Reviewer Guidance

Evaluating the Risks of Drug Exposure in Human Pregnancies

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

April 2005
Clinical/Medical

Reviewer Guidance
Evaluating the Risks of Drug Exposure in Human Pregnancies

Additional copies are available from:

Office of Training and Communications
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

or

Office of Communication, Training, and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
(Tel) 800-835-4709 or 301-827-1800
http://www.fda.gov/cber/guidelines.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

April 2005
Clinical/Medical

TABLE OF CONTENTS

Contains Nonbinding Recommendations

Reviewer Guidance[1] Evaluating the Risks of Drug Exposure in Human Pregnancies

This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

This guidance is intended to help FDA staff evaluate human fetal outcome data generated after medical product exposures during pregnancy.[2] The goal of such evaluations is to assist in the development of product labeling that is useful to medical care providers when they care for patients who are pregnant or planning pregnancy. The review of human pregnancy drug exposure data and assessment of fetal risk (or lack of risk) requires consideration of human embryology and teratology, pharmacology, obstetrics, and epidemiology. Consequently, FDA staff also are encouraged to consult with experts in these fields, as appropriate.

This guidance does not address the assessment of experimental animal reproductive toxicology data. A separate guidance is under development for FDA pharmacology/toxicology reviewers that describes a process using nonclinical data to estimate human developmental and reproductive risks from drug exposure when human data are unavailable.[3]

[1] This guidance has been prepared by the Pregnancy Outcomes Working Group of the FDA Pregnancy Labeling Taskforce in the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

[2] Throughout this document, the terms *medical product* and *drug* include drug and biological products including vaccines.

[3] In November 2001, the FDA issued a draft reviewer guidance on *Integration of Study Results to Assess Concerns about Human Reproductive and Developmental Toxicities*. When finalized, it will represent the Agency's thinking on this issue. (http://www.fda.gov/cder/guidance/index.htm)

Contains Nonbinding Recommendations

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

The term *teratogen* is used to denote the result of a hazard assessment on a particular agent (for purposes of this guidance, a *drug*). The use of the term indicates that the drug has the capacity under certain exposure conditions to produce abnormal development in an embryo or fetus. However, hazard assessment must be put into context. Whether a drug causes abnormal development or not depends not only on the physical and chemical nature of the drug but also on the dose, duration, frequency, route of exposure, and gestational timing involved.

Other factors that may determine whether a particular exposure is likely to be teratogenic in a particular instance include other concurrent exposures and the biological susceptibility of the mother and the embryo or fetus. Throughout this guidance, the term *teratogen* is used to designate products with *teratogenic potential at clinical doses used in humans*. Classifying a drug as a teratogen only indicates that it ***may*** have the potential for producing developmental toxicity given the appropriate conditions. For example, a single 50-mg dose of thalidomide administered on the 26th day after conception has a significant risk of producing a major structural malformation of the embryo. However, that same dose of thalidomide taken in the 10th week after conception does not produce structural malformations, and a 1-mg dose at any time during pregnancy has no observable effect on the developing embryo or fetus (Brent 2001).

About 4 percent (1/28) of babies are born each year with a major birth defect or congenital malformation (March of Dimes 2001). The March of Dimes defines a major birth defect as an abnormality of structure, function, or metabolism that either is fatal or that is present at birth and results in physical or mental disability (March of Dimes 2001). For the majority of major birth defects (about 65 percent), the etiology is unknown (Schardein 2000). Chemically induced birth defects, including those associated with drug exposure, probably account for less than 1 percent of all birth defects; few drugs are proven human teratogens at clinical doses (Koren 1998). Of the thousands of drugs available, only about 20 drugs or groups of drugs (most being anticonvulsants, antineoplastics, or retinoids) are recognized as having an increased risk of developmental abnormalities when used clinically in humans (Schardein 2000). However, since few drugs have been systematically studied to identify their full range of possible teratogenic risks, we cannot assume that current knowledge is complete. The identification of a drug's teratogenic potential is important because drug-induced adverse fetal effects are potentially preventable.

There appears to be a general perception that the use of any drug at any time during pregnancy can harm the developing embryo or fetus. In one study, pregnant women exposed to certain nonteratogenic drugs believed that the risk for major birth defects was 24 percent, which is similar to the known risk from exposure to thalidomide during the gestational time of most sensitivity (Koren 1989). Another study (done in Canada) reported that even health professionals who were shown drug labeling with "reassuring" text (i.e., labeling that explicitly stated that the drug itself does not cause fetal malformations when used in pregnancy) still rated those drugs as having some risk for causing birth defects (Pole 2000). This exaggerated fear could lead to termination of a wanted pregnancy or to unnecessary withholding of needed drug therapy during pregnancy.

Although knowledge of teratogenic potential is a critical part of a drug's benefit/risk profile, pregnant women are rarely included in clinical trials (Mastroanni 1994). There may be inadvertent pregnancy exposures during clinical trials of new products, but available data are usually insufficient to permit an adequately powered statistical analysis. Consequently, when a drug is first marketed there are usually no human data on the effects of in utero drug exposure. The only data on fetal effects initially available in the product labeling usually comes from animal reproductive toxicology studies.

Despite the lack of information on the safety of drug use during pregnancy, most pregnant woman likely will be exposed to drugs. Fetal exposure can occur before a woman knows she is pregnant. Some women enter pregnancy with medical conditions that require continuing drug therapy. New medical problems may develop during, or old ones may be exacerbated by, pregnancy.

Because little is known before marketing about a drug's teratogenic potential, postmarketing surveillance of drug use in pregnancy is critical to the detection of drug-induced fetal effects. With current postmarketing surveillance methods, this process can take considerable time. For example, with thalidomide, it took almost 4 years before the link between thalidomide use during pregnancy and phocomelia was recognized (Lenz 1961, McBride 1961). The recognition of the more subtle pattern of malformations associated with warfarin use during pregnancy took more than 20 years (Pettifor 1975). Some drugs may induce teratogenic effects that are not clinically evident until many years after birth, such as the reproductive tract abnormalities associated with DES exposures in utero (Herbst 1975).

It is important that the FDA and sponsors routinely review all available data on drug exposure during pregnancy and work together to provide up-to-date product labeling that reflects what is known and not known about human fetal risk or lack of risk.

Contains Nonbinding Recommendations

III. CRITICAL FACTORS IN EVALUATING THE EFFECTS OF DRUG EXPOSURE IN HUMAN PREGNANCIES

The following is a list of factors to consider when presented with human pregnancy data and faced with making a determination whether and how the data should be included in the product labeling. A discussion of each factor follows.

- Background prevalence of adverse pregnancy outcomes
- Combined vs. individual rates of birth defects
- Major vs. minor birth defects
- Timing of exposure
- Intensity of exposure
- Variability of response
- Class effects

A. Background Prevalence of Adverse Pregnancy Outcomes

Every pregnancy has a risk of an abnormal outcome regardless of drug exposure. Reproductive toxicologists generally consider the four major manifestations of abnormal fetal development to be growth alteration, functional deficit, structural malformation, and death (Schardein 2000). The purpose of collecting and evaluating data on drug exposure during pregnancy is to address whether a particular drug exposure ***increases*** the risk of abnormal fetal development above the background rate.

Clinical classifications of pregnancy outcomes include live births, elective terminations, and fetal losses, i.e., spontaneous abortions (loss prior to 20 weeks post-conception), and fetal deaths/stillbirths (loss beyond 20 weeks post-conception) (Ventura 2000). Based on national data for 1996 (the most recent year for which data are available) only 62 percent of clinically recognized pregnancies result in a live birth; 22 percent end in elective termination and 16 percent result in spontaneous abortions (1 of 7 known pregnancies) or fetal death/stillbirth (1 of 200 known pregnancies) (Ventura 2000, March of Dimes 2001). Among live births, preterm birth (before the 35th week after conception) and low birth weight (<2500 grams) are also considered adverse pregnancy outcomes occurring in 1 of 8 and 1 of 12 live births, respectively (March of Dimes 2001).

Although the March of Dimes reports an overall background birth defect prevalence of 4 percent of births, this percentage can change depending upon the definition of birth defect (e.g., inclusion of abnormalities of cosmetic importance), the specific population studied, and the time period beyond birth used for detection. The term *birth defects*, as used by the March of Dimes, refers to *congenital anomalies* identified by codes 740-759 of the Ninth Revision of the International Classification of Diseases (ICD-9) (Petrini 1997). The rate of some birth defects will be influenced by whether prenatally diagnosed defects and/or elective terminations are included, and whether prenatal defects are confirmed postnatally. For example, a large percentage of pregnancies with anencephaly are terminated and will not be captured in birth statistics other

than as a termination. Prenatally diagnosed hydronephrosis and ventricular septal defects may resolve before delivery or soon afterward and will be overcounted if included only on the basis of prenatal diagnosis. Unfortunately, prenatally diagnosed defects may not be evaluated postnatally at all if the outcome is a stillbirth or elective termination. This not only precludes confirmation of the diagnosis, but also the ability to ascertain additional less obvious conditions.

To assess whether the therapeutic use of a drug results in an increase in adverse fetal effects, one commonly compares the rate of birth defects seen in a study population to some background rate. However, all-inclusive background rates of all major birth defects may include infants with genetic syndromes and chromosomal abnormalities that may not be caused by a medication or other toxic exposure. Therefore, at best such rates constitute very crude comparisons.

Population estimates can vary considerably by maternal age, race, geographic region, socioeconomic status, and time period (e.g., the availability of diagnostic tools can change over time). In addition, maternal disease states (e.g., diabetes) can influence pregnancy outcomes and fetal development with outcome rates within a particular disease population being very different from rates for the general population. Therefore, when evaluating whether exposure to a drug during pregnancy increases the risk for any adverse pregnancy or fetal outcome, it helps to know the background rate of the outcome in a population as similar as possible to the study population.

It is easier to detect an increased risk for an abnormal outcome that occurs at a relatively high background rate (e.g., spontaneous abortion) than it is to detect an increased risk for an abnormal outcome that occurs at a relatively low background rate (e.g., oral clefts). Consequently, the statistical power of a study to detect an increased risk of adverse fetal effects that normally occur at a low rate is often limited. When collecting pregnancy exposure data on a previously unidentified teratogen that behaves like isotretinoin in causing serious, obvious effects in about 25 percent of exposed live-born infants, it will not take many exposed pregnancies before the problem is detected. However, a teratogen that behaves like valproic acid, which causes neural tube defects at a rate of 1 to 2 percent in exposed infants (a tenfold increase over background), requires a much larger number of exposed pregnancies to detect. One report of 12 normal infants following gestational exposure to valproic acid (Hiilesma 1980) was misleading because the number of exposures was not sufficient to detect an effect size of 1 to 2 percent. It took a retrospective case control study to demonstrate the association (Bjerkedal 1982, Robert 1982).

When reviewing studies or case series, a reviewer should consider whether there are enough exposures to demonstrate an increase in risk if such a risk exists. Any studies reporting no increase in the background rate of birth defects in exposed pregnancies can be viewed with skepticism unless the power of study to detect or rule out a stated level of risk is also included (Ferencz 2000).

B. Combined vs. Individual Rates of Birth Defects

Human teratogens generally increase rates of specific defects or a spectrum of defects. For example, thalidomide causes limb, spine, and central nervous system (CNS) defects; isotretinoin causes ear, CNS, and cardiac defects; valproic acid causes neural tube defects; warfarin causes cartilage defects; and angiotensin II converting enzyme (ACE) inhibitors cause renal functional effects (Mitchell 2000).

In addition to evaluating the overall rate of birth defects in the study population, it is also important to look at rates of individual birth defects. A population that experiences a tenfold increase in a specific rare birth defect (e.g., spina bifida from 0.04 percent to 0.4 percent) as a result of exposure to a teratogen may still have a total birth defect rate (i.e., all malformations) that is not measurably different from that in a reference population.

It has been suggested that the possibility of detecting a teratogen can be increased by grouping individual malformations according to an understanding of their embryologic tissue of origin (Mitchell 2000, Scheuerle 2002). Mitchell uses the example where interference with the normal development of the neural crest would lead to malformations of tissues derived from neural crest cells which, in the earliest stages of embryogenesis, migrate to form a variety of structures including those of the face/ears, parts of the heart, and the neural tube. A case in point is isotretinoin which interferes with neural crest cell migration/development and leads to specific malformations of the ear, heart, and neural tube (Mitchell 2000).

For general information on background population rates, refer to Table 1, which lists the leading categories of birth defects according to the March of Dimes, whereas Table 2 lists the range of reported state-level rates for some individual birth defects as reported to the National Birth Defects Prevention Network.

Table 1. Leading Categories of Birth Defects

Birth Defect	Estimated Prevalence
Heart and circulation	1 in 115 births
Muscles and skeleton	1 in 130 births
Genital and urinary tract	1 in 135 births
Nervous system and eye	1 in 235 births
Respiratory tract	1 in 900 births
Metabolic disorders	1 in 3,500 births

Source: March of Dimes, National Perinatal Statistics (http://www.modimes.com)

Table 2. Birth Defect Surveillance Data from Selected States*

Birth Defect	Range in Rates per 10,000 Live Births
Anencephalus	0.00 – 4.92
Anophthalmia/microphthalmia	0.39 – 3.73
Anotia/microtia	0.19 – 6.43
Atrial septal defect	10.51 – 70.29
Cleft lip with and without cleft palate	4.48 – 22.98
Cleft palate without cleft lip	3.52 – 9.35
Endocardial cushion defect	1.11 – 6.84
Esophageal ataresia/tracheosophageal fistula	0.83 – 5.20
Gastroschisis	0.90 – 6.59
Hydrocephalus without spina bifida	0.59 – 19.34
Hypoplastic left heart syndrome	0.51 – 3.94
Hypospadias and epispadias	1.59 – 46.61
Microcephalus	0.51 – 15.65
Obstructive genitourinary defect	3.43 – 35.26
Omphalocele	0.45 – 3.48
Pulmonary valve atresia/stenosis	1.17 – 20.71
Pyloric stenosis	0.18 – 30.70
Rectal and large intestinal atresia/senosis	1.02 – 9.02
Reduction deformity: lower limbs	0.74 – 5.35
Reduction deformity: upper limbs	0.75 – 5.02
Spina bifida without anencephalus	1.36 – 8.08
Tetralogy of Fallot	1.49 – 8.08
Transposition of great vessels	1.02 – 6.68
Ventricular septal defect	8.41 – 79.18

*Rates represent data contributed to the National Birth Defects Prevention Network from 30 states. Rates for most of the states were based on 5 years of data from 1995 – 1999. However, four states provided data for only 1 to 3 years. Not all states reported data for all defects. State birth defect surveillance programs vary considerably in their methodologies, definitions, and inclusion criteria. Because of these differences, the data cannot be combined to give overall rates for the United States; therefore data are presented as a range of rates.

Source: National Birth Defects Prevention Network, 2002, "Birth Defects Surveillance Data from Selected States, 1995 – 1999," *Teratology*, 66:S129-S211.

C. Major vs. Minor Birth Defects

Most case reports and studies of birth defects focus on *major birth defects* (i.e., those incompatible with life or requiring medical/surgical intervention). The term *minor birth defect* generally refers to minor physical anomalies with less clinical importance that represent deviations from what is considered *normal* and do not have obvious medical, surgical, or

cosmetic consequences. Minor birth defects are more common than major ones, occurring in 14 to 40 percent of live births (Leippig 1987). Clearly, what constitutes a minor birth defect may be subjective, hence the wide range of prevalence estimates.

Minor birth defects may have predictive value in identifying more serious associated problems. Estimates of the predictive value of three or more minor birth defects for an associated major defect range from 19.6 to 90 percent (compared to 1.3 to 2.4 percent of those with no minor birth defects) (Leippig 1987). Identification of a higher-than-expected frequency of infants with three or more minor defects, and especially a specific pattern of these defects, may be a more sensitive screen for possible human teratogens than evaluation of the increased risk of any major birth defect (Chambers 2001). Assessing minor birth defects in children who have a major birth defect is also important because a complete description of all abnormalities can help identify a pattern characteristic of the exposure and improve the ability to attribute causality. A pattern of minor defects may represent the mild end of a spectrum of effects with a major defect, occurring less frequently, at the other end of the spectrum. For example, the spectrum of effects resulting from in utero exposure to carbamazepine ranges from 11 percent of exposed infants with minor craniofacial defects and nail hypoplasia to 1 percent of exposed infants with a neural tube defect (Jones 1989, Hernandez-Diaz 2001).

D. Timing of Exposure

When interpreting data on drug exposure during pregnancy, it is important to consider the timing and duration of exposure and their relationship to windows of developmental sensitivity. Agents that produce adverse effects on the fetus typically do so during discrete sensitive periods of fetal development that vary depending on the particular teratogenic process and target organ. Each part, tissue, and organ of an embryo has a critical period during which its development may be disrupted (see Attachment A). For example, the most critical period for brain development is from 3 to 16 weeks post-conception, but its development may be disrupted after this because the brain is differentiating and growing rapidly (Moore 1998).

When evaluating pregnancy outcome data, it is important to identify the frame of reference for the reported gestational age. Determining the gestational week of exposure based on the date of last menstrual period versus the date of conception can produce a 2-week time difference that can be critical when evaluating an association between a birth defect and drug exposure. In this guidance, when gestational age is mentioned, it refers to time since conception.

During the first 2 weeks after conception the developing embryo is not susceptible to teratogenesis (Moore 1998). Drug exposures during this time period are not known to cause congenital anomalies in human embryos; however, such exposures may interfere with cleavage of the zygote or implantation of the blastocyst and/or cause early death and spontaneous abortion of the embryo (Moore 1998).

In humans, the embryo is most easily disrupted during organogenesis (3 to 8 weeks post-conception) when the tissue and organs are forming. During this time, teratogenic agents may induce gross structural abnormalities readily seen at birth. However, although more common with teratogenic exposures during the main embryonic period, the production of important anatomic defects is not limited to organogenesis, as evidenced by the microcephaly seen with maternal alcohol abuse during the fetal period (Moore 1998).

Later in gestation, the fetus rapidly grows and matures, undergoing active cell growth, differentiation, and migration, particularly in the CNS. Exposures during this period may cause physiologic defects such as minor anomalies of the external ear, growth retardation or functional disorders such as mental retardation. However, important abnormalities can also be produced by late pregnancy exposures such as the fetal alcohol syndrome seen with alcohol abuse, the renal function effects seen with the use of ACE inhibitors, and the cartilage defects seen with the use of warfarin.

Knowledge of the sensitive period for human target organ development facilitates optimal data interpretation. For example, if drug exposure occurred after the critical period of development for an organ, the exposure is an unlikely cause of the organ malformation (e.g., an infant born with transposition of the great vessels that are formed during the first trimester, who was exposed to the drug only in the third trimester).

Evaluating the timing of exposure is also important when assessing the power of a study. For example, consider a drug that causes a tenfold increase in neural tube closure defects, from 0.1 percent to 1 percent. Formation of the neural tube begins about day 18 after conception and, with normal development, the neural tube closes by the end of the fourth week of gestation (Moore 1998). A hypothetical study identifies 100,000 women exposed during the first trimester of pregnancy, but only 1,000 of the women were exposed during the sensitive period. The 1,000 pregnancies with exposure during the sensitive period produce 10 cases of neural tube defects based on a 1 percent rate while 99 cases are seen in the other 99,000 pregnancies based on the background rate of 0.1 percent. The total of 109 affected children from 100,000 exposed pregnancies produces a 0.11 percent rate, which probably would not be appreciated as different from the background rate of 0.1 percent and would not identify the real increase due to the exposure to the drug.

Table 3 lists the sensitive periods for exposure to some known teratogens. However, as a practical matter, the sensitive period for exposure to a drug, if there is one, is usually unknown. In situations where no clear toxicity has been identified, it is common to globally assess risk from first trimester exposures because that is the time of organogenesis. There are two potential sources of error in using this global approach. First, as seen in the previous example, sensitive time periods for a particular problem may make up a small portion of the first trimester. Therefore, if numbers allow, it is recommended that exposures during the first trimester be analyzed by gestational week post-conception of fetal development. Second, drug-induced fetal toxicities may not be limited to the first trimester or may produce abnormalities during more than

one exposure window. If studies exclude second and third trimester exposures, they will not be able to identify potentially important adverse effects that occur later in pregnancy, such as those seen with the ACE inhibitors.

Table 3. Examples of Critical Timing of Exposure for Some Known Teratogens

Teratogen	Critical Timing of Exposure
Thalidomide	Exposure between days 24 to 36 post-conception can produce limb and other defects (Moore 1998).
Diethylstilbestrol (DES)	Exposure before the 9th week post-conception leads to a precancerous vaginal adenosis in 73 percent of female offspring, but in only 7 percent of female offspring exposed after the 17th week post-conception. Clear-cell carcinoma has not been reported in female offspring who were exposed in utero after the 18th week post-conception (Herbst 1975).
ACE inhibitors	Exposure in the 2nd and 3rd trimester of pregnancy is associated with fetal hypotension, renal tubular dysplasia, anuria-oligohydramnios, growth restriction, hypocalvaria, and death (Sedman 1995).
Warfarin	Exposure in the latter half of the 1st trimester (6 to 12 weeks post-conception) produces the greatest susceptibility to skeletal features of fetal warfarin syndrome (Scialli 1995).

E. Intensity of Exposure

Dosing, including frequency and duration of exposure, is also an important consideration. Typically, a drug must cross the placenta and reach the fetus in sufficient concentration to cause an effect. Most nonprotein agents do cross the placenta, the exceptions being highly charged agents or certain very large molecules like heparin and insulin. Maternal changes during pregnancy in absorption, volume of distribution, metabolism, plasma protein binding, and excretion also will affect the extent to which the fetus is exposed. These parameters are dynamic over the course of pregnancy. For example, some products may be more readily transported across the placenta during late gestation than during early gestation because of an increased unbound fraction in maternal circulation, increased utero-placental blood flow, increased placental surface area, or changes in fetal circulation. Agents that undergo relatively little metabolism in the fetus, but are excreted into the amniotic fluid by the well-developed fetal kidney in the third trimester, may have greater exposure as the fetus continually swallows amniotic fluid.

All teratogens have a threshold below which adverse effects do not occur. Conversely, almost all exposures can be toxic to the fetus if the dose is high enough, even if only indirectly through maternal toxicity (e.g., low birth weight or mortality because the mother is too sick from the medication to eat).

Products that interfere with fetal development may produce manifestations along a continuum of response that may be closely connected with a dose-response relationship. A dose-response curve is usually seen in animal studies, but is rarely seen in humans because of the relatively narrow range of standardized clinical doses. An exception is the dose-response effect in humans that can be seen with alcohol. Alcohol at low doses throughout pregnancy is associated with slightly decreased birth weight. At high doses, it has effects on fetal neurologic development, and at progressively higher doses, it is associated with microcephaly and other visible anatomic effects (Ernhart 1987).

F. Variability of Response

People differ in their responses to specific medications, which may at least partially be due to genotypically determined differences in metabolism or receptors. Just as therapeutic or adverse effects related to a given drug do not occur in all exposed individuals, exposures during a sensitive time period known to increase the incidence of adverse pregnancy outcomes may do so only in a fraction of those infants exposed. For example, more than half of infants with similar in utero exposure to phenytoin are unaffected, about one-third show some congenital anomalies, and only 5 to 10 percent develop fetal hydantoin syndrome (Moore 1998).

Although the effects of known teratogens are generally predictable from a population perspective, the nature and extent of effects are not necessarily possible to predict in individual patients under similar conditions. The teratogenicity of an exposure can be influenced by both the maternal and fetal genotypes, which may result in differences in cell sensitivity, placental transport, metabolism, receptor binding, and distribution (Polifka 1999). Even at the same dose in the same gestational window, there can be a range of possible outcomes. Because of this variability, assessment of a drug's potential teratogenesis ought to consider the full range of birth defects.

It is important to remember that the concept of variability extends not only to toxic responses, but also to baseline attributes of populations. Birth weight, for example, varies by race and sex. Normative growth curves developed in Caucasian populations that are used to evaluate African-American or Asian-American babies may result in over-diagnosis of growth impairment. Genetic differences in metabolism of a drug by the mother, the placenta, and the fetus may contribute to variation in how much of a drug's metabolites reach fetal tissue and thereby lead to variable rates and types of toxicity manifested.

G. Class Effects

Understanding the structure/activity relationships and pharmacological mode of action of a class of therapeutic agents in some circumstances can provide a prediction of the possible safety and efficacy of a new agent. However, such knowledge is generally not predictive of human teratogenesis (Mitchell 2000). For example, thalidomide and glutethimide are closely related by chemical structure, but there is no evidence that glutethimide is teratogenic (Heinonen 1977).

While the introduction of a new product from a class of drugs with known human teratogenicity will solicit heightened scrutiny, it cannot be assumed that the product will also be teratogenic. Similar findings in the animal studies for the new product compared to the class would be cause for more concern, whereas clean animal data would lessen the concern.

IV. SOURCES OF HUMAN DATA ON GESTATIONAL DRUG EXPOSURES

Information on human gestational drug exposures will emerge during the postmarketing phase for virtually all drug products. Human pregnancy outcome data will be sent to the Agency either directly by voluntary reporters or via the sponsors as required by federal regulations.[4] The data will come from a variety of sources. For the most part, data will not be derived from controlled clinical trials, but from observational studies.

No single methodology can delineate the complete spectrum of adverse outcomes associated with prenatal exposure to a drug. Therefore, it is important to consider information from all available postmarketing surveillance sources to optimize detection and characterization of the reproductive effects of prenatal drug exposure.

---

[4] 21 CFR 310.305, 314.80, 314.98, 600.80, and 600.81 require sponsors to report adverse drug experiences from all sources (e.g., spontaneous reports, literature, studies). All congenital anomalies are considered serious (§ 314.80(a) and § 600.80(a)) and must be reported to the Agency within 15 days if unexpected (§ 314.80(c)(1)(i) and § 600.80(c)(1)(i)).

The ICH guidance for industry *E2C Clinical Safety Data Management: Periodic Safety Update Reports for Marketed Drugs* provides recommendations on what to include in a periodic report. It specifically lists "positive or negative experiences during pregnancy or lactation" as one safety issue to be explicitly addressed in the Overall Safety Evaluation section of the Periodic Safety Update Report (PSUR). On March 14, 2003, the FDA published a proposed rule that would incorporate these requirements into the FDA's postmarketing safety reporting regulations (68 FR 12406).

21 CFR 314.81(b)(2)(vii), 314.98(c), and 601.70 require a sponsor to submit in an annual report the status of certain types of postmarketing studies, including clinical safety studies, whether the FDA required the study or the applicant committed to it in writing. This requirement would include any epidemiologic study directed to the safety of drug use during pregnancy (e.g., pregnancy exposure registries) to which the applicant committed in writing.

Brief descriptions of the most common types of data and study designs follow. More in-depth information on the strengths and weaknesses of each can be found in pharmacoepidemiology texts elsewhere (Hartzema 1998, Strom 2000).

A. Case Reports

Case reports describe one or a series of clinical observations on drug exposure during pregnancy and subsequent infant effects. Case reports are the most common source of reported congenital anomalies, but can also be the most difficult to interpret. It is true that most pregnancy exposures that produce an increase in developmental risk in humans were first suggested by case reports (Schardein 2000). However, case reports cannot distinguish coincidence from causation and cannot be used to assess teratogenic risk. A series of case reports associating an unusual outcome with a drug exposure might well raise suspicions, as in the case of thalidomide, but follow-up evaluations are always necessary to assess risk. Unlike the case of thalidomide, most case reports that call attention to a putative association are not confirmed by follow-up evaluation. Case reports were the source for allegations regarding teratogenicity with Bendectin, but those allegations were later disproved through epidemiologic studies (Goldberg 1986). Conversely, case reporting of a small number of normal outcomes can delay the identification of an association, as was the case with valproic acid and spina bifida, where a published report of 12 normal pregnancies was too small to detect the now known 1 percent rate of affected children (Scialli 1992).

It is critical to be cautious and objective when evaluating isolated case reports because adverse outcomes tend to be disproportionately reported. The birth of an infant with a birth defect can be devastating. In an effort to understand what went wrong, health care providers and their patients may consider drug and other exposures during pregnancy. In contrast when a normal infant is born, there is no incentive to recall, much less report to the FDA or in the literature, any drug exposures during pregnancy.

Although an individual case report, by itself, can never prove causality, a series of similar reports of a distinct abnormality or group of abnormalities can establish a strong association or signal the need for further research. Most signals based on case reports will need to be further investigated using other pharmacoepidemiologic studies. Attachment B is a checklist that may be useful when evaluating case reports.

B. Epidemiology Studies

Formal epidemiology studies provide the best means of evaluating whether a gestational exposure adversely affects the developing infant. Epidemiology studies can identify associations between a given drug exposure and abnormalities in the newborn, and they can quantify the strength of such associations. Also, although it is impossible to demonstrate absolute safety, epidemiology studies can provide some measure of reassurance if risks are not found to be elevated, and the level of reassurance, like evidence of risk, can be quantified. The degree of reassurance is a function of the sample size (power) of the study. A study may report a lack of

association between a drug and birth defects simply because the study was insufficiently powered due to a small sample size to detect anything other than a very large difference. Thus, results from any pharmacoepidemiology study of birth defects must take into account what level of risk the study was capable of detecting.

In addition to a study's power, possible confounding by maternal indication is an important consideration. For example, birth defects are known to occur with increased frequency in infants born to women with certain medical conditions (e.g., diabetes), independent of the product used to treat the disease. With other diseases, it may not be known whether it is the disease, the drug, or an interaction between the severity of disease and the drug exposure that causes an observed increased risk of birth defects. For example, after years of speculation, it was only recently that a study suggested that the increased risk for embryopathy seen in the offspring of women with epilepsy is associated with gestational exposure to anticonvulsants rather than with epilepsy itself (Holmes 2001).

Other maternal attributes such as age, race, weight, parity, geographic location, and socioeconomic status can cause confounding as well. For example, children born to young mothers have a higher risk of gastroschisis (Nichols 1997), and Hispanics have a higher risk of spina bifida (Lary 1996).

Attachment C contains a checklist that may be useful when evaluating epidemiology studies.

1. Prospective Studies

a. Cohort studies

Cohort studies enroll a group of pregnant women before pregnancy outcome is known and collect information periodically throughout pregnancy on personal demographics, a variety of exposures, including drugs, and potential confounders. Their offspring are examined at birth and followed for a set amount of time. An example is the U.S. Collaborative Perinatal Project (CPP) that enrolled about 52,000 pregnant women between 1959 and 1965, collected detailed information on their pregnancies, and followed the children until age seven (Heinonen 1977). Another pregnancy cohort study, The National Children's Study, currently under development by NIH,[5] plans to study 100,000 children by enrolling pregnant women and following their offspring through 21 years of age. The study will periodically query the pregnant women and then the children about certain exposures, including therapeutic drug use.

The strength of the cohort design is the prospective, systematic collection of data, including exposures, confounders, and outcome information. If the population is large, it allows the study of multiple exposures and multiple outcomes. However, one major weakness of these studies is that generally small numbers of specific birth defects will be seen even in a large study population. This, coupled with the small number of women exposed to specific drugs within the

[5] http://nationalchildrensstudy.gov.

cohort, makes the detection of even a substantial drug-induced increase in the rate of any individual defect very difficult if not impossible.

b. Pregnancy exposure registries

Pregnancy exposure registries, a type of cohort study, prospectively enroll pregnant women exposed to one or more specific drugs of interest (unexposed pregnant women are sometimes included as well) and evaluate the associated pregnancy outcomes. Structured interviews or questionnaires that gather data on drug exposures and other confounders are administered periodically through pregnancy and often up until the infant is 1 year old. Some registries also include an evaluation of all live born infants by a dysmorphologist (Chambers 2001). Adverse pregnancy outcome rates in the exposed cohort are compared to rates in a non-exposed group drawn either from the study itself or from population data.

This design is the most efficient way to collect information on exposures with newly marketed drugs and potentially allows for evaluation of a range of adverse outcomes including patterns or a spectrum of malformations and functional deficits. However, this type of study is likely to undercount spontaneous abortions that occur because a pregnancy registry would never be able to detect drug-induced fetal losses that occur before pregnancy is known.

The length of time an infant is monitored and the source of information about the infant can influence the number of defects detected. One study reported that registries that limited infant data collection to information from the maternal health care provider on the infant's status at birth were less likely to ascertain either multiple defects per case or internal and serious defects, particularly genitourinary defects (Honein 1999).

The choice of an appropriate comparison group is a challenging aspect of pregnancy exposure registries. Comparison to the general population background rate of birth defects, although useful as a first screen, does not take into account the influence of relevant maternal attributes that may be related to the occurrence of birth defects, such as disease state, age, weight, use of alcohol or tobacco, or socioeconomic status. The ideal comparison group would be provided through concurrent enrollment of a group of pregnant women similar in all regards (including the disease leading to drug exposure) except for the specific drug exposure, but this is very difficult to implement.

Registries are often limited by self-referral bias (where women who enroll in the registry may be more or less likely to have malformed infants). Losses to follow-up can also limit the interpretability of registry data. Losses may be related to whether the infant is normal or not, particularly if the mother is not the primary provider of information to the study.

Pregnancy exposure registries tend to be too small or limited to resolve concerns about a drug increasing the risk of specific birth defects; however, they can potentially detect teratogens that, like thalidomide or isotretinoin, are so powerful that they produce defects in a relatively high

proportion of exposed pregnancies. Thus, a cohort of only a few dozen or 100 women exposed to such a drug would be sufficiently large to identify major teratogens. This ability is demonstrated with isotretinoin where 36 women with first trimester exposures were identified prospectively and followed through their pregnancies. Of the 28 live-born infants, five (18 percent) were malformed (Lammer 1985). Furthermore, because pregnancy registries may have the capacity to conduct specific examinations of infants born to exposed and unexposed mothers, including up until the time the child is 1 year old, they can identify outcomes other than major structural malformations, such as a developmental delay.

Postmarketing pregnancy exposure registries are being increasingly used proactively to monitor for major fetal effects. Sponsors may develop pregnancy exposure registries, either on their own initiative or when requested by the FDA as a postmarketing commitment. In 2002, the Agency published industry guidance on establishing pregnancy exposure registries to encourage the a priori development of epidemiologically sound, written, study protocols.[6] The *Pharmacist's Guide to Pregnancy Registry Studies* provides a basic checklist for evaluating pregnancy exposure registries (Weiss 1999). A list of current pregnancy exposure registries enrolling women is available at http://www.fda.gov/womens/registries/default.htm.

2. Retrospective Studies

a. Birth defect registries

Birth defect registries such as the Centers for Disease Control and Prevention's (CDC's) Metropolitan Atlanta Congenital Defects Program (Edmonds 1981) and the International Clearinghouse for Birth Defect Monitoring Systems (International Centre for Birth Defects 2000) monitor temporal and geographic frequencies of birth defects. The concept of these registries is that the introduction of a new major teratogen would lead to an unusual frequency or clustering of particular defects. Although no drug teratogen has ever been detected using a birth defect registry (Khoury 1987), the registries have been used to identify cases for use in case control studies (Robert 1982) and to provide population-based birth defect rates for use as a comparator in pregnancy exposure registries (Honein 1999).

b. Case control studies

Case control studies enroll infants with specific birth defects and collect data on gestational exposures, including drugs. The frequency of drug exposure in the cases is compared with the frequency of exposure to the same drug in a group of controls without the birth defects of interest. Information on drug exposure is usually obtained retrospectively by interviews, questionnaires, or review of medical records. Case control studies can be conducted on an ad hoc basis or as part of an ongoing case control surveillance study such as the Slone

[6] See guidance for industry *Establishing Pregnancy Exposure Registries*. (http://www.fda.gov/cder/guidance/index.htm)

Epidemiology Center's Birth Defects Study (Mitchell 1981) or the CDC's National Birth Defects Prevention Study (Yoon 2001).

Although pregnancy exposure registries are limited to screening for major teratogens on the level of thalidomide or isotretinoin, case control studies have the statistical power to identify teratogens with more modest risks on the level of valproic acid. The main strength of case control studies is their ability to evaluate the risk of rare events, and in the setting of birth defects, this strength means that such studies are highly efficient in identifying the risk of specific birth defects.

Recall bias is always of concern with case control studies. There are limitations to the specificity with which mothers can recall their drug exposures during pregnancy, particularly if a drug is used only occasionally for an indication that may not be easily recalled (e.g., analgesic for a headache). A mother of an infant born with a major birth defect may be more likely to carefully recall all gestational events and exposures than the mother of a normal infant (Werler 1989), and this phenomenon, though poorly documented, must be taken into account in the design and analysis of case control studies. One approach that can minimize the risk of recall bias is the use of infants with a wide range of other defects, as controls may help correct for this bias (Lieff 1999, Mitchell 2000). Another approach is the use of medical or pharmacy records to confirm drug exposures. This would at least confirm that a drug was prescribed or a prescription filled; whether the drug was actually ingested would still be unknown.

V. RESOURCES FOR MORE INFORMATION

FDA staff have access to several external resources that can assist in assessing reproductive toxicities from drug exposures. The online Micromedex Integrated Index, which can be accessed via the FDA Medical Libraries' WebLEARN intranet page, contains the REPRORISK system. This system contains electronic versions of four comprehensive, periodically updated, teratogen information databases, which are scientifically reviewed resources that critically evaluate the literature regarding human and animal pregnancy drug exposures. The four databases are:

- REPROTEXT Reproductive Hazard Reference
- REPROTOX (www.reprotox.org)
- Shephard's Catalog of Teratogenic Agents (Shephard 2001)
- TERIS Teratogen Information System (Friedman 2000)

Other reference texts to consult include:

- Chemically Induced Birth Defects (Schardein 2000)
- Drugs in Pregnancy and Lactation. A Reference Guide to Fetal and Neonatal Risk (Briggs 1998)
- The Women, Their Offspring, and the Malformations (Heinonen 1977)

Within the FDA, it is important for clinical reviewers to consult each other as well as pharmtox reviewers, drug safety evaluators, epidemiologists, and statisticians. In particular, the FDA's Pregnancy Labeling Team (PLT) is available as a consultative resource. The PLT ensures that the FDA maintains access to experts in teratology, birth defects epidemiology, or obstetric clinical pharmacology who are approved Special Government Employees (SGEs) and available for consultation by request.

VI. OVERALL ASSESSMENT OF POSTMARKETING HUMAN DATA

No set formula exists that can prove that an association between a gestational drug exposure and adverse pregnancy outcome is, in fact, a cause-effect relationship. There are no known medications for which it can be said that all exposed pregnancies would be adversely affected and only a few medications for which it is known that a large proportion of pregnancies would be affected. If a relationship between medication use and adverse outcome exists, it is more typical for the relationship to represent a relatively small increase over the background risk. Therefore, it is probably not useful to approach the review of data with the question: Is this medication teratogenic? What is useful is to consider the following:

- Does there appear to be an increased risk in exposed infants compared to the background rate seen in unexposed pregnancies?
- If so, what is the apparent likelihood of increased risk of adverse outcome associated with the medication?

Evidence from all sources, including human data from case reports, epidemiology studies, and animal data, should be considered collectively to determine the strength of the relationship. To test the possibility that an association is causal, there are six commonly used assessments that may be helpful to apply to any accumulated data. These criteria for evaluating the causal nature of an association were introduced by Sir Austin Bradford Hill (Hill 1965) and subsequently modified for use in teratology as summarized by Scialli (Scialli 1992).

- *Strength of the Association* — What is the statistical likelihood that the association did not occur by chance alone? Is the same level of association seen in several studies?
- *Consistency of the Association* — Is the same association seen in several reports or studies? In different populations?
- *Specificity of the Association* — How often does the drug exposure occur without causing the effect, and how often does the effect occur separately from drug exposure?
- *Appropriate Timing* — Does the association make chronological sense? Was the drug taken at the critical time in development to affect the target organ?
- *Dose-Response Relationship* — Does the likelihood and magnitude of response increase with dose of the drug? Although a dose-response relationship is not often seen in human reports or studies because the same dose or a narrow range of doses is

typically used, in animal studies the production of toxicity is expected to be dose dependent.

- *Biological Plausibility* — Does the association make biological sense? Does the drug cross the placenta? Assessment of biologic plausibility requires consideration of the dose of a potentially toxic agent, its pharmacology, maternal and fetal metabolism, mechanism and targets of toxicity, variability of response, and windows of likely effect. For example, a product that is known to alter cardiac tissue development, but not known to affect vasculature, skeletal muscle, or bone development, is unlikely to have been responsible for limb reduction defects in the case of a mother exposed only in the third trimester. The type of defect seen and the timing of exposure make this implausible.

VII. LABELING

As part of the Periodic Safety Update Report (PSUR) sponsors are asked to specifically report on "positive or negative experiences during pregnancy or lactation," evaluate new human data as they become available, in the context of what is already known about the reproductive effects of the drug, and, if clinically relevant, communicate conclusions regarding risk or lack of risk associated with gestational exposure in the product labeling.[7] The lack of human data must also be noted in the labeling (21 CFR 201.57(f)(6)(i)(*b*) and (*c*)).

If there is an increased risk associated with the use of the drug during pregnancy, the labeling should describe, to the extent possible, the specific abnormality, the incidence, seriousness, reversibility, and correctability of the abnormality and the effect of dose, duration of exposure, or gestational timing of exposure on the likelihood of risk.

The labeling must also include a description of all adequate and well-controlled studies that have failed to demonstrate a risk from gestational exposure to the drug (§ 201.57(f)(6)(i)(*a*) and (*b*)). Whenever possible, for all critically assessed, valid studies, the labeling should include confidence limits and power calculations to establish the statistical power of the study to identify or rule out a specified level of risk. The labeling should include and be routinely updated with data from ongoing studies, such as pregnancy exposure registries.

The labeling generally should not include isolated case reports unless there has been a conscious, scientific judgment made by the sponsor and FDA reviewer that the quality of the reports and other factors (e.g., consistency with animal findings; information on dose, duration, and timing of gestational exposure; or biologic plausibility) support inclusion.

[7] See ICH guidance for industry *E2C Clinical Safety Data Management: Periodic Safety Update Reports for Marketed Drugs*. (http://www.fda.gov/cder/guidance/index.htm)

REFERENCES

Bjerkedal, T, A Czeizel, J Goujard, et al., 1982, Valproic Acid and Spina Bifida, Lancet, 2:1096.

Brent, RL, 2001, Addressing Environmentally Caused Human Birth Defects, Pediatrics in Review, 22(5):153 – 165.

Chambers, CD, SR Braddock, GG Briggs, et al., 2001, Postmarketing Surveillance for Human Teratogenicity: A Model Approach, Teratology, 64(5):252-61.

Edmonds, LD, PM Layde, LM James, et al., 1981, Congenital Malformations Surveillance: Two American Systems, Int J Epidemiol, 10:247-52.

Ernhart, CB, RJ Sokol, S Martier, et al., 1987, Alcohol Teratogenicity in the Human: A Detailed Assessment of Specificity, Critical Period, and Threshold, Am J Obstet Gynecol, 156:33-9.

Ferencz, C, C Loffredo, and PD Wilson (letter), 2000, Epidemiologic Methods in the Search for Causes of Human Birth Defects, Teratology, 61:243-5.

Friedman, JM, and JE Polifka, 2000, Teratogenic Effects of Drugs: A Resource for Clinicians (TERIS), second edition, Baltimore: Johns Hopkins University Press.

Goldberg, JD, and MS Mitchell, 1986, The Value of Case Reports in Human Teratology, Am J Obstet Gynecol, 154:479-82.

Hartzema, AG, M Porta, and HH Tilson, editors, 1998, Pharmacoepidemiology: An Introduction, third edition, Cincinnati, OH: Harvey Whitney Books.

Heinonen, OP, D Slone, and S Shapiro, 1977, The Women, Their Offspring, and the Malformations, In: Kaufman, DW, ed., Birth Defects and Drugs in Pregnancy, Littleton, MA: Publishing Sciences.

Herbst, AL, DC Poskanzer, SJ Robboy, et al., 1975, Prenatal Exposure to Stilbestrol: A Prospective Comparison of Exposed Female Offspring with Unexposed Controls, N Eng J Med, 292:334-339.

Herbst, AL, RE Scully, and SJ Robboy, 1975, Effects of Maternal DES Ingestion on the Female Genital Tract, Hospital Practice, October, 51-57.

Hernandez-Diaz, S, MM Werler, AM Walker, et al., 2001, Neural Tube Defects in Relation to Use of Folic Acid Antagonists During Pregnancy, Am J Epidemiol, 153:961-8.

Hiilesma, VK, AH Bardy, ML Granstrom, et al., 1980, Valproic Acid During Pregnancy, The Lancet, 1:883.

Hill, AB, 1965, The Environment and Disease: Association or Causation?, Proc R Soc Med, 58:295-300.

Holmes, LB, EA Harvey, BA Coull, et al., 2001, The Teratogenicity of Anticonvulsant Drugs, N Engl J Med, 344:1132-8.

Honein, MA, LJ Paulozzi, JD Cragan, et al., 1999, Evaluation of Selected Characteristics of Pregnancy Drug Registries, Teratology, 60:356-64.

International Centre for Birth Defects, 2000, Annual Report 2000, Rome: International Centre for Birth Defects. (http://www.icbd.org)

Jones, KL, RV Lacro, KA Johnson, and J Adams, 1989, Patterns of Malformations in the Children of Women Treated with Carbamazepine During Pregnancy, N Engl J Med, 320:1661-6.

Khoury, MJ and NA Holtzman, 1987, On the Ability of Birth Defects Monitoring to Detect New Teratogens, Am J Epidemiol, 126(1):136-143.

Koren, G, A Pastuszak, and S Ito, 1998, Drugs in Pregnancy, N Engl J Med, 338(16):1128-1137.

Koren, G, M Bologa, D Long, et al., 1989, Perceptions of Teratogenic Risk by Pregnant Women Exposed to Drugs and Chemicals During the First Trimester, Am J Obstet Gynecol, 160:1190-94.

Lammer, EJ, DT Chen, RM Hoar, et al., 1985, Retinoic Acid Embryopathy, N Engl J Med, 313:837-41.

Lary, JM and LD Edmonds, 1996, Prevalence of Spina Bifida at Birth — United States, 1983-1990: A Comparison of Two Surveillance Systems, In: CDC Surveillance Summaries, April 19, MMWR, 45(SS-2):15-26.

Leiff, S, AF Olshan, MM Werler, et al., 1999, Selection Bias and the Use of Controls with Malformations in Case-Control Studies of Birth Defects, Epidemiology, 10:238-41.

Lenz, W, 1961, Kindliche Missbildungen nach Medikament-Einnahma Wahrend der Graviditat, Dtsch Med Wochenschr, 86:2555.

Leppig, KA, MM Werler, CI Cann, et al., 1987, Predictive Value of Minor Anomalies, I. Association with Major Malformations, J Pediatr, 110: 530-7.

March of Dimes Birth Defects Foundation, 2001, Fact Sheets: Birth Defects, Miscarriage, and Stillbirth; Leading Categories of Birth Defects; and Perinatal Profiles: Statistics for Monitoring State Maternal and Infant Health. (www.modimes.org).

Mastroianni, AC, R Faden, and D Federman, editors, 1994, Risks to Reproduction and Offspring, In: Women and Health Research: Ethical and Legal Issues of Including Women in Clinical Trials, Washington, DC: Natl Acad Sci Press, 175-202.

McBride, WG, 1961, Thalidomide and Congenital Abnormalities, Lancet, 2:1358.

Mitchell, AA, 2000, Special Considerations in Studies of Drug-Induced Birth Defects, In: Strom, BL, editor, Pharmacoepidemiology, third edition, England: John Wiley & Sons, Ltd., pp. 749-63.

Mitchell, AA, L Rosenberg, S Shapiro, and D Slone, 1981, Birth Defects Related to Bendectin Use in Pregnancy. I. Oral Clefts and Cardiac Defects, JAMA, 245(22):2311-4.

Moore, KL and TVN Persaud, 1998, The Developing Human: Clinically Oriented Embryology, sixth edition, Philadelphia: W.B. Saunders.

National Birth Defects Prevention Network (NBDPN), 2002, Birth Defects Surveillance Data from Selected States, 1995-1999, Teratology, 66:S129-S211.

Nichols, CR, JF Dickinson, and PJ Pemberton, 1997, Rising Incidence of Gastroschisis in Teenage Pregnancies, J Matern Fetal Med, 6:225-9.

Petrini, J, K Damas, and RB Johnston, 1997, An Overview of Infant Mortality and Birth Defects in the United States, Teratology, 56(1/2):8-9.

Pettifor, JM and R Benson, 1975, Congenital Malformations Associated with the Administration of Oral Anticoagulants During Pregnancy, J Pediatr, 86; 459-62.

Pole, M, A Einarson, N Pairaudeau, et al., 2000, Drug Labeling and Risk Perceptions of Teratogenicity: A Survey of Pregnant Canadian Women and Their Health Professionals, J Clin Pharmacology, 40:573-7.

Polifka, JE and JM Friedman, 1999, Clinical Teratology: Identifying Teratogenic Risks in Humans, Clin Genet, 56:409-420.

Robert, E and P Guiband, 1982, Maternal Valproate Exposure and Neural Tube Defects, Lancet, 2:937.

Schardein, JL, 2000, Chemically Induced Birth Defects, third edition, New York: Marcel Dekker, Inc.

Scheuerle, A and H Tilson, 2002, Birth Defect Classification by Organ System: A Novel Approach to Heighten Teratogenic Signaling in a Pregnancy Registry, Pharmacoepi Drug Safety, 11:1-11.

Scialli, AR, 1992, A Clinical Guide to Reproductive and Developmental Toxicology, Boca Raton, FL: CRC Press.

Scialli, AR, A Lione, and CKB Padgett, 1995, Reproductive Effects of Chemical, Physical and Biologic Agents, REPROTOX, Baltimore: Johns Hopkins University Press.

Sedman, AB, DB Kershaw, and TE Bunchman, 1995, Recognition and Management of Angiotension Converting Enzyme Inhibitor Fetopathy, Pediatr Nephrol, 9(3):382-5.

Shephard, TH, 2001, A Catalog of Teratogenic Agents, 10th edition, Baltimore: Johns Hopkins University Press.

Strom, BL, editor, 2000, Pharmacoepidemiology, third edition, England: John Wiley & Sons, Ltd.

Ventura, SJ, WD Mosher, SC Curtin, et al., 2000, Trends in Pregnancies and Pregnancy Rates by Outcome, United States, 1976-96, National Center for Health Statistics, Vital Health Stat, 21(56).

Weiss, SR, CE Cooke, LR Bradley, and JM Manson, 1999, Pharmacist's Guide to Pregnancy Registry Studies, J Am Pharm Assoc, 39:830-4.

Werler, MM, BR Pober, K Nelson, et al., 1989, Reporting Accuracy among Mothers of Malformed and Nonmalformed Infants, Am J Epidemiol, 129:415-21.

Yoon, PW, SA Rasmussen, MC Lynberg, et al., 2001, The National Birth Defects Prevention Study, Public Health Reports, 116 (Supp. 1):32-40.

Contains Nonbinding Recommendations

ATTACHMENT A

Critical Periods in Human Development

Attachment A, which appears on the following page, was taken from *The Developing Human: Clinically Oriented Embryology* (Moore, KL and TVN Persaud, 1998, sixth edition, Philadelphia: W.B. Saunders Company, p. 548) and reproduced with permission. For an accurate interpretation of the diagram, it should be printed in color.

CRITICAL PERIODS IN HUMAN DEVELOPMENT*

*Mauve denotes highly sensitive periods when major birth defects may be produced.

Contains Nonbinding Recommendations

ATTACHMENT B

Considerations when Evaluating Case Reports of the Effects of Drug Exposure During Pregnancy

- Is there information on drug dose?
- Is there information on dates of drug therapy?
- Is the date of conception or the LMP given?
- Does the drug cross the placenta?
- Was the drug given at the appropriate time during gestation to cause the reported fetal effect or problem?
- Were any other medications given to the mother? If so, are they associated with any birth defects?
- What maternal attributes are included (e.g., age, race, other)? Could they be associated with the reported fetal effect?
- What is known about the mother's medical history?
- What is known about maternal and paternal family histories of birth defects?
- If the drug was given for treatment of a maternal disease, what is known about fetal effects associated with that maternal disease?
- Does the reported fetal effect have any familial or genetic associations?
- Did a physician evaluate and confirm the infant's condition? What is the level of expertise of the person who made the diagnosis?
- What types of developmental problems were seen in the animal studies? Is the reported human effect similar in presentation or mechanistically?
- Are there any other similar human reports in AERS or in the literature?
- What do the REPRORISK system and other reference texts say about the drug?
- Are similar fetal effects seen with pharmacologically similar drugs?

NOTE: If important data are missing from the case report, the initial reporter should be contacted.

Contains Nonbinding Recommendations

ATTACHMENT C

Considerations when Evaluating Epidemiology Studies on the Effects of Drug Exposure During Pregnancy

Exposure information

- Are dose, duration, and timing of exposure included?
- How are exposures reported and/or confirmed (e.g., patient report, health care professional report, medical/pharmacy records)?
- With case control studies, what is the length of time between birth and obtaining information on exposures during pregnancy?

Loss to follow-up

- What is the proportion lost to follow-up?

Outcome ascertainment

- What outcomes were evaluated (e.g., spontaneous abortion, still birth, congenital anomalies, neurodevelopmental toxicities)?
- When is outcome ascertained?
- Who is the source of outcome information (e.g., mother, maternal health care provider, infant health care provider)?
- Are reported adverse outcomes confirmed with medical records?
- Are normal live births confirmed with medical records?
- What is the length of follow-up on the infant (e.g., only at birth, up to 1 year of age)?

Additional factors and influences on pregnancy outcome (potential confounders)

- What additional information is collected (e.g., indication, concomitant medications, medical and family history, previous pregnancies and outcomes, socio-demographics, smoking, alcohol consumption)?

Comparison group

- What type of comparison group was used?
- Did the study have an internal comparison group?
- With case control studies, what is the impact of the comparison group on recall bias?

Representativeness

- How representative of the U.S. general population is the population covered by the study?

Analysis and Interpretation

- How was the risk of each outcome calculated? Were all exposures in the denominator eligible to be in the numerator?

- How does the loss to follow-up influence the study results?
- How are the data grouped for analysis based on timing of exposure (e.g., by trimester, by weeks of gestation)?
- How does the study address concerns raised by animal studies? Are there other relevant outcomes that the study was not able to address?

Power

- What level of risk is the study powered to evaluate?
- Are confidence intervals provided?