Medical databases in studies of drug teratogenicity: methodological issues

Vera Ehrenstein\textsuperscript{1}  
Henrik T Sørensen\textsuperscript{1}  
Leiv S Bakketeig\textsuperscript{1,2}  
Lars Pedersen\textsuperscript{1}  
\textsuperscript{1}Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; \textsuperscript{2}Norwegian Institute of Public Health, Oslo, Norway

Abstract: More than half of all pregnant women take prescription medications, raising concerns about fetal safety. Medical databases routinely collecting data from large populations are potentially valuable resources for cohort studies addressing teratogenicity of drugs. These include electronic medical records, administrative databases, population health registries, and teratogenicity information services. Medical databases allow estimation of prevalences of birth defects with enhanced precision, but systematic error remains a potentially serious problem. In this review, we first provide a brief description of types of North American and European medical databases suitable for studying teratogenicity of drugs and then discuss manifestation of systematic errors in teratogenicity studies based on such databases. Selection bias stems primarily from the inability to ascertain all reproductive outcomes. Information bias (misclassification) may be caused by paucity of recorded clinical details or incomplete documentation of medication use. Confounding, particularly confounding by indication, can rarely be ruled out. Bias that either masks teratogenicity or creates false appearance thereof, may have adverse consequences for the health of the child and the mother. Biases should be quantified and their potential impact on the study results should be assessed. Both theory and software are available for such estimation. Provided that methodological problems are understood and effectively handled, computerized medical databases are a valuable source of data for studies of teratogenicity of drugs.

Keywords: databases, birth defects, epidemiologic methods, pharmacoepidemiology

Introduction
In Western countries, more than half of pregnant women take prescription medication, and nearly all pregnant women use over-the-counter medications, vitamins or other dietary supplements.\textsuperscript{1–4} Drugs that are safe for adults may be teratogenic for the developing fetus. The majority of drugs or their metabolites cross the placental barrier,\textsuperscript{5} and metabolites may be more fetotoxic than their source substances, as was noted in the case of thalidomide-induced phocomelias.\textsuperscript{6} Because pregnant women rarely participate in randomized studies of medicines, evidence from observational studies is central in establishing safety of prenatal drug exposure.\textsuperscript{7}

Since birth defects are rare, assembling cohorts to observe their occurrence is expensive in terms of time, money, and resources, leading to widespread use of the case-control design. Case-control studies, often based on interviews or questionnaires, are susceptible to selection and recall bias, and they do not allow estimation of absolute risks (prevalences) of birth defects. Existing medical databases are increasingly being used to conduct pharmacoepidemiologic cohort studies, including studies of drug teratogenicity.\textsuperscript{8} Medical databases, some of which have been in existence...
for decades,\textsuperscript{9,10} prospectively and routinely record health data, enabling relatively quick and inexpensive analyses of data from large populations.\textsuperscript{8} Such analyses allow direct estimation of birth defects’ prevalences, while estimates of association are less susceptible to selection or recall bias obtained in studies with primary data collection.\textsuperscript{11} Medical databases have their own limitations, however, which must be considered when interpreting findings of studies based on their data.\textsuperscript{5}

In this paper, we first provide examples of established North American and European medical databases suitable for studying teratogenicity of drugs, and then describe common epidemiologic biases as manifested in studies based on such databases.

**Medical databases for studying teratogenicity of drugs**

Medical databases are data repositories that contain health-related data, including electronic medical records, administrative databases such as claims records, and registries of diseases and rendered health services.\textsuperscript{5,12} Medical databases are typically maintained for surveillance or reimbursement, meaning that the influx of data into them is decoupled from research purposes. Therefore linkage of records from different databases covering the same population may be required in order to combine in the same dataset information on prenatal drug exposure, occurrence of birth defects, and relevant covariates. Independence of data collection from research hypotheses reduces risk of self-selection bias, but limits the variables available for analysis to those routinely collected by the databases.

In the United States, Medicaid, a health care plan for low-income persons, maintains claims databases of its enrollees. Low-income pregnant women, and children under the age of 6 years, are eligible for Medicaid coverage, which includes access to prescribed drugs.\textsuperscript{16} Each state administers its own Medicaid program, and patient eligibility and available services vary from state to state.\textsuperscript{17}

Private insurers in the US also maintain claims databases that compile data on maternal use of prescribed drugs and on birth defects as a part of an overall diagnostic record. For example, the health management organization (HMO) Research Network combines data maintained by several managed-care health plans (such as Kaiser Permanente or Harvard Pilgrim Health Care).\textsuperscript{18} The network is part of the national initiative under the aegis of the Department of Health and Human Services “to increase awareness of the benefits and risks of therapeutics”.\textsuperscript{18} Each participating health plan maintains computerized databases of member enrollment, filled prescriptions, and diagnoses made during outpatient visits and hospitalizations.\textsuperscript{15}

In Canada, the Saskatchewan Health Services Databases cover 99% of the population of the Saskatchewan province (about 1 million persons or 3.2% of the Canadian population). Residents of the province have universal health coverage, including prescription medication reimbursement. Drug teratogenicity may be studied by linkage of databases on vital statistics (live and still births), outpatient prescriptions, and hospitalization services. The linked data are available for research use, provided they remain unidentifiable to ensure data protection.\textsuperscript{19}

The United Kingdom’s Medicines and Healthcare products Regulatory Agency (MHRA) hosts the General Practice Research Database (GPRD), which is one of the largest and the most complete databases, containing medical records on more than 4 million patients, including prescribed medications, referrals and diagnoses made during hospitalizations and visits to general practitioners (GPs).\textsuperscript{13,14} The database was established in 1987, and its data are linkable to the UK’s other medical databases.\textsuperscript{13}

In the course of the past two decades, all Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden) have established prescription databases,\textsuperscript{21–23} tracking prescription medications dispensed in outpatient pharmacies. The earliest database (in Denmark) was established in 1989, while the newest (in Sweden) was launched in 2005 (for a recent review of Nordic prescription databases, see review by Furu and colleagues\textsuperscript{21}). In Sweden, maternal use of medication in pregnancy is also available, since 1995, as measured by self-reporting during the first antenatal visit.\textsuperscript{24}

Drawbacks of North American databases maintained by health insurers or based on residence include nonuniform eligibility; selective coverage (eg, the poor or the employed); and potential loss to follow-up if patients cease to be eligible for coverage after changes in income, employment, or residence. By contrast, medical databases in Nordic countries are derivatives of universal and uniform health coverage of welfare states.\textsuperscript{20} Thus, in contrast to the North American databases, membership in a Nordic medical database is independent of income, employment, or residence.\textsuperscript{21}

Sources of data on birth defects and other reproductive outcomes in Nordic countries include birth registries,\textsuperscript{3,24,25} hospital discharge registries,\textsuperscript{26} and registries of congenital malformations.\textsuperscript{24,26,27} Birth registries, with records dating back to the 1960s or 1970s,\textsuperscript{3,25} typically record birth defects discovered immediately after birth, and therefore must be
supplemented with data from hospital registries and registries of congenital malformations. Hospital registries, congenital malformation registries, or induced abortion registries can also be used to ascertain reproductive outcomes other than live or still births. Miscarriages, and elective or therapeutic pregnancy terminations, including those done after prenatal diagnosis, and some data on malformations also may be available. Data from medical databases can be linked to other registries containing demographic, social, and labor-market data. This is especially true for Denmark, whose network of population databases has been described as “the most complete and interwoven collection of statistics touching on almost every aspect of life”.

A crucial advantage of Nordic databases is the possibility of across-the-board data linkage via unique identification number, assigned at birth and encoding date of birth and sex, which follows each citizen “from cradle to grave”. Birth registry records contain the maternal identification number, which is a necessary link for unambiguous ascertainment from prescription databases of maternal drug intake during pregnancy.

Teratology information services (TIS) counsel newly pregnant women, or women who are trying to conceive, regarding safety of medication use. The European Network of Teratology Information Services lists about 25 European and South American TIS, and a similar number is listed by the US-based Organization of Teratology Information Specialists of the United States and Canada. TIS record the women’s demographic, obstetrical, medical, and drug-exposure history. During the year after the expected delivery, the TIS conducts a follow-up interview, collecting data on malformations. Reporting to TIS is initiated by women and is thus not systematic. Therefore, despite availability of large numbers of computerized records, TIS-based studies on teratogenicity of drugs are similar to epidemiologic studies with primary data collection in their susceptibility to self-referral bias, and nonrandom losses to follow-up. Furthermore, TIS may cover diverse geographic areas, making it difficult to establish a reference for an expected number of malformations in the source population.

Medical databases (summarized in Table 1) are widely used for addressing teratogenicity of drugs. Examples include use of the GPRD study on anticonvulsants; the Tennessee Medicaid study on angiotensin-converting enzyme inhibitors; a TIS-based study of prenatal loradine exposure; studies of antidepressants from US claims databases; the Saskatchewan Healthcare Databases; the population databases of Sweden, Denmark, and on TIS.

**Bias in studies of medical databases**

Large sample sizes, obtainable from medical databases, may reduce random error around the resulting estimates, but systematic error remains a problem. All three main types
of epidemiologic bias – selection bias, information bias, and confounding – occur in studies of medical databases.

Selection bias
In an ideal cohort study of drug teratogenicity, an investigator would recruit cohorts of women exposed and unexposed to a given agent before conception and examine birth defects detected among the fetuses throughout gestation, at birth, and several years postnatally. In such an ideal setting, the incidence rate of a birth defect is the number of all fetuses or neonates with that defect, detected at any time during the follow-up, divided by the total person-time contributed by all fetuses at risk. In reality, neither reproductive outcomes nor total person-time contributed by the initial conceptuses is fully observable (respectively, the numerator and the denominator of the incidence rate).

Two major sources of selection bias are relevant to the study of drug teratogenicity: spontaneous fetal loss (extrauterine pregnancy, miscarriage, stillbirth), and induced abortion. Malformations associated with first-trimester miscarriages are not always recorded or observable (Table 2).28 Elective pregnancy terminations during the first trimester are usually unrelated to medication use or suspected birth defects.41,42 Second-trimester induced abortions are commonly carried out after diagnosis of malformation by prenatal diagnosis. The proportion of malformations diagnosed prenatally varies geographically (eg, 25% in Croatia vs 88% in Paris)42 and according to local availability of relevant procedures.44 Furthermore, rates of second-trimester pregnancy terminations depend on local laws, severity of birth defect, and long-term prognosis.41 Up to 94% of fetuses with prenatally diagnosed fatal malformations (eg, anencephaly) are aborted, compared with 30% to 40% of fetuses with treatable malformations (eg, diaphragmatic hernia or transposition of great arteries).43 In Sweden, 60% of spina bifida cases were diagnosed at elective termination of pregnancy between 18 and 22 weeks’ gestation, and the level of ascertainment of spina bifida was inversely related to achieved gestational age at pregnancy end (birth or pregnancy termination).28 In contrast, diagnosis of cleft palate rarely led to pregnancy termination.28 In the US, between the 1970s and the 1990s, use of ultrasonography or amniocentesis for prenatal diagnosis has increased from 7% to nearly 90%, while the rate of elective abortions for any malformation increased from 0.8% to 18%, with a larger absolute increase among terminations for nonfatal malformations.45 Finally, access to and utilization of prenatal diagnosis may depend on a pregnant woman’s socioeconomic status, race/ethnicity, or age.46 Thus, severity of selection bias resulting from second-trimester pregnancy terminations may vary according to geography, type of malformation, timing of termination, calendar time, and maternal characteristics.

In summary, in cohort studies of birth defects, inability to observe birth defects at all reproductive outcomes represents loss to follow-up of a potentially nonrandom subgroup of embryos and fetuses. Selection bias ensues if either the medication or the malformation affects an embryo’s survival until malformation can be observed. Such bias can cause a spurious apparent association between drug exposure and medicinal agent or, alternatively, lead to erroneous conclusions about the lack of an association.47 To reduce selection bias, whenever possible, all observable reproductive outcomes should be ascertained as well as malformations detected both prenatally and at birth.48

Information bias
In database studies of drug teratogenicity, relying on dispensed prescription information to ascertain drug use in pregnancy may lead the investigator to erroneous assumptions regarding the fact, the timing, and the dosage of medication intake.49 Such misclassification is an important limitation given the short duration of gestation and even shorter duration of developmental “critical periods”, during which birth defects can plausibly occur as a result of drug exposure. A major
drawback of prescription registries is lack of data on adherence once medication is dispensed. Adherence may be indirectly measured by the number of filled prescriptions. Furthermore, medication dispensed during hospitalization or in outpatient clinics are not recorded in prescription registries potentially leading to under-ascertainment of medication use. Nondifferential misclassification of maternal drug exposure, if severe, may nullify the observed estimate of effect, if an effect exists. The direction of bias resulting from differential misclassification of maternal medication use is unpredictable.

Presence of a birth defect is also subject to misclassification. The proportion of true cases of birth defects captured by electronic sources (completeness, analogous to sensitivity) may vary widely by type of anomaly and type of data source. In the Saskatchewan Health Databases, for example, data in the hospitalization records may be suitable for studying only major birth defects. Imperfect sensitivity of birth defect measure leads to underestimation of true prevalence of birth defects, but imperfect sensitivity alone does not bias a relative estimate of effect. If no other bias is at work, relative estimate of effect will be unbiased in the absence of false-positive records of birth defects (100% specificity), which is usually the case for electronic records of birth defects.

In summary, data on medication use and occurrence of birth defects in medical databases are of varying quality, depending on method of data collection and on the type of medication and birth defect under study. A researcher embarking on a study of teratogenicity should obtain information about validity of data on the variables of interest in a selected data source.

Confounding
Predictors of medication use by a pregnant woman that are independent risk factors for a given birth defect can confound the estimate of association between the medication and the birth defect under study. Examples of potential confounding factors include geography, maternal age, race, socioeconomic status, and the disease for which the medication is prescribed.

Unmeasured or unknown confounding cannot be controlled in an analysis, except indirectly, if unmeasured traits happen to correlate with measured and controlled characteristics. Residual confounding, which can be viewed as a special case of unmeasured confounding, occurs when controlling for a variable used to measure a confounding factor does not completely remove confounding by that factor. This may occur when the variable is misclassified owing to poor measurement or inadequate categorization. The estimate of effect adjusted for a misclassified version of a confounder is biased in the direction of confounding. If adjusting for a misclassified confounder variable attenuates the crude estimate, adjustment for a perfectly measured confounder is expected to result in further attenuation, while amplification of effect estimate by a misclassified confounder variable indicates that the true effect may be larger than the apparent one.

Confounding by indication is common in studies of unintended effects of drugs, because of the difficulty in separating the effect of a given drug from the effect of the disease for which the drug is given (the indication). Thus, a maternal disease itself – rather than medication used to treat it – may increase risk of malformation in offspring. To counter confounding by indication, one may examine risks of birth defects among offspring of mothers taking the same medication prescribed for different indications and among offspring of women with similar indications taking different drugs. These methods may only partially address confounding by indication since use of different medications for the same indication may vary according to severity or etiology of disease, both of which may affect fetal risks. One way to address confounding by indication is by taking advantage of the time-sensitive nature of the relation between drug exposure and the possible birth defect. For example, causation between cardiac malformations and drug exposure cannot be inferred if the drug exposure occurred only during second and third trimester, ie, after the heart had been formed.

Selection bias, information bias, and confounding are all at work simultaneously in a given epidemiologic study, and may bias estimates in the opposing directions. It is therefore difficult to know the magnitude and direction of the net bias. Theory and software have been developed to quantify the impact of study estimates by unmeasured confounding and misclassification of study variables. The methods are based on subjecting study results to an “array of informed assumptions” about the source and the magnitude of systematic error. Many available methods tend to apply to simple situations, such as those characterized by dichotomous study variables. However, even rough quantification of bias is an improvement over sometimes insufficiently justified assertions and beliefs regarding its direction and impact.

Conclusion
With respect to teratogenicity of drugs, any effect – harmful, neutral, or protective – has important implications for pregnant women and their offspring. Bias masking a true teratogenic drug effect would result in continued use of a harmful agent, while bias creating false appearance of teratogenicity...
may limit treatment options available to pregnant women. These could include treatments for chronic conditions that may themselves detrimentally affect pregnancy outcome if left untreated.

Provided that methodological problems are understood and effectively handled, computerized health care databases are a valuable source of data for cohort studies of teratogenicity of drugs.

**Disclosures**

The authors report no conflicts of interest in this work.

**References**


60. Fox MP, Lash TL, Silliman RA. A sensitivity analysis to separate bias due to confounding from bias due to predicting misclassification by a variable that does both. *Epidemiology*. 2000;11(5):544–549.