chapter 6

Reproductive Risk Assessment

CONTENTS

List of Tables

Chapter 6

Reproductive Risk Assessment

INTRODUCTION

Health risk assessment is the use of scientific evidence to estimate the likelihood of adverse effects on the health of individuals or populations from specific exposures to hazardous materials and conditions. Although risk assessment is often confused with risk management, the two are different. Risk assessment attempts to evaluate the probability of occurrence of biologically significant events, while risk management determines the possible actions that can or should be taken to respond to an assessment of significant risk. This chapter discusses some of the complexities in reproductive risk assessment; risk management is the subject of chapter 7. Ethical issues surrounding the difficulty of separating value judgments from the risk assessment process are discussed in the background paper, Ethical ***Issues in Reproductive Health Hazards in the Workplace,*** prepared for this report (see appendix F).

Several government agencies are charged with the regulation of harmful substances and thus with risk assessment and/or risk management. A number of measures designed to centralize and standardize the risk assessment and management processes have been proposed (reviewed in ref. 5). Because these agencies have differing mandates based on the legislation underlying their authority and the types of substances and environments that are of concern, the feasibility of centralizing the risk assessment and management processes among them is uncertain. But there is the potential for establishing guidelines that can make the procedures and assumptions used in risk assessment and management processes explicit.

Health risk assessments always involve scientific uncertainties. It is not possible to predict the likelihood of a particular health effect from a given exposure situation without some degree of uncertainty regarding the exact number of people who may be affected. Scientific decisions regarding use of particular models and dose-response curves, for example, carry with them judgments that can ultimately result in different assessments of risk and thus different risk management policies. Critical steps in the risk assessment process frequently require not only scientific information, but also judgment, experience, intuition, and common sense.

THE RISK ASSESSMENT PROCESS

The risk assessment process usually contains four steps (18): hazard identification, dose-response assessment, exposure assessment, and risk characterization.

Hazard Identification

The first step in risk assessment is hazard identification, the qualitative analysis of all available animal and human data to determine whether, and at what dose, an agent is or is not likely to cause reproductive impairment. Hazard identification determines the potential of an agent to do harm, not the probability that harm will, in fact, occur (7).

Part of the task of hazard identification is to determine whether the toxin is a reproductive or developmental toxin, or both. In general, reproductive toxins are substances that affect adults. They can cause a range of effects from genetic change to systemic damage. They may act directly on reproductive organs or impair reproductive health by damaging other systems (neural, endocrine, or circulatory). Developmental toxins affect the offspring of individuals. They can cause delays in growth, malformations, cancer, behavioral

changes, or death of the embryo/fetus (see chapter 3). once the existence of a hazard has been established, the remaining steps of risk assessment-dose-response assessment, exposure assessment, and risk characterization-can begin.

Dose-Response Assessment

In dose-response assessment the relationship between the magnitude of exposure and the probability of human health effects is determined. This step nearly always involves the evaluation of animal studies that test the effects observed in a range of doses. Also involved in this process is the task of extrapolating the effects of the high doses used in animal studies to lower doses or the actual exposure levels that humans are likely to encounter. Interpretation of results is extremely complex because particular reproductive outcomes or endpoints may be difficult to observe, and numerous other variables (e g., age, sex, lifestyle) may affect response in humans. Scientists must take account of differences in reproductive function and structure among animal species and between animals and humans; different in-utero and post-utero development; and different rates of metabolism and excretion of toxins.

Exposure Assessment

Exposure assessment identifies the population segments potentially exposed to the agent, including their composition and size as well as the magnitude, frequency, and duration of potential exposure to the agent. These data are often difficult to obtain.

Exposure to a reproductive health hazard must occur for the hazard to have an effect. Exposure may be: 1) acute (one-time) exposure, 2) episodic (recurrent but discrete) exposure, or 3) chronic (constantly present) exposure. Acute or episodic exposures are often relatively high doses over short periods of time, while chronic exposures are usually low doses over longer periods of time. Chronic exposure may also be characterized by high doses over long periods of time.

The timing and route of exposure can be very important to normal fetal development. The exposure may be of brief duration, but if it occurs at a critical point of development of the embryo/fetus, the effects can be profound. A toxin can have different effects because of the route of exposure. Some toxins have their greatest detrimental impact when inhaled. There can also be indirect exposure. The spouse, a developing embryo/fetus, or children of a worker can be exposed to substances carried home on clothing or equipment.

Reliable estimates of the number of workers potentially exposed to harmful substances and the specific substances to which they are exposed are not currently available. However, the National Institute for Occupational Safety and Health (NIOSH) is in the process of tabulating the results of an update of the 1972-74 National Occupational Hazard Survey to estimate the numbers of workers potentially exposed to specific substances. Preliminary tabulations should be available by late 1985. The information will be tabulated by sex but not by age. Estimates of exposure are extremely difficult to obtain because workers maybe exposed to more than one substance and trade secrets make identification of substances difficult and time-consuming.

Estimates of human risk are complicated by individual differences in susceptibility to the effects of various levels of exposure, and the likelihood of time lag between hazard exposure and reproductive effect. Lifestyle characteristics such as smoking or alcohol consumption can increase the risk of reproductive impairment and may act additively or synergistically with hazards to which people are exposed in the workplace. Workers who have health problems associated with lower socioeconomic status may cluster in industries where hazards to their reproductive systems are more likely to be present. And people vary in their susceptibility to various harmful agents,

Risk Characterization

In this final step the data from dose-response assessment and exposure assessment are combined to estimate the actual risk from the agent. The strengths and weaknesses in each phase of the assessment are presented and summarized as a part of this step, along with the assumptions and extent of uncertainties encountered in the

process. The critical component is the estimate of the level of uncertainty in the conclusions (19,23).

The transition from each step in the process is a decision point that affects allocation of resources, If the hazard assessment indicates that a hazard does not exist, resources can then be allocated to another task. If, following the risk characterization phase, a substantial risk is identified, risk management decisions must begin (see chapter 7).

DATA USED IN REPRODUCTIVE RISK ASSESSMENT

The signal that a chemical, physical, or biological agent may warrant risk assessment can come from several sources. For a new chemical, evidence r-nay surface from toxicological tests carried out by the manufacturer in order to submit a Premanufacture Notification to the Environmental Protection Agency (EPA). However, this is an unreliable source from which to derive data on reproductive or developmental health hazards because test requirements do not specify reproductive endpoints that must be examined (28). Health hazard evaluations and NIOSH or EPA research also serve as input for risk assessments, as noted later in this chapter. Two primary sources of information are epidemiological and toxicological studies published in scientific journals.

Epidemiological Studies

Epidemiology is the study of relationships between the frequency and distribution, and the factors that may influence frequency and distribution, of diseases and injuries in human populations. The underlying tenet of epidemiology is that diseases are not distributed randomly in a population but tend to cluster (26). These groups or clusters of disease can be studied in order to discover whether the clusters are, in fact, random, or are linked to some causal factor or factors.

Epidemiology studies can have a macro or micro level of focus; both levels are important. Macro-level studies, usually surveillance systems or programs, involve large samples and are important for measuring baseline rates of reproductive endpoints such as normal and low birth weight or the frequency of congenital malformations in large segments of the population. In contrast, micro-level studies are usually concerned with a subpopulation (workers, for example) at risk because of exposure to a substance. Micro-1evel studies can take various forms, depending on the endpoints or group of individuals being studied.

Epidemiological studies can be divided into three broad classes: descriptive, analytical, and experimental. Descriptive and analytical studies are more often utilized for studying reproductive impairment.1 (For further discussion of study designs see ref. 2.)

Descriptive studies

There are two types of descriptive studies. The first, case reports (also called observational epidemiology), can highlight the occurrence of a cluster of cases of reproductive impairment, which may indicate that a potential problem exists. These are often clinical reports from occupational health physicians. The detection of infertility in DBCP-exposed men in a pesticide-manufacturing plant in California, as noted in chapter 2, is an example of this type of study. An earlier example is the detection of rubella as a causative agent of birth defects by an Australian ophthalmologist, who observed congenital cataracts in many of the offspring of his patients. When his investigations revealed that their mothers had contracted rubella during their pregnancies, he became the first to clearly implicate this disease as the cause of cataracts and other birth defects (24). This ap-

‘Experimental studies are difficult to undertak**en industrial settings because subjects must be assigned to treatment groups. For ethical reasons, investigators must usually accept the situation as it exists with regard to exposure, and then identify appropriate comparison groups. Data from clinical trials are reviewed in the risk assessment process if they are pertinent, however. For example, results from clinical trials (experimental studies) of estrogen contraceptives are reviewed to help delineate the risk of exposure to estrogen compounds in the workplace.**

preach has two major disadvantages, however: the damage from the hazard has already occurred, and the studies are serendipitous in nature. Some hazards may thus go undetected or may already have affected large numbers of people by the time they are finally detected.

The second type of study, surveillance, is important for the detection of certain kinds of reproductive dysfunction. As indicated previously, surveillance systems are usually large-scale enterprises that produce information on baseline rates in the total population. Large-scale malformation surveillance programs, for example, are an important source of information on the occurrence of birth defects. U.S. programs include the Birth Defects Monitoring Program and the Metropolitan Atlanta Congenital Defects Surveillance Program conducted by the Centers for Disease Control (CDC). (A review of State and national surveillance and monitoring programs appears in ref. 24.)

Welldesigned surveillance systems have several advantages (10,24). First, they provide background incidence and prevalence rates for large numbers of persons. These background rates are valuable in detecting changes in the frequency of reproductive endpoints. Increased frequencies in time or geographical area can be checked to determine whether a true increase exists and follow-up investigations can be initiated to ascertain the cause. Second, time trends can be monitored and reproductive endpoints of specific interest can be targeted for careful investigation. Third, surveillance can provide reassurance about the absence of problems. Since the inception of birth defects surveillance programs around the world, no new teratogen has yet been initially identified in a surveillance system. Although this may indicate that the systems are not sensitive enough, most experts believe that they are adequate and that new developmental effects would have been recorded had they occurred (10,24). The major disadvantage of surveillance systems is their expense.

Micro-level concerns are the focus of monitoring studies. In these programs a population at risk can be identified and followed over time in order to detect an outcome of interest. Relatively small groups, such as persons in particular employment groups, or persons working at factories manufacturing specific products, can be studied. Monitoring systems have an advantage in that they permit observation of a population that is exposed to suspect substances. For example, a birth defects monitoring system for the Rhone-Alps region of France was able to detect an association between maternal valproic acid ingestion and the occurrence of infants born with lumbosacral neural tube defects. Valproic acid is an anticonvulsant that was used by pregnant women (3,22).

The American Petroleum Institute (10) commissioned a review of reproductive health surveillance and monitoring activities both within and outside the industry.[2] The nine U.S. oil companies that have monitoring systems have several characteristics in common: 1) reproductive monitoring is built into the existing employee health system, 2) provision is made for computer storage and editing of the data, 3) there is computer linkage to personnel records and some type of exposure data, and 4) all intend some type of analysis of this data. None have as yet analyzed the data or determined the types of statistical analyses to be used. (A summary of these systems appears in ref. 10.)

Analytical Studies

Analytical studies test for an association between exposure and outcome or result. There are three types of analytical studies: cross-sectional, case-control, and cohort. Analytical studies look for an association between an agent (e.g., exposure to a potentially harmful substance) and a particular outcome (e.g., increased rate of spontaneous abortion or lowered sperm counts). This is done by comparing a group or groups of exposed individuals with matched control groups. Cross-sectional studies compare exposed groups with control groups at one point in time; case-control studies compare individuals with a particular outcome with controls and look at prior exposure

[2]Thirty-nine companies were **surveyed; 27 reported little or no activity, 3 refused to participate, and 9 agreed to be interviewed. See** (10) for details.

in the two groups; cohort studies follow groups that differ in amounts of exposure and look for differences in the frequency of particular outcomes in each group. (Further discussion of these studies appears in refs. 2 and 26.)

General Considerations in Epidemiological Studies

The results of epidemiology studies may be invalid because of the complexity of factors that must be taken into consideration in the design and implementation of the studies. These factors include:

Design of the Study.–The design of the study is crucial. If the study has been improperly designed, the investigator may not be able to answer the research question or the research may take longer than necessary. Selection of the appropriate control group is also crucial. If control groups are not carefully matched with exposed groups, study results may be invalid.

Measurement of Reproductive Endpoints.—The measurement of the reproductive endpoints under study must be valid and reliable. Most reproductive endpoints are extremely difficult to measure. For example, investigators studying male infertility are not in agreement as to which tests of semen characteristics best measure infertility (validity), and test results of semen characteristics vary from laboratory to laboratory (reliability). Another endpoint, the spontaneous abortion rate, is extremely difficult to study. It has been estimated that only about 31 percent of all fertilized eggs survive to term: about 16 percent do not make the first cell division, another 15 percent are lost during the first week, and a further 27 percent during implantation. By the time of the first missed menstrual period, only about 42 percent of the fertilized eggs have survived (14,36). Many women thus spontaneously abort without realizing that they have been pregnant.

Recall bias must be considered. It is extremely difficult for all individuals to recall past events accurately,

Many reproductive endpoints are extremely rare in the population. Congenital malformations diagnosed at birth occur in about 3 percent of all births. Thus the study of a particular congenital malformation requires large numbers of births (see later discussion of sample size), and diagnoses can vary among physicians and hospitals. Many reproductive endpoints have several causes, only some of which may occur in the workplace environment.

Multiple endpoints can be affected by a particular toxicant, and there is usually no way to predict which outcomes are most likely. For example, alcohol consumption can increase the frequency of infertility, low birth weight, spontaneous abortion, congenital malformation, and developmental delay. By contrast, genetic effects may result in a variety of outcomes but show no particular pattern since genetic pathways can be affected at random (35).

The reproductive endpoints for which population frequencies are available in the United States are listed in table 6-1. No population frequencies are available for sexual dysfunction, menstrual problems, semen quality, and childhood cancer.

Table 6.1.— Reproductive Endpoints for Which Population Estimates are Available

Endpoint	Population survey[a]
1. Infertility of male and female origin.	NSFG, PYS
2. Conception delay.	NSFG, PYS
3. Birth rate	NSFG, NNS, NFMS, PYS
4. Pregnancy complications . . .	NSFG, NNS, NFMS, PYS
5. Gestation at delivery (prematurity, postmaturity) . .	NSFG, NNS, NFMS
6. Early fetal loss (c28 weeks gestation)	NSFG, NNS, NFMS, PYS
7. Late fetal loss (>28 weeks gestation)	NSFG, NNS, NFMS, PYS
8. Sex ratio	NSFG, NNS, PYS
9. Birth weight	NSFG, NNS
10. Apgar score	NNS
11. Congenital defect	NNS
12. Infant morbidity and mortality	NSFG, NNS
13. Childhood morbidity and mortality	NNS, NFMS, PYS

[a]NSFG = 1982 National Survey of Family Growth; NNS = 1980 National Natal ity Survey; NFMS = 1980 National Fetal Mortality Survey; PYS = Parries Youth Survey.

NOTE: These surveys also contain data on the following related topics: onset of menses, fertility expectations, birth spacing, contraceptive use, sterilization, care+ seeking for infertility, prenatal care, spontaneous and induced abortions, maternal smoking and alcohol consumption, chronic diseases, and venereal infections in pregnancy.

SOURCE: Adapted from M. Hatch, V, Stefanchik-Scott, and Z.A. Stein, "Surveillance of Reproductive Health in the U. S.: A Survey of Activity Within and Outside Industry, " unpublished, prepared for the American Petroleum Institute, December 1983.

Indications of the prevalence of some of these endpoints are available from tumor registries or individual studies from infertility and prenatal clinics (10).

Many individuals, especially workers, are reluctant to cooperate in studies because they consider them an invasion of their privacy. Some workers also believe that their medical records may be used to compromise their work status or possibilities for promotion. In addition, companies may not wish to participate in a study either because they employ their own epidemiologists or they are concerned about the liability ramifications if substances to which their employees are exposed are found to be associated with adverse effects. All of these considerations must be carefully evaluated by the investigator and must also be taken into account by those who review results of epidemiological studies during the risk assessment process.

Key Factors

The size of the sample must be adequate to demonstrate at a given level of statistical significance that there is an association between exposure and outcome variables. Three important factors are interrelated: the power of the test, the sample size needed to show a significant difference, and the presence of confounding variables.

Power.—Power is the probability of detecting a specified difference in effect between experimental and control groups. The power of a given study is determined by the sample size, background incidence of the endpoint(s) measured, and the variance of the endpoints. Power is directly related to sample size and inversely related to background incidence and variance. Power is very important because the higher the power of a test, the stronger the possible conclusions regarding the exposuremutcome relationship. If the test lacks sufficient power, two possible errors can occur:

1. the results indicate that an exposure is associated with an outcome when, in fact, there is no association (Type I error); and
2. the results show no association between the exposure and an outcome when an association in fact exists (Type II error).[3]

The probability of a Type I error is estimated with a test statistic called alpha. Before an association is said to be significant, the probability of its occurring as a result of chance sampling fluctuations (i.e., the probability of a Type I error) must be less than some predetermined value, called the statistical significance level (12,13,27).

The power is often low in studies of worker populations because the sample sizes are small. Study results, therefore, can erroneously show that exposure is not associated with the reproductive outcome when it may be.

The investigator selects the power of the test by choosing the probabilities of these two possible errors. Once this has been done, the investigator determines the frequency of the endpoint in the population in order to choose a sample of sufficient size to meet the power constraints already set (26).

Sample Size.—The adequacy of the sample size is directly related to the frequency of the reproductive endpoint in the population. If the frequency is small, for example, less than 15 percent, large samples are needed. In addition, the investigator must decide how much of a difference is a significant difference. For example, if the frequency is 15 percent, a far larger sample size would be required to show that 18 percent is a significant difference than to show that a doubling (30 percent) is a significant difference.

The frequencies of selected adverse reproductive outcomes and the sample sizes necessary to show that a twofold difference in those rates is significant are shown in table 6-2. For example, in order to detect a twofold increase in the spontaneous abortion rate (during the period from the

[3]Type I and II errors are often defined slightly differently because the researcher is testing a null hypothesis, that is, that there is no association between two variables. **The error of rejecting the null hypothesis of no association when the hypothesis is true is a Type I error. The error of not rejecting the null hypothesis when it is in fact false is a Type II error.**

Table 6-2.—Sample Size Required to Detect Twofold Increase in Adverse Reproductive Outcomes[a]

Outcome	Sample size[b]
Impaired fertility:	
No conception after 1 year unprotected intercourse	322 couples
Pregnancy loss:	
Spontaneous abortion (s20 weeks gestation)	322 pregnancies
Stillbirths	)
Birth/developmental defect:	
Low birth weight	586 live births
Major birth defects (all)	631 live births
Neural tube defects	1,819 live births
Severe mental retardation	8,986 live births
Chromosomal abnormalities	17,902 live births
Infant (s 1 year) death	1,856 live births

[a]Alpha = 0.05, beta = 0.20.
[b]Divided evenly between exposed and unexposed groups.

SOURCE: M. J. Rosenberg and L.H.Kuller, "Reproductive Epidemiology: What Are the Problems in Methodology?" *Reproductive Health Policies in the Workplace,* Proceedings of Symposium held on May 10-11, 1982, Pittsburgh, PA, Family Health Council of Western Pennsylvania, Inc., 1983, pp. 201-226.

point at which a pregnancy is recognized to 20 weeks gestation), 161 pregnancies are needed in both the exposure group and the control group. In order to study this many pregnancies, the investigator must draw on a large population. Using plausible assumptions about the birth rate and number of working women, the investigator would have to draw from a population of more than 11,000 workers to find a sufficient number of pregnancies to study (24).

Confounding Factors.—A confounding factor is a variable that is correlated with both exposure and outcome. It can therefore partially or wholly account for an apparent effect of the exposure levels under study or mask an underlying true association. Confounding factors include lifestyle variables such as smoking or alcohol consumption, or ascribed characteristics such as ethnic status or age.

Maternal age, for example, can be a confounding factor. In a hypothetical study of the relationship between cumulative occupational radiation exposure and Down syndrome, the case group might contain a greater number of workers with high cumulative exposure than the control group. Because older radiation workers would be expected to have greater cumulative radiation exposure than younger workers, the risk of Down syndrome would appear to be associated with cumulative radiation exposure when it may in fact have been due to the greater age of the exposed group. In this case, maternal age would be a confounding variable since it would be associated both with the risk of Down syndrome and with cumulative radiation exposure (26).

A confounding variable that is often overlooked in studies of developmental effects is paternal exposure. If the possibility of paternally mediated effects is not considered, invalid conclusions regarding maternally mediated effects on the embryo/fetus may result.

Toxicology Studies

Toxicology studies include in vitro and whole animal tests of suspected hazards that allow the investigator to examine the roles of dose and routes of exposure. While extrapolation to humans is a complicated task, these studies, properly executed and interpreted, can ***predict*** an association with agents to which humans are exposed, in contrast to epidemiology studies, in which the humans will already have been affected by exposure to the hazard.

Although evidence from studies on humans is often used to refute or confirm results from animal screening tests, toxicology studies are necessary for several reasons (20):

- Experimental studies that deliberately expose humans to potentially toxic chemicals are ethically unacceptable, except in special circumstances (e.g., clinical trials for new pharmaceuticals) where there is extensive evidence from animal studies and informed consent has been given.
- Epidemiological studies of workers exposed to a chemical already in production, or reports of adverse reactions to substances, are available for only a small number of chemicals (see chapter 4).
- Even in epidemiological studies of exposed humans, results are difficult to interpret because of factors such as the lack of large enough samples and good exposure data, difficulty in measuring endpoints, and confounding variables.
- Although epidemiological studies are valuable, tests on animals have proven to be an important source of data on human risk.

Single Generation and Multigeneration Studies

Animal tests for reproductive and developmental toxicity are divided broadly into two categories: single generation studies and multigeneration studies. Single generation studies were primarily devised to test the safety of new drugs to help prevent repetition of such occurrences as the thalidomide disaster, i.e., a test of one application, usually of a high dose. Multigeneration studies were devised to test the safety of food additives and unintentional food-processing contaminants such as pesticides and packing material residues; i.e., screening for effects of chronic exposure, usually at smaller doses. These studies are conducted for two purposes:

1. to investigate mechanisms of action of toxic chemicals on various reproductive processes, and/or
2. to screen chemicals in order to identify those that may present hazards to humans exposed to them (20).

These tests are often used to evaluate the safety of chemicals before clinical trials or commercial production, sometimes without full review of their suitability as models for occupational or environmental exposures (l). (Descriptions of single generation and multigeneration study designs appear in refs. 1,4, and 20.)

General Considerations of Toxicology studies

Design, Conduct, and Interpretation of Tests.—Evaluation of results of toxicity testing must include such considerations as the species to be selected; dosage, route, and timing of exposure; the number of animals to be used; the selection of positive and negative controls; the toxicokinetics (rates of metabolism and excretion of chemicals) of the animals being used; the endpoints under study; and whether appropriate statistical analyses have been carried out. (Discussion of these considerations appears in refs. 4 and 20.) (For discussion of experimental protocols for toxicity testing see refs. 4,11,16,17,19,21,31,32, 34,37.)

Differences in Structure and physiology Among Animal Species and Humans.—Although reproductive processes in the mouse, rat, hamster, guinea pig, rabbit, dog, and rhesus monkey are broadly similar to those in humans, there are a number of differences in anatomy, physiology, and timing of exposure that need to be taken into account when interpreting experimental results. For example, there are substantial interspecies differences in the structure of the placenta (table 6-3). Dogs and some other species have the most tissues separating fetal and maternal blood, followed by humans and female primates, who have more than rodents and rabbits. Humans differ from experimental species in the timing and development of the placenta and in metabolism and pharmacokinetics of toxic chemicals.

The physiology of pregnancy in rodents and humans differs markedly. In rodents, for example, pituitary function is essential during the first half of the pregnancy in rodents, whereas in humans it is not required once conception has occurred (l).

Concordance Between Animals and Humans.—There are two types of concordance, that of effect and that of dose. Concordance of effect is the extent to which the types of effects observed in humans are matched by similar or related effects observed in animals, while concordance of dose is the extent to which animals and humans are affected at similar dose levels (20).

Table 6-3.—Tissues Separating Fetal and Maternal Blood

Maternal tissue			Fetal tissue			
Endothelium	Connective tissue	Epitheliums	Trophoblast	Connective tissue	Endothelium	
Epitheliochorial . . . . . +	+	+	+	+	+	Pig, horse, donkey
Syndesmochorial. . . . . +	+	–	+	+	+	Sheep, goat, cow
Endotheliochorial . . . . +		–	+	+	+	Cat, dog
Hemochorial. . . . . –	–	–	+	+	+	Woman, monkey
Hemoendothelial . . . . . –		–	–	–	+	Rat, rabbit, guinea pig

SOURCE: 1. C. T. Nisbet and N. J. Karch, Clrernica/ Hazards to Human Reproducflon (Park Ridge, NJ: Noyes Oata Corp., 1983), p. 94.

A basic tenet of toxicology is that effects observed in experimental animals can be used to infer likely effects (or lack of effects) in humans, with appropriate consideration of biological differences between species. And, in general, animal models do have good predictive value for humans (see chapter 4). For example, in reproductive toxicology studies, substances that affect menstrual cycles in monkeys and estrous cycles in rodents also affect menstrual cycles in humans (tables 6-4 and 6-5). Effects on fertility in rodents also seem to be a good indicator of effects in humans; most of the original work on contraceptive agents was carried out on rodents (1). However, interpreting effects of toxic doses on sexual behavior and pregnancy from animals to humans is far more complex. There are so many differences in sexual behavior between humans and animals that special care must be exercised not to misinterpret results.

Selection of the proper species is extremely important because one or even several animal species may give "false negative" results. The experience with thalidomide is a case in point. Effects similar to the phocomelic-type limb deformities observed in humans were observed in a few breeds of rabbits and seven species of primates. Thalidomide has been tested in 10 strains of rats, 15 strains of mice, 11 breeds of rabbits, 2 breeds of dogs, 3 strains of hamsters, 8 species of primates, and in cats, armadillos, guinea pigs, swine, and ferrets. Developmental effects were only occasionally produced in any of these species. However, there were fertility effects: prenatal mortality was high in rabbits, and there was a low conception rate in rats (20). This underscores the importance of selecting the appropriate species, examining other endpoints as indicators of toxic effects, and of performing human epidemiology studies to corroborate the information from ani-

Table 6.4.—Selected Examples of Reproductive Toxic Effects Common to Animals and Humans

Compound	Effect in animals	Effect in humans
Benzene	Estrous cycle disturbance: rat	Menstrual disorders
Styrene	Estrous cycle disturbance: rat	Menstrual disorders
Chlordecone (Kepone)	Testicular atrophy, decreased fertility: mouse, rat, rabbit, both sexes, females more affected	Decreased sperm count and motility, abnormal morphology
Chloroprene	Testicular damage, decreased sperm count, dominant lethal mutations: mouse, rat, cat	Decreased libido, impotence, decreased sperm count, motility, abnormal morphology. Increased spontaneous abortion in wives
DBCP	Testicular atrophy, decreased fertility, dominant lethal mutations: rat, rabbit, guinea pig	Testicular atrophy, decreased sperm count, decreased fertility
Arsenic	Embryolethal, teratogenic: mouse, hamster, rat	Low birth weight, spontaneous abortions
Carbon monoxide	Fetotoxic, low birth weight, poor postnatal development and brain damage: rodent, rabbit, sheep, pig, monkey	Fetotoxic, low birth weight, fetal brain damage
PCB	Low birth weight, high perinatal and postnatal mortality, poor postnatal growth, skin discoloration: mouse, rat, rabbit, pig, dog, monkey Prolonged estrous cycle: rat Prolonged menstrual cycle: donkey Sc)ontaneous abortion: rhesus monkey	Low birth weight, high postnatal mortality, skin discoloration Menstrual disorders
Lead	Wide spectrum of effects: rats and m~ce, both sexes	Wide spectrum: both sexes
EDB	Sterility: rats, bulls	Reduced fertility in men
Carbon disulfide	Effects on spermatogenesis: rats Early embryonic mortality-increased congenital malformations: rat	Sperm abnormalities Spontaneous abortions: women

SOURCE: Adapted from S. M. Barlow and F. M. Sullivan, *Reproductive Hazards of Industrial Chemica/s* (London: Academic Press, 1982), p. 16; and 1. C. T. Nisbet and N, J. Karch, Chemica/ *Hazards to Human Reproducflon* (Park Ridge, NJ: Noyes Data Corp., 1963), p. 104.

Table 6-5.—Comparison of Reported Developmental Effects of 10 Agents in Humans and in Experimental Animals

Agent	Reported sites in humans	Reported sites in animals
Anesthetic gases	Hemangiomas, hernias, skin, heart	Skeletal defects only: rat, mouse (halothane and NzO)
Smelter emissions (lead and/or arsenic)	Multiple malformations	Multiple malformations: rat, mouse, hamster (lead and arsenic)
PBB	Skin discoloration; enlarged fontanelles	Skin discoloration and iesions: rhesus monkey; enlarged fontaneiies and syndactyiy: pig, dog; negative: rat, rabbit
Alcohol	Facial, CNS	Faciai, dermai, neural, extremities: rat, mouse
Vinyl chloride	Neural tube	Various, inciuding encephalocele: rat
Warfarin	Nose, bones (case reports only)	Negative: mouse, rabbit
Diphenylhydantoin	Cleft lip, cleft palate, other craniofacial, mental deficiency	Cleft lip, cieft paiate, syndactyly, other skeletai defects: mouse; minor kidney anomalies: rhesus monkey
Aminopterin	Multiple malformations	Multipie malformations: sheep, rat
Busulfan	Eye, cleft palate (1 report)	Skeletal, genitai defects: rat
Methotrexate	Skull, ribs, toes (2 reports)	Various: rat, cat, rabbit, mouse
Methylmercury	CNS	CNS, skeletai: rat, mouse, hamster, cat

SOURCE: 1. C. T. Nisbet and N. J. Karch, Chern/ca/ *Hazards to Human Reproduction* (Park Ridge, NJ: Noyes Data Corp., 19S3), pp. 97-98,

mal studies. This case also illustrates the kind of expense and level of research that may be required to determine whether substances are or are not harmful.

Dose-Response Considerations.—There is consensus among developmental biologists that thresholds do exist for the effects of toxic stimuli, unlike carcinogens (1,33). This assumption is based on biological considerations. First, the embryo has some capacity for repair of damaged tissues. Second, at early stages some systems are redundant; duplicate cells die if not used. Third, some cells have the ability to reprogram themselves. And finally, congenital abnormalities are multifactorial in nature; i.e., there is an interaction between genetic and environmental factors that determines whether an effect occurs. This can be illustrated by the action of factors causing cleft palate. Closure of the palate requires a critical balance between the size of the palatal shelves and the distance between them, which in turn depends on the width of the head and the time at which the shelves move up into the horizontal plane to fuse. If this balance is upset, either by altered tissue growth or by delay in movement of the shelves, closure of the palate may never occur (l).

In developmental toxicology testing, the assumption of threshold effects carries with it the determination of no observed effect levels (NOELs)[4] and calculation of margins of safetys or safety factors' in order to extrapolate developmental effects to humans. NOELS are difficult to establish. There is always a background rate of many of the endpoints; i.e., they occur naturally with a nonnegligible frequency. Other traits, such as the weight of an organ or birth weight, are continuously distributed. A value that represents a significant weight reduction or gain must be chosen in order to determine a NOEL. Using smaller sample sizes will yield larger NOEL values. The slope or steepness of the dose-response curve currently plays a small role in the determination of the NOEL. This curve may contain valuable information that is overlooked (6,8).

[4]Anima1s are treated at three dosage levels, a high dose that produces maternal toxicity, at least one intermediate dose, and a low dose that demonstrates **a NOEL. Determining a NOEL is a very complex procedure. Further discussion appears in 8, 12, 20, 22.**

'The margin of safety approach derives a ratio of the NOEL from the most sensitive species to the estimated human exposure level from all potentiai sources.

Whe safety factor approach is intended to derive a calculated exposure level that is unlikely to cause any developmental toxic responses in humans. The safety factor will vary depending on the agent, interspecies differences, and the slope of the dose-response curve. A safety factor of 100 is generally used, assuming a factor of 10 for species variability among test animals, and another 10 for animal-to-human differences. After the safety factor is selected, it is divided into the NOEL obtained from the most appropriate and/or sensitive animal species tested.

REPRODUCTIVE RESEARCH AND RISK ASSESSMENT ACTIVITIES IN GOVERNMENT AGENCIES

Discussion of reproductive research and risk assessment activities in government agencies will be confined to those of EPA, the Occupational Safety and Health Administration (OSHA), and NIOSH because this study focuses on occupational hazards. Research on reproduction in humans and toxicology testing and development of protocols, models, and guidelines is currently carried out in several government agencies.

Generally, OSHA does qualitative risk assessment for reproductive health hazards where data indicate the necessity. Risk assessment procedures have been made explicit in legal challenges to some standards that have been set by OSHA (see discussion in chapter 7). NIOSH, as the research and information support agency established by the OSH Act, is in the beginning phases of making risk assessment guidelines explicit, although it is carrying out research on reproductive impairment, NIOSH ranks disorders of reproduction as sixth of the 10 priority areas for research on work-related diseases and injuries (15).

EPA is currently engaged in developing guidelines for reproductive and developmental risk assessment and is also carrying out research on reproductive health hazards.

Environmental Protection Agency

Data Collection

As detailed in chapter 7, EPA obtains information on reproductive health hazards under a number of statutes. The submission requirements in most of the statutes place the burden of testing chemicals on industry. Under the Toxic Substances Control Act (TSCA), EPA receives basic data on the chemical identity of substances, their production volume, and worker exposure to the substances. The EPA Office of Toxic Substances also receives Premanufacture Notifications that help to determine the developmental (teratogenic) or mutagenic potential of proposed commercial substances. In addition, the agency receives notices when significant adverse reactions are observed in employees exposed to a substance and receives notices when substantial risks of significant environmental and health effects are observed by manufacturers.

EPA obtains data on pesticides under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). In order to collect testing information on environmental and human health effects of products not subject to recent review, EPA has implemented a program for reregistration of pesticide products "licensed" under FIFRA over the past 40 years. This program requires teratogenicity testing in two animal species generally rats and rabbits). The program also utilizes limited means of obtaining information on adverse health effects in workers.

EPA may also collect information on reproductive health hazards as part of the Clean Air Act, the Resource Conservation and Recovery Act, the Atomic Energy Act, and possibly Superfund.

In general, however, these laws provide very little basis for the ***swvstematic*** collection of reproductive health hazard data, and virtually no regulatory authority for monitoring or collecting information on toxic occupational exposures.

Data Bases

In addition to data handling submissions, EPA participates in several independent data collection activities. The most comprehensive data base is the Chemical Substances Information Network (CSIN), which was established under TSCA and is currently maintained by EPA and the Council on Environmental Quality. CSIN'S broad information base includes data on reproductive health hazards, structure, effects, uses, production, and pertinent regulatory requirements of many chemicals. Another data system, the Chemical Information System, maintained within the National Institutes of Health (NIH), contains the Scientific Parameters in Health and the Environment, which is a group of integrated data bases.

The Department of Energy's Oak Ridge National Laboratory provides internal data services on chemicals that are known or suspected reproduc-

tive health hazards via a data base called the Environmental Mutagen and Environmental Teratogen Information Center.

Internal EPA Research

The Health Effect Research Laboratory (HERL) in Research Triangle Park, North Carolina, provides research support for the Office of Research and Development's (ORD) reproductive health hazard assessments. Within HERL, the Developmental Biology Division conducts research in developmental toxicology and reproductive toxicology. For example, when there is disagreement concerning the toxicity of a particular substance being considered for regulation, the division will perform the research necessary to resolve the dispute. The division also reviews certain compounds for their reproductive effects. While the division does not perform risk assessments per se, it assesses the exposure of a compound, supplies input for risk assessment models, and makes recommendations concerning standards for a substance's continued use.

The agency also relies on CDC and FDA for research on reproductive health hazards. When specific substances are being considered for regulation, information on reproductive health hazards is exchanged under FIFRA and TSCA with OSHA and, to a more limited extent, with the Consumer product Safety Commission (CPSC). The National Toxicology Program, under the supervision of the Public Health Service in the Department of Health and Human Services (DHHS), in which EPA is a participating agency, may also provide assistance through its coordination and monitoring of interagency research, testing, and method development.

In some limited instances, EPA may employ outside contractors to perform certain tests to provide data necessary for risk assessments being performed by the agency.

Peer Review procedures

EPA risk assessments and the resulting regulatory decisions undergo peer review in several ways. At the request of agency officials, risk assessments performed within ORD are reviewed by professionals in the field both within and outside the Agency. Occasionally individuals in other agencies are informally requested to review ORD'S risk assessment work.

The second review method for risk assessment is through internal agency procedures and informal case-by-case referrals to different program offices. These are also not mandated by any particular statute. Red-border review[7] of regulatory actions is perhaps the most visible review of risk assessments within the agency. Before any regulatory proposal is published by EPA, a regulatory package is assembled by the program office with responsibility for the action and is distributed for review and approval to each assistant administrator in EPA.

Risk assessments are also reviewed on an informal basis within EPA by intra-agency task forces formed on a case-by-case basis to review particular chemicals. Risk assessments on reproductive health hazards are also regularly referred to the Developmental Biology Division in Research Triangle Park, North Carolina. For ionizing radiation, the agency has traditionally relied on periodic reviews conducted by the National Academy of Sciences at the agency's request. Finally, risk assessments are reviewed by independent advisory groups established pursuant to the environmental statutes themselves or to the Environmental Research and Development Act.

Assessment of Reproductive Health Hazards

Under TSCA and FIFRA, the Office of Pesticide Programs (OPP) and the Office of Toxic Substances (OTS) are responsible for analyzing the industry data submitted to EPA. .Risk assessments are performed in OPP by the Hazard Evaluation Division and in OTS by the Health and Environmental Review Division. These offices are staffed by toxicologists, biologists, and statisticians. Scientists working in one of these branches are sometimes unaware of work being done in their functionally equivalent branch.

---

[7] "Red border review" denotes intra-agency EPA procedures for the review of all agency rulemaking proposals by all assistant administrators in EPA. The term comes from the fact that these proposed regulatory actions are routed through EPA in red folders,

Other EPA program offices do not generally conduct their own risk assessments of particular substances. They rely instead on the Office of Health and Environmental Assessment in ORD if a risk assessment is required. An exception is the Office of Radiation Programs, which maintains it own health effects staff. In ORD, the Reproductive Effects Assessment Group (REAG), staffed by 15 scientists (reproductive and developmental toxicologists, epidemiologists, pharmacologists, biologists, and geneticists), conducts reproductive risk assessments for most program offices other than OPP and OTS. They also perform some risk assessments for OPP and OTS on a case-by-case basis. OPP and OTS risk assessments are generally reviewed by the Assistant Administrator of ORD only if a regulatory action is proposed and proceeds through red-border review. This is to assure consistency of all risk assessments done by EPA.

Risk assessment procedures for reproductive health hazards, while appearing to be fairly consistent among offices, are still perceived as problematic by the agency's officials.

EPA proposed Risk Assessment Guidelines

At the request of the former administrator, ORD is developing six specialized risk assessment guidelines: 1) mutagenicity, 2) developmental toxicology, 3) exposure, 4) carcinogenicity, 5) complex mixtures, and 6) male and female reproductive impairment. REAG has the responsibility for three: developmental toxicants, mutagens, and male/female reproductive effects.[8] REAG anticipates drafting the Male/Female Reproductive Effects Risk Assessment Guidelines by 1986.

In the developmental toxicology guidelines, EPA, for the most part, continues to recommend safety factors and margins of safety in risk assessment determinations, but acknowledges that more research needs to be done on mathematical modeling from dose-response curves. REAG and the Office of Research are currently developing methodology in this area. EPA officials expect the guidelines to be constantly revised as new advances are made in the science.

REAG staff have also been contributing developmental toxicology and reproductive toxicology guidelines to the Interagency Risk Management Council. (Member agencies include the Food and Drug Administration, the U.S. Department of Agriculture, NIH, OSHA, CPSC, and EPA). The goal of this council is to attempt the drafting of consistent policies across all executive regulatory agencies, This effort had been expected to take 2 years, but is now stalled because of a lack of resources.

Conclusions

EPA's collection of data and research on reproductive health hazards appears disjointed. Probably because of programmatic divisions within the agency, data developed under one statute are often not routinely shared with offices carrying out other statutory responsibilities. Although this may be a consequence of the fact that EPA operates under several different legislative mandates, it may inhibit regulatory consideration of chemicals with potential for reproductive effects in different exposure situations that are covered by different mandates. It may also lead to duplication of internal and external testing.

Data retrieval systems appear to offer one avenue for the coordination of this information. One system, the Status Report of Chemical Activities published through the Toxics Information Series, is a particularly useful model in this regard. The status report lists, by chemical, testing being performed on a particular substance, the statutory authority under which it is being performed, and a contact person within the agency. It also indicates whether a regulatory action is being contemplated or has been taken.

[8] Four **of the proposed guidelines were published in the Federal Register, vol. 49, No. 227, Nov. 23, 1984: Carcinogen Risk Assessment, p. 46294; Exposure Risk Assessment, p. 46304; Mutagenicity Risk Assessment, p. 46314; and Health Assessment of Suspect Developmental Toxicants**, p. 46324.

NATIONAL INSTITUTE FOR OCCUPATIONAL SAFETY AND HEALTH

NIOSH is the research agency created by the OSH Act of 1970. NIOSH is a part of the CDC, which is a part of the Public Health Service which, in turn, is a part of DHHS. The director of NIOSH is appointed by the Secretary of HHS for a term of 6 years. NIOSH has no authority for promulgating or enforcing standards (risk management) but is responsible for conducting research and making recommendations to the Department of Labor pursuant to the OSH Act and the Federal Mine Safety and Health Act.

NIOSH research may begin at the urging of the Secretary of HHS, or on the initiative of the Director of N1OSH. An employer or employee request may also lead to a safety and health evaluation. In all its activities, NIOSH approaches the development and evaluation of standards with the intent of providing optimum protection for employees, whereas OSHA'S mandate is to examine the potential costs and benefits (see chapter 7).

NIOSH has responsibility for several major activities:

1. develop criteria for recommended occupational safety and health standards,
2. conduct educational programs to provide an adequate supply of qualified personnel,
3. conduct informational programs on the importance of the use of adequate safety and health equipment,
4. conduct Health Hazard Evaluations, and
5. conduct industrywide studies of the effects of chronic or low-level exposures.

NIOSH has been criticized from several directions. OSHA has criticized it for the inadequacy of criteria documents for OSHA standard-setting. The General Accounting Office has criticized the quality of its criteria documents and Health Hazard Evaluation program. Labor groups have stated that it is unresponsive to worker requests. Management representatives have claimed that Health Hazard Evaluations are too aggressively pursued, and NIOSH research is of poor quality (for further discussion, see (29)). Recent directors of NIOSH have worked to improve the quality of NIOSH research.

Reproductive Health Hazard Research

Former and current NIOSH officials agree that NIOSH has been slow to study reproductive health hazards. This has been due, in part, to budgetary and personnel problems. In the last few years the issue of reproductive health hazard research has received higher priority (so). Current research activities are listed in table 6-6.

NIOSH has pursued several approaches for studying the adverse effect of occupation on human reproductive systems. First, NIOSH has accessed several large data bases that include information on occupation and has linked these data with State or city vital statistic and birth records, permitting an analysis that attempts to determine whether adverse pregnancy outcomes are associated with specific types of occupations. Second, NIOSH has been investigating the effects of specific exposure on both female and male reproductive function.

To study the effects on the female reproductive system, information on pregnancy outcomes from State or city records or information on pregnancy outcomes from a questionnaire administered to the mother is obtained and analyzed to determine if specific occupational exposures are associated with adverse pregnancy outcomes such as miscarriage, low birthweight, or malformations,

To study the effects on the male reproductive system, one of two strategies has been used: 1) a similar approach to the one described for study of effects following female exposure, except that the analysis determines whether adverse pregnancy outcomes of spouses are associated with specific occupational exposures of males; and 2) an evaluation of specific semen quality parameters. The parameters considered include sperm count, sperm motility, sperm morphology, and specific hormone activity. The meaning of these semen quality parameters in terms of actual adverse pregnancy outcomes is not known at present, but the study of these parameters is believed to document the effects of specific exposures.

Table 6-6.—NIOSH Reproductive Health Hazards Research

Subject of study/ suspected hazard	Status of research/ workers studied	As of Aug. 1, 1984
1. Oryzalin	Males	Completed
2. Carbon disulfide	Males and male workers' wives	Completed
3. Organic compounds (wastewater treatment workers)	Males	Completed
4. PCBS	Females	Completed
5. heavy metals (uranium workers)	Male workers' wives	Completed
6. DBCP	Males	Completed
7. Pharmaceutical estrogen	Males	Completed
8. Pharmaceutical lab workers.	Females	Completed
9. EDB (2 studies)	Males	1 completed 1 in progress
10. Lead	Males	Nearly completed
11. Chemotherapeutic drugs	Females	1 study completed, hazard alert in preparation
12. Glycol ethers	Males	Field work completed, analysis in progress
13. Human semen characteristics.	Male	Proposed
14. VDTS	Females	In progress
15. Dioxin	Males	Development stage
16. Ethylene oxide	Males and females	Proposed
17. Organo-tin compounds	Males	Interest
18. Butadiene	Males	Interest
19. Radiofrequency	Females	Abandoned (problem with cohorts) (but being reactivated)

NOTE This list excludes some reports of health hazard evaluations based on clusters of negative reproductive outcomes (e.g., spontaneous abortions).

SOURCE: Office of Technology Assessment

With respect to developmental toxicology, NIOSH has been conducting research on the effects of chemicals on the offspring of laboratory animals (rats) exposed during gestation. The tests used to determine the developmental effects examine both instinctive and spontaneous behavior. Using these study designs, NIOSH has studied several glycol ethers and industrial alcohols. The findings have shown that behavioral effects in the offspring can appear in the absence of other signs of toxicity in both the dam and the offspring.

NIOSH has a collaborative effort with the National Toxicology Program to test dose-response characteristics of selected chemicals for reproductive toxicity (30)

Reproductive Risk Assessment

Since NIOSH is a scientific and technical research agency, it approaches health hazard control with the view of providing maximum protection for workers. Thus, although it does not determine whether a risk is "significant" in the legal sense, it does attempt to quantify the magnitude of risk. Because the courts are requiring that OSHA standards contain increasingly detailed risk assessment, NIOSH has just initiated a formal section for quantitative risk assessment in the criteria documents division. Because the agency currently has little expertise in this field, it is working with consultants to develop the capability to better quantify the need for standards. One of the goals of the new section is to develop working groups in various subject areas and, where needed, to use outside experts to assist with risk assessments.

Exposure Estimates

NIOSH is in the process of surveying industries in order to estimate the numbers of individuals exposed to hazards, In contrast to an earlier survey, this is a representative sample of establishments selected from Dun & Bradstreet files. Sup-

elementary samples of establishments from other files have been selected for the Standard Industrial Classifications determined to be inadequately covered by Dun & Bradstreet. The sample of establishments will constitute an unbiased random sample of industries in the United States. The sample design is based on a decision to maximize the reliability of estimates of numbers of employees exposed to hazards. Estimates by industry or estimates of the number of firms with hazards have been assigned lower priority. Information will be available by sex but not by age. Some data and tabulations are expected to be available by late 1985 (9)25).

Conclusions

Although NIOSH is carrying out a fair amount of research on reproductive health hazards, it lags behind the efforts of EPA in the development of reproductive and developmental risk guidelines. It is increasing this latter capability in response to court challenges of OSHA standards.

CHAPTER 6 REFERENCES

1. Barlow, S. M., and Sullivan, F. M., ***Reproductive Hazards of Industrial Chemicals (New*** York: Academic Press, 1982).
2. Bloom, A.D. (cd.), ***Guidelines for Studies of Human Populations Exposed to Mutagenic and Reproductive Hazards,*** Proceedings of a Conference held Jan. 26-27, 1981, Washington, DC, March of Dimes Birth Defects Foundation, White Plains, NY, 1981.
3. Brown, N. A., Kao, J., and Fabro, S., "Teratogenic Potential of Valproic Acid," ***Lancet*** 1:660-661, 1980.
4. Christian, M. S., Galbraith, W. M., Voytek, P., and Mehlman, M. A., "Assessment of Reproductive and Teratogenic Hazards," ***Advances in Modern Environmental Toxicology, vol.*** HI (Princeton, NJ: Princeton Scientific Publishers, Inc., 1983).
5. Coodley, C., "Risk in the 1980's: New Perspectives on Managing Chemical Hazards," San ***Diego L, Rev.*** 21(5):1015-1044, September-October 1984.
6. Crump, K. S., "A New Method for Determining Allowable Daily Intakes," ***Fund. and Applied Toxi. 4:854-871, 1984.***
7. Deisler, P. F., Jr,, "A Goal~riented Approach to Reducing Industrially Related Carcinogenic Risks," ***Drug Metabolism Reviews 13(5):875-91*** 1, 1982.
8. Gaylor, D. W., "The Use of Safety Factors for Controlling Risk," ***J. Toxicol. Environ. Heahh 11:329-336, 1983.***
9. Hanson, R., and Edmonds, J., ***Sample Design of the National Occupational Hazard Survedv II (NOH II),*** draft report, contract No. 210-80-0057, submitted to NIOSH, Cincinnati, OH, Westat, Rockville, MD, January 1981.
10. Hatch, M., Stefanchik-Scott, V,, and Stein, Z. A., "Surveillance of Reproductive Health in the U. S.: A Survey of Activity Within and outside Industry)" unpublished, prepared for the American Petroleum Institute, December 1983.
11. Health and Welfare Canada, ***The Testing of Chemicals for Carcinogem"city, Mutagenicity, and Teratogenicitey,*** published by authority of the Honorable Marc Lalonde, Minister of Health and Welfare (Ottowa, Ontario: 1975).
12. Kimmel, C. A., ***Reliability of Experimental Studies for Predicting Hazards to Human Development,*** National Center for Toxicological Research, Technical Report for Experiment No. 6015, Jefferson, AR, January 1984.
13. Kline, J., Stein, Z., Strobino, B., Susser, M., and Warburton, D., '(Surveillance of Spontaneous Abortions: Power in Environmental Monitoring, " ***Prev. Med. 9:267-274, 1977.***
14. Miller J. F., Williamson, E., Glue J., Gordon, Y. B., et al., "Fetal Loss After Implantation: A Prospective Study," ***Lancet 11:554-556, 1980.***
15. ***Morbidity and Mortality Weekly Report,*** "Leading Work-Related Diseases and Injuries—United States," 34(35):537-540, Sept. 6, 1985.
16. National Academy of Sciences, ***Principles for Evaluating Chemicals in the Environment,*** Washington, DC, P72, 1975.
17. National Academy of Sciences, ***Principles and Procedures for Evaluating the Toxicitbv of Household Substances*** (Washington, DC: National Academy Press, 1977).
18. National Academy of Sciences, ***Risk Assessment in the Federal Government: Managing the Process*** (Washington, DC: National Academy Press, 1983).
19. National Academy of Sciences, ***Toxicity Testing, Strategies to Determine Needs, and Priorities*** (Washington) DC: National Academy Press, 1984).
20. Nisbet, I.C.T., and Karch, N. J., Chemical***Hazards to Human Reproduction*** (Park Ridge, NJ: Noyes Data Corp., 1983).
21. Organization for Economic Coo~eration and De-

velopment, OECD ***Short-Term and Long-Term Toxicology Groups,*** final report, Dec. 31, 1979

22 Robert, E., and Guibaud, P., "Maternal Valproic Acid and Congenital Neural Tube Defects, " ***Lancet 11:937-38, 1982***

23. Rodricks, J.~~., and Tardiff, R. C., "Risk Assessment: How It's Done, and How It Might Be," *Chemtech.* 14:394-397, July 1984

24. Rosenberg, M. J., Feldblum, P. J., Shy, C. M., and Marshall, E. G., "Epidemiologic Surveillance of Occupational Effects on Reproduction)') OTA contract report, 1984.

25. Seta, J., personal communication, NIOSH, 1985.

26. Sever, L. E., and Hessol, N. A., "Overall Design Considerations in Male and Female Occupational Reproductive Studies," ***Reproduction: The New Frontier in Occupational and Environmental Health Research (New*** York: Alan R. Liss, Inc., 1984), pp. 15-47.

27. Stein, Z., Hatch, M., Kline J., and Warburton, D., "Epidemiologic Considerations in Assessing Health Effects at Toxic Waste Sites," ***Assessment of Health Effects at Chemical Disposal Sites,*** Proceedings of a symposium held June 1-2, 1981, New York, W. W. Lowrance (cd.) (New York: The Rockefeller University Press, 1981), pp. 126-145.

28. U.S. Congress, Office of Technology Assessment, ***The Information Content of Premanufacture Notices–Background Paper, OTA-BP-H-17*** (Washington, DC: U.S. Government Printing Office, April 1983).

29, U.S. Congress, Office of Technology Assessment, ***Preventing Illness and Injury in the Workplace, OTA-H-257*** (Washington, DC: U.S. Government Printing Office, April 1985).

30. U.S. Department of Health and Human Services, National Institute of Occupational Safety and Health, ***Program of the National Institute for Occupational Safety and Health, for Fiscal Years 1984-89,*** DHHS (NIOSH) publication No. 84-107, February 1984.

31 U.S. Environmental Protection Agency, "Proposed Guidelines for Registering Pesticides in the U. S., Hazard Evaluation: Humans and Domestic Animals," ***Federal Register 43:37335-37402, 1978.***

32. U.S. Environmental Protection Agency, "Proposed Guidelines for Registering Pesticides in the U. S.: Hazard Evaluation: Humans and Domestic Animals," ***Federal Register 45:50544-50546, 1980.***

33. U.S. Environmental Protection Agency, "Guideline for the Health Assessment of Suspect Developmental Toxicants, " draft, October 1984.

34 U.S. Food and Drug Administration, "Advisory Committee on Protocols for Safety Evaluation: Panel Report on Reproduction Studies in Safety Evaluation of Food Additives and Pesticide Residues," ***Toxicol. Appl. Pharmacol. 16:264-296, 1970.***

35. Warburton, D., "Selection of Human Reproductive Effects for Study," ***Assessment of Health Effects at Chemical Disposal Sites,*** Proceedings of a symposium held on June 1-2, 1981, New York, William W. Lowrance (cd.) (New York: Rockefeller University Press, 1981), pp. 106-124.

36. Witshi, E., "Teratogenic Effects From Overripeness of the Egg, " ***Congenital Malformation,*** F.C. Fraser and V.A. McKusick (eds.) (Amsterdam: Excerpta Medica, 1970), pp.157-169.

37. World Health Organization (WHO),***Principles for the Testing of Drugs for Teratogenicity,*** ITS Tech. Ser. No. 364 (Geneva: WHO, 1967).