the remaining data gaps. We conclude the article with a discussion of future directions on this topic. We define augmented pregnancy data as data from (1) administrative claims databases linked to other data sources or (2) other electronic data sources that generally are thought to provide richer clinical or reproductive information compared to claims databases, such as population-based pregnancy or birth registries and electronic health records (EHRs).

Data Considerations in Medication Safety in Pregnancy Research

Assessing medication safety in pregnancy using observational data poses a number of unique challenges not seen in other types of epidemiological studies. This is because certain medications may be harmful to women or infants only if they are taken during specific gestational periods. For example, cardiac malformations originate in the first trimester, so late pregnancy exposure is not relevant when studying these outcomes. There are seven critical data components in observational studies of medication safety in pregnancy (Table 1). The validity of the study relies heavily on the ability to identify and analyze sufficiently large data sources that provide accurate and complete information on these data components.

Table 1 Data consideration in medication safety in pregnancy research

- (1) Information to identify pregnancy, including live birth, spontaneous abortion, pregnancy termination, and other pregnancy episodes
- (2) Information that allows mother-infant linkage
- (3) Information on medication exposure, including time of initiation and cessation of treatment
- (4) Information on gestational age, including start and end of pregnancy, to determine timing of exposure relative to gestational age
- (5) Information on maternal and birth outcomes, including birth defects
- (6) Information on potential confounding factors, including indication for use, comedication, lifestyle, and reproductive factors
- (7) Information on long-term follow-up of infants or mothers, if long-term effect of prenatal exposure is of interest

Administrative Claims Data Only and Its Application in Medication Safety in Pregnancy Research

Administrative claims data from commercial health insurance companies, government-sponsored health plans, and health maintenance organizations in the USA, Canada, and some Asia-Pacific, South American, and European countries have been used for studies of medication use and safety in pregnant women. Created primarily for administrative, financial, and reimbursement purposes, these databases include information on claims submitted by healthcare providers for payment and records of patient encounters with healthcare systems. They generally contain information in the following domains: (1) plan enrollment, e.g., beginning and end dates; (2) demographics, e.g., birth date and sex; (3) outpatient pharmacy dispensing, e.g., codes to determine the drug product, dispense date, amount dispensed, and days’ supply; (4) inpatient and outpatient medical encounters, e.g., dates of service, admission and discharge dates, and providers or facilities; (5) diagnoses recorded for each encounter, e.g., diagnosis codes recorded as International Classification of Diseases, 9th or 10th Revision, Clinical Modification (ICD-9-CM or ICD-10-CM) codes; and (6) procedures done at each encounter, e.g., procedure codes to identify surgeries or laboratory tests ordered. The databases typically have unique personal identifiers for tracking of individuals across various data domain files.

Some of these databases include data from millions of individuals, allowing researchers to conduct studies in large, well-defined insured populations. They also collect detailed information on prescription medications, diagnoses, and procedures performed. The information is generally collected prospectively and not subject to recall bias. We discuss the strengths and limitations of administrative claims data in medication safety in pregnancy research below.

Strengths of Using Administrative Claims Data in Studying Medication Safety in Pregnancy

Information to Identify Pregnancy Recent studies have shown that administrative claims databases in the USA and Canada can accurately identify most pregnancies resulting in a live birth [1•, 2, 3]. Some studies have also shown that these databases can also identify other pregnancy episodes (e.g., spontaneous abortion) [1•, 2, 3, 4•]. The ability to efficiently identify pregnancies in large, population-based cohorts is a tremendous advantage compared to identification using voluntary pregnancy registries for specific drugs, patient interviews, or teratology information services. Most of the studies that rely on these more traditional epidemiological data sources enroll a relatively small number of women who volunteer or agree to participate in the study. The time and cost to collect the data using these traditional methods are generally much greater than using administrative claims data. In addition, these studies may be more prone to selection bias if participation is more (or less) likely among those exposed to a medication who have an outcome of interest.

Information That Allows Mother-Infant Linkage To be useful for research of medication safety in pregnancy, information from the mother and the infant must be linked to each other. For a number of administrative claims databases, researchers have determined these mother-infant linkages using a variety of methods, such as unique family identification numbers included in the health plan enrollment data and definitive or probabilistic name and address matching. The linkage rates ranged from approximately 45 to 90% [5–8]. One example of a claims database that has created these internal mother-infant linkages to conduct studies of medication use and safety in pregnancy using administrative data only is the Medicaid Analytic Extract (MAX) files, a US government database including more than one million mother-infant pairs, comprising mostly low-income or chronically disabled enrollees [5].

Information on Medication Exposure Studies have shown substantial underreporting of prescription medication use using paper-based questionnaires or interviews in retrospective studies [9–12]. Electronic pharmacy dispensing data from administrative claims databases provides an advantage over patient interviews or surveys as the information is collected free from recall bias. The information is available on the exact date of dispensing along with other information, such as amount dispensed and days' supply.

Information on Gestational Age Pregnancy-related diagnosis and procedure codes can be used to estimate trimesters and weeks of gestation in administrative claims databases [1•, 2, 3, 13•, 14, 15•]. A number of algorithms have been validated

against the gestational age information in birth certificates or medical records [2, 13•]. They have been shown to perform well in estimating gestational age and classifying prenatal exposure status, particularly for chronically used medications.

Information on Maternal and Birth Outcomes The encounter files in administrative claims databases contain diagnosis and procedure codes that can be used to identify medically attended maternal and birth conditions. The validity of these codes varies by outcome [15•, 16].

Information on Confounding Factors Administrative claims data provides information on a number of potential confounding factors, including maternal demographics, maternal conditions that lead to medical attention, and prescription drug exposures other than the drug of interest.

Information on Long-Term Follow-Up of Infants or Mothers Follow-up of women and their children in administrative claims databases generally can continue after birth as long as they remain members of the health plan.

Limitations of Administrative Claims Data in Studying Medication Safety in Pregnancy

There are a number of limitations with using only administrative claims data for pregnancy research. A number of possibilities exist for misclassification of exposure. These databases generally do not capture information that provides more accurate estimation of pregnancy start date, such as last menstrual period and clinical or obstetric estimate of gestational age. Although claims-based gestational age algorithms have been shown to perform well on average (e.g., a difference of a few days in average gestational age), individual woman's gestational length can be misclassified by a considerable amount (e.g., weeks), potentially resulting in exposure misclassification [13•, 17]. While the electronic pharmacy files document that a medication was dispensed, information on whether or not the medication was taken as directed is not available. While some studies have shown that prescription fillings are valid estimates for actual medication use [18], others have reported that noncompliance for some medications may be common among pregnant women, especially in the first trimester [19–21]. Such noncompliance may pose a major threat to the validity of medication safety in pregnancy research. Inpatient medication exposures and over-the-counter (OTC) use are often incompletely documented or not captured at all.

Misclassification of certain birth outcomes is also possible. The validity of claims-based algorithms for some birth defects may be questionable, requiring confirmation through medical record review [15•, 16]. For example, Cooper et al. showed that the positive predictive value of claims-based outcome

algorithms varied between 34% for hydrocephaly and 93% for oral clefts [16]. Some of these databases may have relatively short follow-up (<1 year) of the infant and mother due to health plan turnover, limiting their ability to study long-term effects of medication use. The state Medicaid databases in the USA typically have shorter enrollment of children, while integrated delivery systems generally have longer follow-up than other commercial health plans.

Information on certain potential confounding variables of interest is also not available or poorly documented in administrative claims data. For example, maternal characteristics including reproductive history, maternal level of education, and OTC folate use or dietary folate consumption are generally not captured. While the enrollment file may contain information on other characteristics such as race/ethnicity and the encounter files may have diagnoses or procedures related to tobacco use, alcohol use, and obesity, the information is generally incompletely recorded.

Augmented Pregnancy Data

Due to the limitations of using only administrative claims data to evaluate the safety of medication use in pregnancy, a number of researchers have linked administrative claims databases to other data sources. These data sources include other databases within healthcare organizations, registries created by various organizations, and data collected by direct contact with the mothers, children, or healthcare providers. Depending on the data source, additional or more accurate information on gestational age, selected maternal and birth outcomes, or potential confounders may be recorded. In some European and North American countries, population-based EHRs and birth registries are available for pregnancy research and can be analyzed alone. We describe examples of some of these data sources and the information available below and in Table 2.

Electronic Health Records

Integrated healthcare delivery systems within the USA have linked administrative claims data with ambulatory or inpatient EHRs for a wide array of research studies, including studies of medication use in pregnancy [26, 27]. Several European general practice EHR databases are also available for pregnancy research [28, 29]. EHRs from healthcare providers and healthcare systems may provide information on important potential confounders not well-recorded in administrative claims data, including tobacco use, alcohol use, blood pressure, and height and weight. Information from clinical encounter data and laboratory test results may also provide data to identify potential maternal or birth outcomes of interest, such as head circumference and test result-confirmed gestational diabetes

and preeclampsia. Since the data are recorded as part of clinical care and are collected prospectively, studies using EHRs are typically not subject to recall bias.

National or Regional Birth, Death, Fetal Death, and Malformation Registries

Some researchers have linked administrative health plan or prescription data to national or regional government birth registries to identify data not available or incompletely captured in claims or prescription data [4••, 22, 30–32]. In the USA, birth certificate data may be obtained from many state departments of public health for research purposes [22, 33]. These data sources include more accurate estimates of gestational age and estimation methods (e.g., obstetric or clinical estimates) and other maternal, paternal, and birth information that may be important for determining potential confounders (e.g., tobacco use, parity, gravidity, maternal and paternal education, race/ethnicity), or birth outcomes (e.g., birth weight). Fetal death reports provide similar information for stillbirths. In the USA, augmenting administrative claims data with birth certificate data has increased the mother-infant linkage rates or has been used as the primary linkage method for some health plans [34]. Health plans that maintain birth registries have generally been able to use the registry information to link $\geq 95\%$ of deliveries/mothers identified in administrative data [34]. Healthcare organizations in the USA have also linked their administrative claims data to national or regional government death registries for epidemiological studies, including studies of medication safety in pregnancy [27, 28]. This information can be used to identify maternal and infant mortality and cause of death.

The Quebec Pregnancy Cohort in Canada combines data from four electronic databases [4••]. The Régie de l'Assurance Maladie du Québec (RAMQ) database provides administrative health insurance information on medical encounters and medication use. The Med-Echo database records all acute care hospitalizations and includes information on gestational length and birth weight (clinical database). The birth registry of the Institut de la Statistique du Québec (ISQ) provides additional demographic information on the mother, father, and infant (e.g., education level, marital status, birth weight, gestational age, parity, ethnicity) for live births and stillbirths. The Ministère de l'éducation, des loisirs et des sports du Québec (MELS) database provides information on use of specialized services at elementary schools such as speech therapist or psychologists. Linkage of these databases allows for identification of a large pregnancy cohort with comprehensive information on medication exposures, pregnancy and birth outcomes (including longer-term outcomes), and potential confounders of interest for studies of medication safety in pregnancy.

Table 2 Examples of availability and source of information in administrative claims and augmented pregnancy databases

Data elements	Examples of augmented pregnancy databases					
	Administrative claims data	Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) [22]	Quebec Pregnancy Cohort [4••]	Nordic birth registries [23••]	European birth registries	EUROMedCAT ^b
Type of pregnancy Live births	Yes; based on diagnosis and procedure codes	Yes; based on diagnosis and procedure codes	Yes; based on diagnosis and procedure codes	Yes; mandatory reporting to national registry	Yes; not mandatory for all registries	Yes; only infants with congenital malformations, voluntarily
Stillbirths	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Yes; mandatory reporting of all pregnancies ≥ 22 weeks	Partial; varies by registry	Yes; minimum gestational age varies by registry
Spontaneous abortions	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; any that occur after 22 weeks gestation (12 weeks in Norway)	Partial; varies by registry	No
Elective/therapeutic abortions	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; varies by registry	Partial; varies by registry	Yes; illegal in some countries
Mother-infant linkage	Yes; based on unique family identification numbers, name and address matching, or birth certificate data	Yes; based on unique family identification numbers or name and address matching	Yes; based on unique family identification numbers or name and address matching	Yes; based on national identification numbers	Yes; mostly probabilistic linkage methods	Yes
Medication exposure Prescription drugs	Yes; based on dispensing data	Yes; based on dispensing data	Yes; based on dispensing data (some validated against real-time maternal report) [18]	Partial; reporting may be incomplete in birth records. Pharmacy records available through separate linkage	Yes; mostly based on pharmacy dispensing data; some on prescription data from GPs	Yes; based on medical records, dispensing data, or maternal interviews
Over-the-counter drugs	Partial; only if prescribed	Partial; only if prescribed	Partial; only if prescribed	Partial; reporting may be incomplete in birth records.	No	Partial; not available in most registries
Inpatient drugs	Partial; based on procedure data	Partial; based on procedure data	Partial; based on questionnaire	Partial; as reported to antenatal care provider	Partial	Partial; based on medical/obstetric records
Actual gestational age	No; estimated based on diagnosis and procedure codes	Yes; based on birth certificate data	Yes; first day of last menstrual period for all pregnancies regardless of pregnancy outcomes validated against patient charts [24•]	Yes; based on ultrasound for majority of pregnancies or reporting of last menstrual period if ultrasound is not available	Partial; estimated in some registries, available in most registries	Yes
Maternal outcomes Preeclampsia	Yes; based on diagnosis codes and drug use ^a	Yes; based on diagnosis codes and drug use ^a	Yes; based on diagnosis codes and drug use ^a	Yes; reported by antenatal care provider and/or patient registry	Partial; registered by antenatal care provider and/or GPs	No; might be reported as maternal illness
Gestational diabetes	Yes; based on diagnosis codes and drug use ^a	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes and drug use ^a	Yes; reported by antenatal care provider and/or patient registry	Partial; registered by antenatal care provider and/or GPs	No; might be reported as maternal illness
Postpartum depression	Yes; based on diagnosis codes and drug use ^a	Yes; based on diagnosis codes and drug use ^a	Yes; based on diagnosis codes and drug use ^a	Yes; though linkage to patient registry and based on diagnosis codes	Partial; only in registries with GP data	No
Postpartum hemorrhage	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes ^a	No	No

Table 2 (continued)

Data elements	Examples of augmented pregnancy databases				EUROMedicAT ^b
	Administrative claims data	Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) [22]	Quebec Cohort [4•]	Nordic birth registries [23••]	
Birth outcomes					
Congenital malformation	Typical US administrative claims data	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes ^a and birth certificate data	Yes; reported by antenatal care provider and/or patient registry	Partial; registered by antenatal care provider and/or GPs
Birth weight	Yes; based on diagnosis codes ^a	Yes; based on diagnosis codes ^a	Yes; diagnosis codes validated against medical charts [25•]	Yes; reported by antenatal care provider or patient registry; some countries maintain separate malformations registry	Yes; patient registries, congenital anomaly databases, or infant records from GPs
Birth length	Partial; based on diagnosis codes ^a	Yes; based on birth certificate data	Yes; validated against medical charts [24•]	Yes; reported by antenatal care provider	Generally yes; registries or GP records
Head circumference	No	No	Partial	Yes; reported by antenatal care provider	Partial
Prematurity	Partial; based on diagnosis codes ^a	Yes; based on birth certificate data	Partial	Yes in some countries	Partial
			Yes; based on gestational age validated against medical charts [24•]	Yes; as reported by antenatal care provider and/or calculated from ultrasound or last menstrual period-based gestational age estimate	Yes; patient registries or GP data
Apgar score	No	Yes; based on birth certificate data	Partial	Yes; 1 and 5 min Apgar scores	No
Maternal covariates					
Age	Yes; based on demographic info	Yes; based on demographic info	Yes; based on demographic info	Yes; based on demographic info	Yes; based on demographic info
Medical conditions	Yes; based on diagnosis and procedure codes and drug use	Yes; based on diagnosis and procedure codes	Yes; based on diagnosis and procedure codes and drug use	Partial; information varies by registry but additional info available through linkage to other registries or through chart review	Partial; based on medical records for most registries
Reproductive history	Generally no	Yes; based on birth certificate data	Yes; based on diagnosis and procedure codes, drug use, or questionnaire	Yes	Partial; varies by registry
Tobacco use	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes and birth certificate data	Partial; based on questionnaire	Partial; recorded in some registries but reporting is voluntary	Partial; varies by registry
Alcohol use	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; based on questionnaire	Partial; available in some registries or by linkage to other registries	Partial; varies by registry
Body mass index	Partial; based on diagnosis and procedure codes	Partial; based on diagnosis and procedure codes	Partial; based on questionnaire	Partial; recorded in Finland and Sweden, availability varies in other countries	Partial; varies by registry
Folic acid supplementation	Partial; based upon prescription dispensing data (high dose)	Partial; based upon prescription dispensing data (high dose)	Yes for high dosage prescription (>5 mg/day); partial for other dosage and based on questionnaire	Partial; varies by registry, formulation (high vs. low dose) and year	Partial; based upon prescription data (high dose)

Table 2 (continued)

Data elements	Administrative claims data	Examples of augmented pregnancy databases
Typical US administrative claims data	Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) [22]	Quebec Pregnancy Cohort [4••] Nordic birth registries [23••] European birth registries EUROMediCAT ^b
Long-term follow-up Neurodevelopment, somatic childhood illness	Partial; based on diagnosis and procedure codes ^a	Yes; based on diagnosis codes, and school performance Partial; based on diagnosis codes ^a by linkage to national patient registry Partial, varies by registry No

GP general practitioner

^a May require confirmation through medical record review

^b Only included infants with congenital malformations

All Nordic countries (Norway, Denmark, Sweden, Iceland, and Finland) have mandatory reporting of births to national organizations known generally as medical birth registries [35–38]. The registries began capturing comprehensive population-based data as early as 1967 (Norway) or as late as 1987 (Finland). These registries include live born infants and stillbirths occurring after 22 weeks and routinely collect information on gestational age at birth, birth weight, congenital malformations, and the reproductive history of the mother. These registries also gather some information on the type of birth and delivery characteristics and complications, but the type and quality of the information vary. Select variables from birth registries, including serious pregnancy complications and gestational age, have been validated through medical chart reviews [15••, 25•, 35, 39–42]. Medication use during pregnancy is recorded by the prenatal care provider, and may or may not include data on OTC medications. This information can be linked to national prescription registries, which are comprehensive records of pharmacy dispensing or reimbursement [43].

The Nordic countries also have registries on termination of pregnancies that record elective or therapeutic abortions occurring after 12 weeks of gestation, with varying information on the reason for the abortion. National patient registries include diagnosis and procedure codes for all contacts with the healthcare system; the birth and patient registries generally require special linkage, with a possible exception of Denmark, where the birth registry and national patient registry merged in 1995. Patient registries present an opportunity to carry out research on longer-term outcomes of medication use during pregnancy, for mothers and their offspring.

In many other European countries, similar national and regional registries exist, but reporting of births is sometimes not mandatory [44]. Most of these registries do not include personal identifiers such as a national identification numbers, which may hamper correct linkage with other data sources. The EUROMediCAT project, a collaborative initiative under the European Surveillance of Congenital Anomalies (EUROCAT), includes 15 medical birth registries in 13 countries and 7 healthcare databases in 5 countries covering a population of 7.2 million births from 1995 to 2012 [45]. Within this project, national or regional congenital anomaly registry data were linked to primary care or prescription administrative databases [46]. The registries within the network collect data on live births, fetal deaths, and pregnancy terminations with congenital anomalies, including information on date of birth, gestational age, and maternal age, medication use, and comorbidities.

Due to their universal healthcare systems, Canada and some European countries (including the Nordic countries) generally have very low turnover rate in their databases. As discussed above, these countries have combined multiple databases with high linkage rates for pregnancy research. These

linked databases are highly advantageous for evaluating long-term effects of prenatal exposure such as effects on neurodevelopment (e.g., diagnoses of attention deficit hyperactivity disorder or autism spectrum disorders) and other conditions (e.g., childhood cancer), especially if a diagnosis is available and coded [47].

Surveys

Researchers have also supplemented administrative claims data with data collected through interviews or questionnaires of the mothers, infants, or healthcare providers. For example, healthcare organizations in the USA have collected data on pregnant women through self-administered questionnaires or interviews [48]. For the Quebec Pregnancy Cohort, self-administered questionnaire data are collected bi-annually for a random sample of pregnancies ending with a live birth [49]. Information collected includes lifestyle factors (e.g., smoking, alcohol, and physical activity), socio-demographic information, OTC medication use, weight and height at the beginning of the pregnancy, weight gain during pregnancy, natural health product use, folic acid intake, and reproductive history including use of assisted reproduction techniques such as in vitro fertilization.

Norway and Denmark have both carried out large questionnaire-based birth cohort studies, each with more than 100,000 pregnancies [49–51]. The medication exposure data in these cohort studies are particularly rich, as they include both prescription and OTC drugs, as well as indications for use of each drug. Further, these studies include psychometric instruments to assess maternal mental health during pregnancy as well as children's neurodevelopment at various ages [24, 52, 53]. Data collection for a similar cohort study, the PRegnancy and Infant DEvelopment (PRIDE) Study, is currently ongoing in The Netherlands [54]. In this cohort, the detailed data from web-based questionnaires on prescription and OTC medication use are enriched with dispensing data from pharmacies. These questionnaire-based studies are routinely linked to birth registry data, and may also be linked to patient registries and prescription drug registries.

Biological Samples

A large number of pregnancy or birth cohorts now include biobanks (<http://www.birthcohorts.net/>). These biological samples could be used to study the effect of underlying genetic and metabolic status on the risk of adverse pregnancy outcomes. Recently, the Quebec Pregnancy Cohort has started collecting saliva on mothers and children already present in the cohort (more than 300,000 women and 260,000 children), as well as mothers and children who will be entering the cohort in the future [49]. Availability of biological specimens for Nordic countries varies, with

Iceland and Finland both maintaining population-wide serum banks and Denmark and Sweden both storing dried blood spots from phenylketonuria (PKU) testing [50].

Discussion

Administrative claims data is increasingly being used to study medication safety in pregnancy due to its ability to identify large populations in a timely, efficient manner. Due to limitations of studies using only administrative claims data, a number of other data sources have been used to augment claims data or as a stand-alone data source. In general, augmenting administrative claims data with information from other electronic data sources or through primary data collection can help improve data completeness. However, these augmentation approaches also have their own constraints, as we discuss below.

EHRs generally record medications prescribed to the mother. As with administrative claims data, information on whether or not the medication was actually taken is not available. EHRs also have incomplete capture of OTC medication use, diagnoses, procedures, and medication prescriptions made outside the healthcare system. Although registries that capture stillbirths and elective or therapeutic abortions can provide more complete capture of all pregnancies, they may not capture early pregnancy losses given that many will occur among clinically unrecognized pregnancies. Studies that include only live births or very late pregnancy loss may sometimes systematically underestimate medication exposure and its associated risk, if women who take the medication are more likely to suffer early pregnancy loss or elect to terminate the pregnancy [55, 56].

In principle, information not available in administrative claims data could be obtained from other data sources or by contacting mothers, children, or healthcare providers. In practice, however, a typical pregnancy study can only collect such information in a subset of the study population, due to incomplete overlap in populations in various electronic data sources, the cost of primary data collection, and the ability to identify or contact the mothers, children, or their healthcare providers. Therefore, missing data may remain an issue in the augmented data. It is unlikely that the data will be missing completely at random [57]; therefore, appropriate statistical methods, such as multiple imputation [58, 59] and inverse probability weighting [60, 61], will need to be employed to adjust for missing data. An external or internal validation sample with more accurate exposure or covariate information can allow researchers to adjust effect estimates for incomplete or missing information [62–65].

Another methodological issue with augmenting administrative claims data with information from primary data collection is that women or children who participate in the study may be systematically different from those who do not, which

may result in selection bias [66, 67] or affect the generalizability of the study findings. Appropriate analytic methods, such as inverse probability weighting [66], can be used to account for selection bias if factors associated with willingness to participate are measured. Finally, a single data source may not have enough number of pregnancies or outcomes of interest; therefore, pooling of multiple data sources for specific pregnancy studies is often needed to achieve sufficient statistical power [22, 68, 69, 70].

Conclusions

In summary, administrative claims data offers a number of advantages in studying medication safety in pregnancy. Its limitations can be partially addressed by linking it with other electronic data sources. Supplementing administrative claims data with information from primary data collection may also be advantageous. However, other data sources, such as population-based birth registries, prospective cohort studies, or case-control studies, may sometimes be more appropriate for specific research questions. Regardless of data sources, rigorous assessment of data quality and completeness and use of appropriate analytic methods are always recommended in medication safety in pregnancy research.

Compliance with Ethical Standards

Conflict of Interest Anick Bérard, Hedvig M.E. Nordeng, Mollie E. Wood, Marleen M.H.J. van Gelder, and Sengwee Toh each declare no potential conflict of interest.

Susan E. Andrade reports a grant from Pfizer, Inc., outside the submitted work.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1.•• Naleway AL, Gold R, Kurosky S, Riedlinger K, Henninger ML, Nordin JD, et al. Identifying pregnancy episodes, outcomes, and mother-infant pairs in the Vaccine Safety Datalink. *Vaccine*. 2013;31(27):2898–903. doi:10.1016/j.vaccine.2013.03.069. **Validates algorithms that use electronic health plan data to identify live births, spontaneous abortion, elective abortions, stillbirths, and other pregnancy episodes in the Vaccine Safety Datalink. Using information in medical charts as the reference, the authors find that the algorithms have high validity in identifying these pregnancy episodes.**

2. Margulis AV, Palmsten K, Andrade SE, Charlton RA, Hardy JR, Cooper WO, 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):335–42. doi:10.1002/pds.3743.
3. Raebel MA, Ellis JL, Andrade SE. Evaluation of gestational age and admission date assumptions used to determine prenatal drug exposure from administrative data. *Pharmacoepidemiol Drug Saf*. 2005;14(12):829–36. doi:10.1002/pds.1100.
- 4.•• Berard A, Sheehy O. The Quebec Pregnancy Cohort—prevalence of medication use during gestation and pregnancy outcomes. *PLoS One*. 2014;9(4):e93870. doi:10.1371/journal.pone.0093870. **Describes in detail the Quebec Pregnancy Cohort and its underlying databases linkages. It also gives baseline frequencies for the most used medications during pregnancy and the most prevalent adverse birth outcomes such as major malformations, prematurity, and low birth weight. In addition, it identifies important determinants of medication exposure during pregnancy.**
5. Palmsten K, Huybrechts KF, Mogun H, Kowal MK, Williams PL, Michels KB, et al. Harnessing the Medicaid Analytic eXtract (MAX) to evaluate medications in pregnancy: design considerations. *PLoS One*. 2013;8(6):e67405. doi:10.1371/journal.pone.0067405.
6. Davis RL, Rubanowice D, McPhillips H, Raebel MA, Andrade SE, Smith D, et al. Risks of congenital malformations and perinatal events among infants exposed to antidepressant medications during pregnancy. *Pharmacoepidemiol Drug Saf*. 2007;16(10):1086–94. doi:10.1002/pds.1462.
7. Cole JA, Modell JG, Haight BR, Cosmatos IS, Stoler JM, Walker AM. Bupropion in pregnancy and the prevalence of congenital malformations. *Pharmacoepidemiol Drug Saf*. 2007;16(5):474–84. doi:10.1002/pds.1296.
8. Mines D, Tennis P, Curkendall SM, Li DK, Peterson C, Andrews EB, et al. Topiramate use in pregnancy and the birth prevalence of oral clefts. *Pharmacoepidemiol Drug Saf*. 2014;23(10):1017–25. doi:10.1002/pds.3612.
9. de Jong-van den Berg LT, Waardenburg CM, Haaijer-Ruskamp FM, Dukes MN, Wesseling H. Drug use in pregnancy: a comparative appraisal of data collecting methods. *Eur J Clin Pharmacol*. 1993;45(1):9–14. doi:10.1007/BF00315343.
10. Newport DJ, Brennan PA, Green P, Ilardi D, Whitfield TH, Morris N, et al. Maternal depression and medication exposure during pregnancy: comparison of maternal retrospective recall to prospective documentation. *BJOG*. 2008;115(6):681–8. doi:10.1111/j.1471-0528.2008.01701.x.
11. van Gelder MM, van Rooij IA, de Walle HE, Roeleveld N, Bakker MK. Maternal recall of prescription medication use during pregnancy using a paper-based questionnaire: a validation study in the Netherlands. *Drug Saf*. 2013;36(1):43–54. doi:10.1007/s40264-012-0004-8.
12. Sundermann AC, Hartmann KE, Jones SH, Torstenson ES, Velez Edwards DR. Validation of maternal recall of early pregnancy medication exposure using prospective diary data. *Ann Epidemiol*. 2016; doi:10.1016/j.annepidem.2016.11.015.
- 13.• Li Q, Andrade SE, Cooper WO, Davis RL, Dublin S, Hammad TA, et al. Validation of an algorithm to estimate gestational age in electronic health plan databases. *Pharmacoepidemiol Drug Saf*. 2013;22(5):524–32. doi:10.1002/pds.3407. **Validates an algorithm that uses delivery date and diagnosis codes to define gestational age at birth using electronic data from eight health plans participating in the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP). Using gestational age obtained from the infant birth certificate file as the reference, the authors find that the gestational age algorithm correctly classifies prenatal medication exposure**

- status in most live born deliveries, but trimester-specific misclassification may be higher for medications not intended for long-term use.
14. Andrade SE, Toh S, Houstoun M, Mott K, Pitts M, Kieswetter C, et al. Surveillance of medication use during pregnancy in the Mini-Sentinel Program. *Matern Child Health J.* 2016;20(4):895–903. doi:10.1007/s10995-015-1878-8.
 - 15.●● Andrade SE, Scott PE, Davis RL, Li DK, Getahun D, Cheetham TC, et al. Validity of health plan and birth certificate data for pregnancy research. *Pharmacoepidemiol Drug Saf.* 2013;22(1):7–15. doi:10.1002/pds.3319. **Evaluates the validity of health plan and birth certificate data for pregnancy research in the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP). The authors find that health plan and birth certificate data can be useful to accurately identify select infant outcomes (e.g., preterm birth, select cardiac defects), maternal diagnoses (e.g., asthma), and newborn, maternal, and paternal characteristics (e.g., birth weight, race/ethnicity, prior obstetrical history). The findings also underscore the importance of medical record review for validation for select other outcomes.**
 16. Cooper WO, Hernandez-Diaz S, Gideon P, Dyer SM, Hall K, Dudley J, et al. Positive predictive value of computerized records for major congenital malformations. *Pharmacoepidemiol Drug Saf.* 2008;17(5):455–60. doi:10.1002/pds.1534.
 17. Toh S, Mitchell AA, Werler MM, Hernández-Díaz S. 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):633–40. doi:10.1093/aje/kwm367.
 18. Zhao JP, Sheehy O, Gorgui J, Bérard A. Can We Rely on Pharmacy Claims Databases to Ascertain Maternal Use of Medications during Pregnancy? *Birth Defects Res.* 2017 Apr 3;109(6):423–431.
 19. Olesen C, Sondergaard C, Thrane N, Nielsen GL, de Jong-van den Berg L, Olsen J. Do pregnant women report use of dispensed medications? *Epidemiology.* 2001;12(5):497–501.
 20. Lupattelli A, Spigset O, Bjornsdottir I, Hameen-Anttila K, Mardby AC, Panchaud A, et al. Patterns and factors associated with low adherence to psychotropic medications during pregnancy—a cross-sectional, multinational web-based study. *Depress Anxiety.* 2015;32(6):426–36. doi:10.1002/da.22352.
 21. de Jonge L, de Walle HE, de Jong-van den Berg LT, van Langen IM, Bakker MK. Actual use of medications prescribed during pregnancy: a cross-sectional study using data from a population-based congenital anomaly registry. *Drug Saf.* 2015;38(8):737–47. doi:10.1007/s40264-015-0302-z.
 22. Andrade SE, Davis RL, Cheetham TC, Cooper WO, Li DK, Amini T, et al. Medication exposure in pregnancy risk evaluation program. *Matern Child Health J.* 2012;16(7):1349–54. doi:10.1007/s10995-011-0902-x.
 - 23.●● Langhoff-Roos J, Krebs L, Klungsoyr K, Bjarnadottir RI, Kallen K, Tapper AM, et al. The Nordic medical birth registers—a potential goldmine for clinical research. *Acta Obstet Gynecol Scand.* 2014;93(2):132–7. doi:10.1111/aogs.12302. **Gives a comprehensive overview of the Nordic birth registries, such as the history of the registries, the data available for each country (including dates from which data are available), and the potential strengths and weaknesses of data, including information on validity and quality. For readers interested in learning more about the Nordic birth registries, this is an excellent starting point and primer.**
 24. Mathiesen KS, Tambs K. The EAS temperament questionnaire—factor structure, age trends, reliability, and stability in a Norwegian sample. *J Child Psychol Psychiatry.* 1999;40(3):431–9. doi:10.1111/1469-7610.00460.
 - 25.● Vilain A, Otis S, Forget A, Blais L. Agreement between administrative databases and medical charts for pregnancy-related variables among asthmatic women. *Pharmacoepidemiol Drug Saf.* 2008;17(4):345–53. doi:10.1002/pds.1558. **Validates gestational age in the Quebec Pregnancy Cohort against medical and hospital charts. The study finds that gestational age taken within the Quebec Pregnancy Cohort, regardless of pregnancy outcome (spontaneous or planned abortion, delivery, stillbirth), is a valid measure of actual gestational age. This increases the validity of studies performed within this pregnancy cohort.**
 - 26.● Blais L, Berard A, Kettani FZ, Forget A. Validity of congenital malformation diagnostic codes recorded in Quebec’s administrative databases. *Pharmacoepidemiol Drug Saf.* 2013;22(8):881–9. doi:10.1002/pds.3446. **Reports validity of the diagnosis codes used for major malformations within the Quebec Pregnancy Cohort. It concludes that all ICD-9 and ICD-10 codes used to study major malformations within the Quebec Pregnancy Cohort are valid. The validity of these codes depends on the healthcare system in Canada and might not be extrapolated to other cohorts. This increases the validity of studies performed within this pregnancy cohort.**
 27. Ross TR, Ng D, Brown JS, Pardee R, Hornbrook MC, Hart G, et al. The HMO research network virtual data warehouse: a public data model to support collaboration. *EGEMS (Wash DC).* 2014;2(1):1049. doi:10.13063/2327-9214.1049.
 28. Andrade SE, Raebel MA, Boudreau D, Davis RL, Haffner K, Pawloski PA, et al. Chapter 12: health maintenance organizations/health plans. In: Strom BL, Kimmel SE, Hennessy S, editors. *Pharmacoepidemiology.* 5th ed. Chichester: Wiley-Blackwell; 2012.
 29. Devine S, West S, Andrews E, Tennis P, Hammad TA, Eaton S, et al. The identification of pregnancies within the general practice research database. *Pharmacoepidemiol Drug Saf.* 2010;19(1):45–50. doi:10.1002/pds.1862.
 30. Charlton RA, Neville AJ, Jordan S, Pierini A, Damase-Michel C, Klungsoyr K, et al. Healthcare databases in Europe for studying medicine use and safety during pregnancy. *Pharmacoepidemiol Drug Saf.* 2014;23(6):586–94. doi:10.1002/pds.3613.
 31. Pedersen LH, Petersen OB, Norgaard M, Ekelund C, Pedersen L, Tabor A, et al. Linkage between the Danish National Health Service Prescription Database, the Danish Fetal Medicine Database, and other Danish registries as a tool for the study of drug safety in pregnancy. *Clin Epidemiol.* 2016;8:91–5. doi:10.2147/CLEP.S98139.
 32. Zomerdijk IM, Ruiter R, Houweling LM, Herings RM, Straus SM, Stricker BH. Dispensing of potentially teratogenic drugs before conception and during pregnancy: a population-based study. *BJOG.* 2015;122(8):1119–29. doi:10.1111/1471-0528.13128.
 33. Andrade SE, Haffner K, Rosofsky R, Dublin S. Birth certificate data matching for the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) program: survey of state and city departments of public health. 2012. https://www.sentinelssystem.org/sites/default/files/data/ComplementaryData/Mini-Sentinel_PRISM_Birth-Certificate-Data-Matching_Survey-State-City-Dept-Public-Health.pdf. Accessed November 12 2016
 34. Johnson KE, Beaton SJ, Andrade SE, Cheetham TC, Scott PE, Hammad TA, et al. Methods of linking mothers and infants using health plan data for studies of pregnancy outcomes. *Pharmacoepidemiol Drug Saf.* 2013;22(7):776–82. doi:10.1002/pds.3443.
 35. Kristensen J, Langhoff-Roos J, Skovgaard LT, Kristensen FB. Validation of the Danish Birth Registration. *J Clin Epidemiol.* 1996;49(8):893–7.
 36. Irgens LM. The Medical Birth Registry of Norway: a source for epidemiological and clinical research. *Scand J Rheumatol Suppl.* 1998;107:105–8. doi:10.1080/03009742.1998.11720780.
 37. Axelsson O. The Swedish medical birth register. *Acta Obstet Gynecol Scand.* 2003;82(6):491–2. doi:10.1034/j.1600-0412.2003.00172.x.

38. Gissler M, Haukka J. Finnish health and social welfare registers in epidemiological research. *Nor Epidemiol*. 2004;14(1):113–20.
39. Sneider K, Langhoff-Roos J, Sundtoft IB, Christiansen OB. Validation of second trimester miscarriages and spontaneous deliveries. *Clin Epidemiol*. 2015;7:517–27. doi:10.2147/CLEP.S85107.
40. Klungsoyr K, Harmon QE, Skard LB, Simonsen I, Austvoll ET, Alsaker ER, et al. Validity of pre-eclampsia registration in the medical birth registry of Norway for women participating in the Norwegian mother and child cohort study, 1999–2010. *Paediatr Perinat Epidemiol*. 2014;28(5):362–71. doi:10.1111/ppe.12138.
41. Kubon C, Sivertsen A, Vindenes HA, Abyholm F, Wilcox A, Lie RT. Completeness of registration of oral clefts in a medical birth registry: a population-based study. *Acta Obstet Gynecol Scand*. 2007;86(12):1453–7. doi:10.1080/08037050701645090.
42. Klemmensen AK, Olsen SF, Osterdal ML, Tabor A. Validity of preeclampsia-related diagnoses recorded in a national hospital registry and in a postpartum interview of the women. *Am J Epidemiol*. 2007;166(2):117–24. doi:10.1093/aje/kwm139.
43. Kieler H. Nordic databases to evaluate medications in pregnancy. *Therapie*. 2014;69(1):65–9. doi:10.2515/therapie/2014009.
44. Gissler M, Mohangoo AD, Blondel B, Chalmers J, Macfarlane A, Gaizauskiene A, et al. Perinatal health monitoring in Europe: results from the EURO-PERISTAT project. *Imform Health Soc Care*. 2010;35(2):64–79.
45. EUROMedCAT. <http://euromedicat.eu/whatiseuromedicat>. Accessed January 4 2017
46. de Jonge L, Garne E, Gini R, Jordan SE, Klungsoyr K, Loane M, et al. Improving information on maternal medication use by linking prescription data to congenital anomaly registers: a EUROMedCAT study. *Drug Saf*. 2015;38(11):1083–93. doi:10.1007/s40264-015-0321-9.
47. Wettermark B, Zoega H, Furu K, Korhonen M, Hallas J, Norgaard M, et al. The Nordic prescription databases as a resource for pharmacoepidemiological research—a literature review. *Pharmacoepidemiol Drug Saf*. 2013;22(7):691–9. doi:10.1002/pds.3457.
48. Schmidt MD, Freedson PS, Pekow P, Roberts D, Sternfeld B, Chasan-Taber L. Validation of the Kaiser Physical Activity Survey in pregnant women. *Med Sci Sports Exerc*. 2006;38(1):42–50.
49. Olsen J, Melbye M, Olsen SF, Sorensen TI, Aaby P, Andersen AM, et al. The Danish National Birth Cohort—its background, structure and aim. *Scand J Public Health*. 2001;29(4):300–7. doi:10.1177/14034948010290040201.
50. Magnus P. The Norwegian Mother and Child Cohort Study (MoBa)—new research possibilities. *Nor Epidemiol*. 2007;17(2):107–10.
51. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update: the Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol*. 2016;45(2):382–8. doi:10.1093/ije/dyw029.
52. Novik TS. Validity of the Child Behaviour Checklist in a Norwegian sample. *Eur Child Adolesc Psychiatry*. 1999;8(4):247–54. doi:10.1007/s007870050098.
53. Richter J, Janson H. A validation study of the Norwegian version of the Ages and Stages Questionnaires. *Acta Paediatr*. 2007;96(5):748–52. doi:10.1111/j.1651-2227.2007.00246.x.
54. van Gelder MM, Bretveld RW, Roukema J, Steenhoek M, van Drongelen J, Spaanderman ME, et al. Rationale and design of the PRegnancy and Infant DEvelopment (PRIDE) Study. *Paediatr Perinat Epidemiol*. 2013;27(1):34–43. doi:10.1111/ppe.12023.
55. Liew Z, Olsen J, Cui X, Ritz B, Arah OA. Bias from conditioning on live birth in pregnancy cohorts: an illustration based on neurodevelopment in children after prenatal exposure to organic pollutants. *Int J Epidemiol*. 2015;44(1):345–54. doi:10.1093/ije/dyu249.
56. Grzeskowiak LE, Gilbert AL, Morrison JL. Investigating outcomes associated with medication use during pregnancy: a review of methodological challenges and observational study designs. *Reprod Toxicol*. 2012;33(3):280–9. doi:10.1016/j.reprotox.2012.01.006.
57. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581–92.
58. Rubin DB. Multiple imputation for nonresponse in surveys. New York, NY: John Wiley & Sons; 1987.
59. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med*. 1991;10(4):585–98.
60. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. *J Am Stat Assoc*. 1994;89(427):846–66.
61. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health*. 2006;60(7):578–86.
62. Palmsten K, Huybrechts KF, Michels KB, Williams PL, Mogun H, Setoguchi S, et al. Antidepressant use and risk for preeclampsia. *Epidemiology*. 2013;24(5):682–91. doi:10.1097/EDE.0b013e31829e0aaa.
63. Wood ME, Frazier JA, Nordeng HM, Lapane KL. Prenatal triptan exposure and parent-reported early childhood neurodevelopmental outcomes: an application of propensity score calibration to adjust for unmeasured confounding by migraine severity. *Pharmacoepidemiol Drug Saf*. 2016;25(5):493–502. doi:10.1002/pds.3902.
64. MacLehose RF, Olshan AF, Herring AH, Honein MA, Shaw GM, Romitti PA. Bayesian methods for correcting misclassification: an example from birth defects epidemiology. *Epidemiology*. 2009;20(1):27–35. doi:10.1097/EDE.0b013e31818ab3b0.
65. van Gelder MM, Donders AR, Devine O, Roeleveld N, Reefhuis J. Using Bayesian models to assess the effects of under-reporting of cannabis use on the association with birth defects, national birth defects prevention study, 1997–2005. *Paediatr Perinat Epidemiol*. 2014;28(5):424–33. doi:10.1111/ppe.12140.
66. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615–25.
67. Nilsen RM, Vollset SE, Gjessing HK, Skjaerven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatr Perinat Epidemiol*. 2009;23(6):597–608. doi:10.1111/j.1365-3016.2009.01062.x.
68. Hansen C, Andrade SE, Freiman H, Dublin S, Haffenreffer K, Cooper WO, et al. Trimethoprim-sulfonamide use during the first trimester of pregnancy and the risk of congenital anomalies. *Pharmacoepidemiol Drug Saf*. 2016;25(2):170–8. doi:10.1002/pds.3919.
69. Andrade SE, McPhillips H, Loren D, Raebel MA, Lane K, Livingston J, et al. Antidepressant medication use and risk of persistent pulmonary hypertension of the newborn. *Pharmacoepidemiol Drug Saf*. 2009;18(3):246–52. doi:10.1002/pds.1710.
70. Kieler H, Artama M, Engeland A, Ericsson O, Furu K, Gissler M, et al. Selective serotonin reuptake inhibitors during pregnancy and risk of persistent pulmonary hypertension in the newborn: population based cohort study from the five Nordic countries. *BMJ*. 2012;344:d8012. doi:10.1136/bmj.d8012. **One of the first major pharmacoepidemiology studies that linked birth and prescription registry data within and between Nordic countries, allowing for a population-based study of relatively rare exposures and outcomes.**