#86-05

# Improving Assessment of the Effects of Environmental Contamination on Human Reproduction

## BACKGROUND REPORT

The Working Group on Human Reproductive Outcomes

Child Trends, Inc.
2100 M Street, NW
Washington, DC

June, 1986

# Improving Assessment of the Effects of Environmental Contamination on Human Reproduction

The Working Group on Human Reproductive Outcomes
Child Trends, Inc.
June, 1986

BACKGROUND REPORT

## INTRODUCTION

Public concern with potential environmental hazards is high. Almost daily there are newspaper accounts of chemical spills, leaks of dangerous gases or radiation, improperly disposed waste materials, air pollution, and contaminated water. There are also frequent allegations of excessive numbers or clusters of poor reproductive outcomes, such as miscarriages or certain birth defects, in particular communities or populations.

The recognition in the 1960's of a connection between limb reduction deformities and thalidomide taken early in pregnancy created widespread anxiety about other drugs or environmental agents that might produce birth defects (Edmonds et al, 1981). The proliferation of modern chemicals, drugs, electronic devices, and other conveniences designed to make life more liveable has stimulated continuing concern over the potential consequences of human ingenuity. For example, more than 60,000 chemicals are in common use, yet most of them have not been tested for their effects on human reproductive outcomes.

The association between environmental contamination and health effects are often unclear, however, and concerned citizens frequently request that studies be done to determine whether environmental pollution is causing health problems in their communities. The Environmental Protection Agency (EPA), the agency generally turned to by concerned citizens, currently has no systematic procedures or guidelines for coping with these requests. Such requests are apt to be received by regional offices, and these offices usually act autonomously in an ad hoc manner in responding to them. The lack of coordination between regional and federal offices, as well as the ad hoc approach to these serious health questions, have been a source of dissatisfaction both within EPA and to communities around the United States.

## THE WORKING GROUP ON HUMAN REPRODUCTIVE OUTCOMES

In 1985, EPA and the National Science Foundation provided grant funds to Child Trends, a non-profit research organization in Washington, D.C., to form a Working Group on Human Reproductive Outcomes. The purpose of the Working Group was to help EPA consider steps it might take to strengthen its capability for assessing reproductive effects of exposure to

environmental contamination and determining whether clusters of adverse outcomes exist and can be attributed to environmental contamination.

The Working Group on Human Reproductive Outcomes was composed of a diverse set of specialists from a variety of fields, including: epidemiology, human genetics, obstetrics, pediatrics, reproductive toxicology, survey research, and sampling statistics. A number of federal agencies were also represented at the Working Group meetings, including the Environmental Protection Agency, the National Center for Health Statistics, the Centers for Disease Control, and the Division of Maternal and Child Health of the U.S. Public Health Service. Members of the Working Group and agency representatives are listed at the end of this report.

The charge to the panel was to focus on the statistical data needed to clarify the association between environmental contamination and undesirable reproductive outcomes. More specifically, the Group was asked to suggest ways to improve data on poor reproductive outcomes. They were not asked to deal with the equally vexing problem of improving data on human exposure to environmental contaminants.

Undesirable reproductive outcomes include the following:

- non-voluntary infertility or subfecundity; (Infertility is the inability to conceive a pregnancy after one year of unprotected intercourse. Subfecundity is a more general term that includes couples for whom it may be difficult, but not impossible, to conceive.)

- miscarriages/spontaneous abortions; (These terms refer to pregnancies that end in embryonic or fetal death before the 20th or 28th week of the pregnancy, depending upon the definition.)

- stillbirths; (This term refers to pregnancies that end in a fetal death after the 20th or 28th week of pregnancy, depending upon the definition.)

- birth defects in live or stillbirths; (This refers to any of a number of abnormal conditions ranging from major chromosomal disorders, such as Down's syndrome, to minor abnormalities such as birthmarks.)

- low birthweight, prematurity, intrauterine growth retardation, and other signs of difficulties, such as low Apgar scores, in live births; (Infants weighing less than 2,500 grams [or about 5 1/2 pounds] at birth are said to be low birthweight infants. Infants born live prior to 37 weeks of pregnancy are said to be premature. Babies who are small for

their gestational age are said to have experienced intrauterine growth retadation. Low birthweight is typically a symptom of both prematurity and intrauterine growth retardation. Apgar scores, named for Dr. Virginia Apgar, who devised these scales, are assessments of the infant's heart rate, respiratory rate, muscle tone, cry, and color at birth to summarize the baby's physical condition. The scores are typically taken at 1 and 5 minutes after birth.)

The decision to focus on undesirable reproductive outcomes was made for several reasons. First, reproductive health problems have received considerably less attention at EPA than cancer, respiratory diseases, and other disorders that can be associated with environmental contamination. Yet, as noted above, reproductive problems and birth defects are of great concern to the public. All of these negative outcomes pose emotional costs to the individuals affected, and often entail considerable financial burdens as well. In addition, birth defects carry substantial societal costs.

Apart from recognizing the costs of the reproductive problems themselves, improved tracking of changes in reproductive outcomes over time and across geographical areas could prove to be advantageous in alerting the public to the development of dangerous environmental conditions. This is because many agents that are associated with reproductive problems can also be expected to be associated with other health effects. And the time period between exposure and the occurrence of a miscarriage or birth defect is usually shorter than the latency period for the development of diseases such as cancer. Thus, an effective means of detecting changes in reproductive outcomes could serve as an early warning system for other health hazards as well. Of course, in order to understand the causes of any changes in rates of undesirable reproductive outcomes, data on exposures to environmental contaminants, occupational hazards, drugs, etc., would also be needed.

## EPA'S NEEDS AND THE DIFFICULTIES OF OBTAINING ADEQUATE DATA

In order to be able to determine whether negative reproductive outcomes are unusually high in a given area, EPA needs good national baseline data on the outcomes of interest, data that can be analyzed according to a variety of individual characteristics such as race and age. Good baseline data do not currently exist, however. Measurement problems are a major reason for the lack of high quality data. The following pages summarize much of what is known about the frequency and causes of birth defects and other negative reproductive outcomes, as well as the reasons for the difficulties in obtaining good baseline data.

## THE STATE OF OUR KNOWLEDGE ABOUT BIRTH DEFECTS

Although birth defects loom large for those who are affected, in actuality specific types of major defects are uncommon. (See Table 1 for various estimates of the frequency of occurrence of ten specific types of birth defects.) Taken as a group, however, the number of individuals affected is not small. A recent National Institutes of Health report estimates that major defects occur with 3 to 5 percent of births (National Institutes of Health, 1979). Thus, between 110,000 and 184,000 of the 3,680,577 babies born in 1982 in the U.S. had a serious birth defect (Baldwin and Nord, 1984). Some of these were genetic disorders, such as cystic fibrosis, muscular dystrophy, and sickle cell anemia; some were congenital malformations, of which spina bifida is probably the most familiar; and some were Down's syndrome and other lesser known chromosomal abnormalities. The rate of defects nearly doubles by age one because of subsequent discoveries of serious, but not immediately detectable, effects (Jason and Glick, 1981). The rarity of many of the specific defects, however, makes it difficult to detect changes in their incidence and variability across a population.

Just because specific birth defects are rare, however, does not mean that birth defects have trivial consequences for society. Congenital malformations, for example, are the leading cause of infant mortality in developed countries (Edmonds et al, 1981). Hereditary diseases and congenital malformations account for at least one-fifth of all infant deaths in the U.S. (National Institutes of Health, 1979). In addition, they are a large component of child hospitalizations and sickness, and they have substantial economic consequences (Edmonds et al, 1981). Some researchers estimate that future costs of custodial care only for children born in 1978 with serious birth defects could run more than two billion dollars (Selle et al, 1979). Furthermore, chromosomal aberrations and developmental defects also account for a major proportion of spontaneous abortions (National Institutes of Health, 1979). The impact of a birth defect on family functioning and the well-being of families can be substantial (Earhart and Sporakowski, 1984).

Specific causes of birth defects are not well known. Genetic causes probably account for about one-quarter of them; another 10 percent are known to be related to viruses, drugs, radiation, and chemicals; but the remaining 65 percent are unexplained (Jason and Glick, 1981).

TABLE 1:

BIRTH DEFECTS*

Type of Defect

| | Anencephaly | Spina Bifida | Hydrocephaly | Cleft Lip (+/- Cleft Palate) | Cleft Palate | Rectal/Anal Atresia | Hypospadias | Limb Reduction | Congenital Hip Dislocation | Down's Syndrome |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Rate Per 10,000 Live Births | | | | | | |
| Range of Estimates** | 2.4-8.6 | 4.5-9.0 | 3.6-12.9 | 7.8-18.0 | 4.7-12.9 | 2.1-4.1 | 11.2-32.6 | 3.3-14.3 | 5.7-26.1 | 5.6-15.0 |
| Source of Estimates | | | | | | | | | | |
| BDMP 1970-81(1) | 4.3 | 5.9 | 4.4 | 14.8 | 10.7 | 3.4 | 23.2 | 3.3 | 20.6 | 8.0 |
| 1970-83(1) | 4.1 | 5.7 | 4.6 | 14.7 | 10.7 | 3.3 | 24.1 | 3.4 | 21.9 | 8.0 |
| NCHS 1973-1978(1) | 2.0 | 3.3 | 1.7 | 10.2 | 6.9 | 1.6 | 5.9 | 2.3 | 1.6 | 3.7 |
| Md (9/83-8/84)(2) | 2.6 | 4.5 | 5.0 | 5.1 | 4.7 | 2.1 | 11.2 | 14.3 | 5.9 | 7.2 |
| Missouri (1979-80)(3) | | | | | | | | | | |
| Multisource | 3.2 | 5.6 | | 14.8 | | | | | 19.4 | 5.9 |
| Hospital Discharge | 2.6 | 5.0 | | 13.3 | | | | | 26.1 | 5.6 |
| Nebraska(1) | 4.9 | | 5.0 | 10.3 | 11.9 | 2.9 | 12.8 | 3.6 | 8.0 | 8.2 |
| BC Canada 1980(4) | 2.5 | 5.5 | 9.8 | 7.8 | 12.3 | 3.5 | 24.8 | 5.3 | 42.1 | 10.5 |
| MACDP 1968-81(1) | 6.9 | 9.0 | 10.6 | 17.7 | 12.9 | 4.2 | 25.8 | 7.4 | 10.8 | 9.8 |
| 1968-83(1) | 8.6 | 8.6 | 10.5 | 18.0 | 12.9 | 4.1 | 25.8 | 7.0 | 10.3 | 9.8 |
| 1977-81(1) | 6.7 | 6.7 | 12.9 | 16.5 | 12.3 | 3.8 | 28.4 | 5.9 | 13.0 | 9.8 |
| 1979-83(1) | 6.3 | 6.3 | 11.6 | 18.2 | 11.9 | 3.6 | 26.8 | 5.4 | 10.0 | 10.9 |
| Collaborative Perinatal (1) | 6.2 | 7.3 | 14.1 | 10.5 | 11.3 | | 40.7 | 20.0 | 23.5 | 10.7 |
| Mayo Clinic 1951-63(5) | 2.4 | 6.5 | 3.8 | 11.4 | 10.9 | 5.7 | 32.6 | 4.7 | 5.7 | 14.7 |
| Stein & Warburton(6) | 10.4 | | 3.6 | 7.8 | 3.3 | | 26.4 | | 7.8 | 15.3 |
| Hook & Fabia (1958-65, Mass)(7) | | | | | | | | | | 15.0 |
| Hook & Lindsjo (1968-70, Sweden)(8) | | | | | | | | | | 13.2 |
| Gortmaker(9) | | 7.0 | | 16.0 | | | | | | 14.0 |
| Ireys(10) | | 5-13 | | 8-16 | | | | | | |

*Definitions of specific birth defects may vary slightly from source to source.

**Range was determined after dropping highest and lowest value in each column.

## Notes to Table 1

1. Compiled by the Birth Defects Branch, Chronic Diseases Division, Center for Environmental Health, Centers for Disease Control.

2. Maryland Birth Defects Reporting and Information System

3. Missouri Birth Defects Registry

4. Province of British Columbia, Ministry of Health. Division of Vital Statistics. "Health Surveillance Registry: Annual Report." 1981. H.S.R. No. 6. Chronic Disabilities, Congenital Anomalies, Genetic Defects. Printed by Ministry of Health Printing Services. Victoria, 1981.

5. Lloyd Harris, Lois A. Stayura, Perta Ramirez-Talavera, and John F. Annegers. "Congenital and Acquired Abnormalities Observed in Live-Born and Stillborn Neonates." Mayo Clinic Proceedings 50(2):85-90, 1975. Table 1, pp. 86-88.

6. Zena Stein and Dorothy Warburton. "Report of Panel II. Guidelines for Reproductive Studies in Exposed Populations." In Arthur D. Bloom, ed. Guidelines for Studies of Human Populations Exposed to Mutagenic and Reproductive Hazards. White Plains, NY: March of Dimes Birth Defects Foundation, 1981. Table 8, p. 80.

7. Ernest B. Hook and Jacqueline J. Fabia. "Frequency of Down Syndrome in livebirths by single year maternal age interval: Results of a Massachusetts study." Teratology 17(3):223-228, 1978. Table 1, p. 225.

8. Ernest B. Hook and Agneta Lindsjo. "Down Syndrome in live births by single year maternal age interval in a Swedish study: Comparison with results from a New York State study." American Journal of Human Genetics, 30(1):19-27, 1978. Table 1, p. 20.

9. Steven L. Gortmaker and William Sappenfield. "Chronic childhood disorders: Prevalence and impact." In Robert J. Haggerty, M.D., Guest Editor, Pediatric Clinics of North America 31(1):3-18, 1984. Table 1, p. 5.

10. Henry T. Ireys. "Health care for chronically disabled children and their families". In Better Health for Our Children: A National Strategy. The Report of the Select Panel for the Promotion of Child Health. Volume IV, background papers, pp. 321-353, 1981. p. 324.

## THE STATE OF OUR KNOWLEDGE ABOUT OTHER REPRODUCTIVE OUTCOMES

Birth defects are only one manifestation of toxic effects on humans. We know very little of the effects of drugs, chemicals, and environmental contaminants on the whole human reproductive process. Infertility, subfecundity, spontaneous abortion, and stillbirth, may also be consequences of exposure to environmental hazards (Jason and Glick, 1981; Wilcox, 1983; Baird et al, 1985). The problem is to distinguish natural occurrences of these events from occurrences due to exposure to hazardous elements in the environment. (See Table 2 for estimates of the incidence or prevalence of reproductive impairments and negative outcomes other than birth defects in human populations.)

Non-voluntary infertility/subfecundity. Infertility or subfecundity is estimated to affect between 8 and 15 percent of American couples (Daniels and Weingarten, 1979; National Center for Health Statistics, 1985). Stein and Warburton (1981) estimate that 10-15 couples out of every 100 are unable to conceive after one year of unprotected intercourse. Mosher, using data from Cycle III of the National Survey of Family Growth (NSFG), estimated that in 1982, 8.5% of currently married couples with wives 15-44 years of age were infertile -- that is, they had difficulty conceiving. When he included couples where it was difficult or dangerous for the wife to carry a pregnancy to term, the figure rose to 10.8% (Mosher, 1985). Couples in this more inclusive category are said to have impaired fecundity.

Although the data from the NSFG suggest that there has been a decrease in fecundity impairments since 1976 (from 15.7% to 10.8%), the explanation for the change probably lies in the increase in the proportion of couples who are contraceptively sterile. When Mosher excluded them from his calculations, he found no significant changes in impaired fecundity between 1976 and 1982 (Mosher, 1985).

The NSFG also shows that a higher proportion of blacks than whites are classified as infertile -- 13.1% in 1982 compared to 8.1% for whites (NCHS, 1985a). Since 1965 there has been a significant increase in infertility among young black women aged 15-29 and a significant decrease in infertility among black women aged 30-44 (NCHS, 1985a).

There are also large differences in the prevalence of infertility and impaired fecundity by women's parity. For currently married couples in 1982 with wives 15-44 years of age, Mosher estimated that 8.5% were infertile. The equivalent figures for parity 0, 1, 2, and 3 or more were 19.6, 10.8, 5.0, and 3.8, respectively (Mosher, 1985). All these figures are slightly higher for women with impaired fecundity.

Table 2

Estimates of the Incidence or Prevalence
of Various Reproductive Outcomes
(Other than Birth Defects)

| Type of Outcome | Incidence or Prevalence | Source |
|---|---|---|
| Impaired fecundity (Difficulty in conceiving & difficulty or danger in carrying pregnancy to term | 10.8% of married couples couples with wives 15-44 years of age | 1982 National Survey of Family Growth (Mosher, 1985) |
| Infertile (Difficulty in conceiving or failure to conceive) | 8.5% of married couples couples with wives 15-44 years of age | 1982 National Survey of Family Growth (Mosher, 1985) |
| Miscarriage | 13.2% of recognized pregnancies | 1982 National Survey of Family Growth (Pratt et al, 1984) |
| Stillbirth | 2-4% of pregnancies that survive to 28 weeks of gestation | Stein & Warburton (1981) |
| | 9.0 fetal deaths of 20 weeks gestation or older per 1,000 live births | Vital Statistics of the U.S. (1981) |
| Live birth of low birth weight | Less than 2,500 grams: 6.7% of all live births (5.6% of White births; 12.4% of Black births, | Advance Report of Final Natality Statistics, 1984 (NCHS, 1986) |
| | Less than 1,500 grams: 1.2% of all live births (0.9% of White births; 2.6% of Black births) | |
| Live birth with Low Apgar Score | 5 minute Apgar score of less than 9: 12.4% of all live births in 46 reporting States & Washington, D.C. (11.9% of White births; 14.7% of Black births) | Advance Report of Final Natality Statistics, 1984 (NCHS, 1986) |
| | 5 minute Apgar score of less than 8: 3.9% (3.5% of White births; 5.7% of Black births) | |

Table 2 (Cont'd.)

Estimates of the Incident or Prevalence
of Various Reproductive Outcomes
(Other Than Birth Defects)

| Type of Outcome | Incidence or Prevalence | Source |
|---|---|---|
| Multiple births | Twin deliveries: 1.99% of all live births (1.93% of White births; 2.38% of Black births)<br><br>Triplet & other plural deliveries: 0.045% of all live births (0.048% of White births; 0.035% of Black births) | Advance Report of Final Natality Statistics, 1984 (NCHS, 1986) |
| Sex ratio | 1,050 male births per 1,000 female births (Whites: 1,054 males per 1,000 females; Blacks: 1,031 males per 1,000 females) | Advance Report of Final Natality Statistics, 1984 (NCHS, 1986) |

The above figures only approximate the proportion of all women interviewed in a survey who might report that they were infertile or had impaired fecundity. Cycle III of the NSFG was the first survey in that series to interview unmarried as well as married women. Thus, estimates of the prevalence of impaired fecundity among women 15-44 who had ever had sexual intercourse can be made. Of course, women who have always used contraception or who have never tried to become pregnant will be less likely to know whether they have reproductive impairments and thus these estimates are likely to understate the true extent of impaired fecundity. Mosher (1985) found 8.4% reported impaired fecundity. The figure was the same for 0 parity women and only slightly higher (8.5%) for women of 1 or higher parity. Since approximately 86% of the women in the survey had ever had intercourse (Pratt et al, 1984), one might expect 7.2% (=8.4 x .86) of all women aged 15-44 in a survey to report impaired fecundity.

Of couples who seek help from specialists, about 10 percent are infertile for no apparent reason; about half will conceive; and the remaining 40 percent have some problem that permanently prevents them from having children (Daniels and Weingarten, 1979). About 40 percent of diagnosable fertility problems can be attributed to the man, 20 percent to both partners, and the remaining 40 percent to the woman. The most common problems with male infertility are low-sperm count, blocked sperm ducts, and defective sperm. Women's infertility can be due to a wide range of problems. It is not clear to what extent exposure to environmental hazards is reflected in infertility problems or if there are interactions between subtle problems, for example hormonal imbalances, and exposure to environmental contaminants.

Spontaneous abortions/miscarriages. Estimates of the occurrence of spontaneous abortions and miscarriages vary greatly depending upon how soon after conception the counting of losses begins and how far into the pregnancy the counting of losses continues. James using data from earlier studies estimates that approximately 49% of all conceptions perish naturally between fertilization and confinement (James, 1970). He estimates that approximately 35% of all conceptions are lost before the first missed period and an additional 14% of recognized pregnancies are lost. Stein and Warburton estimate that 10 to 20 percent of pregnancies spontaneously abort between the eighth and twenty-eighth week of gestation (Stein & Warburton, 1981).

The NSFG data indicate that 13.2% of the pregnancies which occurred in the last five years ended in miscarriage or stillbirth (Pratt et al, 1984). This figure agrees well with the estimate by James. An important question if a survey approach is adopted, however, is how many pregnancies and miscarriages might be expected in a more delimited time period such as one year. The staff of the NSFG consider the estimates based on the 5 year period to be good, but due to the complex survey design, one year estimates would be unstable.

As pointed out above, chromosomal aberrations and developmental defects account for a major proportion of spontaneous abortions. It is estimated that chromosomal abnormalities occur in one of every 250 conceptions. About 88 percent of these, however, abort spontaneously, often before the woman even knows she is pregnant (Nortman, 1974). About one-third of all conceptions end in unrecognized losses and for every recognized spontaneous abortion it is estimated that there are three unrecognized losses (Wilcox, 1983). With the development of better methods of detecting pregnancy, we will be able to measure early miscarriages more precisely. But early detection does not indicate anything about the cause of the miscarriage.

Stillbirths. Stein and Warburton (1981) estimate that 2-4% of pregnancies that survive to 28 weeks of gestation will end in stillbirth. In a study which included pregnancies over a 40-year period, the overall risk of stillbirth was 17.3 per 1000 women (Wilcox, 1983). In 1981 the U.S. fetal death ratio (the ratio of fetal deaths of 20 weeks of gestation or older to live births) was 9.0 per 1000 live births with a ratio of 8.0 for whites and 12.8 for non-whites (Bureau of the Census, 1984).

Low birth weight/low Apgar scores/multiple births. Beyond potential effects on conception and the chances of successfully carrying a healthy fetus to term, environmental contamination may affect the viability of an infant after birth. Low birth weight and Apgar scores may reflect low-dose chemical exposures that do not result in malformations but that do affect fetal development and health. Several studies suggest that birth weight is an indicator of extra-uterine stresses in the environment including exposures to toxic substances (Jason and Glick, 1981; Williams et al, 1977; Nordstrom et al, 1978). Neonatal mortality may also reflect exposure to environmental contaminants, as may changes in the sex ratio at birth -- the number of males born to the number of females born (Bloom, 1981).

Low birth weight. Approximately 6.8% of live births in 1981 were 2500 grams or less. Only 1.2% were 1500 grams or less (NCHS, 1984). Low birthweight is more common among blacks than whites -- 12.53% of black births weighed 2500 grams or less compared with 5.67% of white births in 1981. It is also more common among less well educated women -- 10.1% of women with 9-11 years of schooling had low birthweight babies compared to 4.7% of women with sixteen years or more of education in 1980 (NCHS, 1985b).

Apgar scores. The 5-minute Apgar has a range from 0 to 10. Only 11.7% of live births in 1980 scored under 9, and only 3.8% scored under 8 (NCHS, 1985b).

Multiple births. In 1980, 68,647 births out of 3,547,362 were multiple births or 1.9% (NCHS, 1985b). The occurrence of multiple births does vary by race with multiple births most common among blacks and least common among Asians (Hellman & Pritchard, 1971).

## REASONS FOR DIFFICULTIES IN OBTAINING BASELINE DATA

Detection and measurement problems are two reasons for our current ignorance. With birth defects, for example, not all defects are recognizable at birth. Even for the major malformations that can be readily identified, not all hospitals or physicians may be equally careful about listing them on the birth certificates. Although most of the U.S. birth registration areas require the reporting of major malformations, some do not (National Center for Health Statistics, 1983). There is a similar problem with fetal deaths.

Measurement and detection problems become more severe when trying to measure spontaneous abortions. Most losses occur in the first 12 weeks of pregnancy, often before a woman even knows she is pregnant. Even if a woman knows she is pregnant, she might not have seen a doctor or been to a hospital so there would be no official record of her loss. Even later fetal deaths are not always recorded. To complicate matters, it is not clear whether early fetal losses and later fetal losses are reflections of the same problem or if they reflect different processes. For EPA's purposes, the question is whether reproductive toxins affect them in similar or different ways.

Without some sense of the natural occurrence of birth defects, pregnancy loss, and other adverse outcomes, it is impossible to determine if their numbers are excessive under particular circumstances. Even if we were able to detect negative outcomes beyond an expected level, the problem would remain how to determine the causes of these outcomes. Not all of them can be attributed to environmental contamination and there may well be multiple causes of the outcomes.

There are other problems that must be faced in trying to establish expected levels of adverse reproductive outcomes. The occurrence of the various outcomes may be related to other variables. Levels of infertility, birth defects due to chromosomal error, and spontaneous abortions, for example, vary by age. Any attempt to establish baseline information must control for the age of the mother at the time of the study. One group of researchers projects a 64 percent increase in the number of American children born with serious chromosomal defects between 1978 and the year 2000, due in large part to the recent shift to older ages at childbearing (Selle et al, 1979). There is also evidence that geographic variations in malformations may in part be due to differences in background

radiation: the latitude/longitude effect indicates a strong solar radiation factor (Jason and Glick, 1981). Additionally, smoking has been linked to spontaneous abortion and to subfecundity (Baird and Wilcox, 1985). Furthermore, to the extent that drugs are in common use and their teratogenicity or effect on human reproduction is not confirmed, they could be confounding elements in any attempt to understand reproductive effects of environmental contamination. As of January 1985 there were over 7,746 drug products approved for manufacture by FDA. (This figure includes multiple versions of a drug whether they be generic equivalents or versions that can be administered orally versus intravenously [Personal communication with FDA information officer]). One would also want to know how many women deliberately aborted because of a suspected birth defect. It is imperative, then, to control for such behavioral and structural differences when trying to establish either baseline information on reproductive outcomes or environmental effects of reproductive outcomes.

As suggested above, Apgar scores and birth weight might profitably be used to explore effects of environmental exposure. These measures have the advantage of being accurately recorded and readily available. Their relationship to environmental exposures, however, is not clear and may be complex. Further, they are only available for live births, so that if exposure to certain hazardous materials increased early miscarriages such effects would be missed.

## CRITERIA FOR AN IDEAL SURVEILLANCE SYSTEM

An ideal adverse reproductive outcomes surveillance system would meet the following criteria.

(1) The system would be based on a well-defined population (Jason and Glick, 1981; Edmonds et al, 1981). The population would be representative and demographically diverse. This criteria would ensure that changes in frequency of outcomes would not be confounded with changes in the structure of the population, such as changes in the ages at which women bear children or in the racial composition of the population over time.

(2) There would we well defined measures of reproductive outcomes.

(3) There would be complete coverage of all outcomes in the study population (Edmonds et al, 1981).

(4) The population would be as large as necessary in order to detect changes in the events. This requirement is necessary because many of the outcomes are relatively rare (Edmonds et al, 1981).

(5) The system would be able to accurately and rapidly report the occurrence of adverse outcomes (Edmonds et al, 1981; Wilcox, 1983).

(6) It would be possible to accurately measure exposure data.

(7) It would be possible to validate both outcome and exposure data (Jason and Glick, 1981).

(8) The data would be of sufficient quality that it could be augmented by other recording systems without unduly complicating data interpretation (Jason and Glick, 1981).

(9) It would be possible to obtain information on individuals such as their age, race, and occupation.

(10) Place of residence, length of time at residence, medical history as well as other potentially important control variables, would be collected.

(11) As far as birth defects are concerned, it would be possible to link maternal and child records, and to link data to original birth certificates to avoid duplication of rare outcomes (Jason and Glick, 1981; Edmonds et al, 1981).

In the real world, however, concessions need to be made to cost, simplicity, and other practical considerations. The question becomes which of these criteria should be kept and which can be ignored or altered without unduly affecting the quality of results. There is also the problem of how well we can actually measure some of the outcomes and events of interest. We have already mentioned the problem of capturing spontaneous abortions which occur very early. There are also problems with measuring environmental exposure. In cluster situations, for example, exposure is suspected because of perceived increases in adverse outcomes. The contamination, however, may no longer be present; it may be of a type that breaks down with time; and, in any case, doses of exposure may vary over time even in the same location (Bloom, 1981).

## EXISTING FEDERAL SYSTEMS THAT ARE USED TO MONITOR REPRODUCTIVE OUTCOMES AND THEIR LIMITATIONS FOR EPA'S PURPOSES

There are two major federal systems that attempt to monitor reproductive outcomes: the Centers for Disease Control's Birth Defects Monitoring Program (BDMP) and the National Center for Health Statistics's Vital Registration System. Although both programs have merit, they also have serious limitations as far as the needs of EPA are concerned. In this section we briefly describe these systems and their limitations.

## Birth Defects Monitoring Program

The Birth Defects Monitoring Program (BDMP) was begun in December 1974 by the Centers for Disease Control (CDC) with the aid of start-up grants from the National Institute for Child Health and Human Development and the National Foundation -- March of Dimes. It relies upon hospital discharge data from over 1,000 U.S. hospitals with obstetric services. These hospitals are self-selected from a set of hospitals using the same medical auditing system. The medical records departments of these hospitals prepare case abstracts for every patient discharged. These case abstracts are compiled, edited, and entered onto computer tape by a non-profit health organization which then provides CDC with summary tapes. Data collected include diagnoses and surgical procedures, sex, race, date of birth, date of discharge, and birth weight.

The BDMP is the largest single source of data on malformed newborns in the United States. In 1981 about 30 percent of all births in the U.S. occurred in hospitals covered by the BDMP -- or about 1 million births. Because the hospitals are self-selected, however, these births were not geographically representative. In 1981, the BDMP captured about 17.3% of all live births occurring in the South, 28.5% of those occurring in the Northeast, 41.0% of those occurring in the North Central states, and 26.4% of those occurring in the West (Edmonds et al, 1981). Since 1981 some of the hospitals have been lost to the BDMP because they no longer use the medical auditing system. Consequently the universe of coverage has not been the same. Staff at CDC, however, have made the assumption that it is cost considerations that stimulated particular hospitals to leave the system rather than anything that might be confounded with birth defects. At the present time, CDC staff are deciding whether to include other willing hospitals in the BDMP (Personal communication with CDC staff).

In spite of its many advantages, the BDMP has some serious limitations from EPA's perspective. The limitations include the following:

- It is not a representative sample of births in the U.S.
- The data are not well-defined geographically - that is, the proportion of births covered varies substantially across regions of the country and participating hospitals in a given state or area are not necessarily representative of all births in that state or area.
- There is no quality-control mechanism to insure that different hospitals diagnose and classify defects in a uniform manner.
- There is no coverage of early miscarriages.

- There is little or no information on important variables such as length of residence in an area, occupation, smoking history, etc. Thus important exposure information is missing.

- There is no opportunity for prospective tracking of pregnancies and their outcomes.

U.S. Vital Registration System

The National Center for Health Statistics (NCHS), Division of Vital Statistics, collects and publishes natality and mortality data that are provided by the states. Certificates of live births and of deaths, completed by the attending physician or other health personnel, are sent by local registrars to the State registrar. States report the data to the federal Division of Vital Statistics, NCHS, on State-coded data tapes or in the form of microfilm copies of birth or death certificates. Data collection is continuous. Monthly and annual reports of provisional data and annual and special reports based on final data are issued. All states have been included in the birth and death registration areas since 1933.

Birth certificates contain a limited number of items including such things as mother's race, education, and usual residence, the infant's birth weight and Apgar scores, and other information. Death certificates contain similar information except that they do not report weight or Apgar scores. Fetal death certificates, however, do ask for the weight of the fetus. The information on residence given on the various certificates gives no sense of the length of time at that residence. Not all states collect all the information recommended by NCHS. Only 4 of the 52 registration areas in the U.S. (the 50 States, the District of Columbia, and New York City) do not report malformations or other anomalies of the child on the live birth certificate and only 6 do not on the fetal death certificate. These numbers, however, indicate nothing about the accuracy of the reporting. Generally birth defects are severely underreported on vital statistics records. One reason is that not all birth defects are recognizable at birth. But there is also a tendency to underreport recognizable but subtle defects. Currently, there is an open ended question on malformations on the birth and fetal death certificate. NCHS is considering introducing a new Standard Birth Certificate which would contain a list of specific malformations that could be checked off by the attending health person.

NCHS is a data disseminating agency. Thus the data are available in several formats and at reasonable cost. Unfortunately, however, there is frequently a long time delay between the collection and the availability of the data. Because the data are first collected by local areas and then transmitted to the State for conversion to machine readable form, there is also ample room for transcription errors and for

biases introduced by differences in local reporting practices, neonatal medical care, nomenclature conventions, etcetera. Measures such as birth weight and Apgar scores, then, are probably more reliably reported than are types of malformations.

Although the birth statistics data in the Vital Registration System capture nearly 100 percent of births occurring in the United States, the data do have the following limitations for EPA's purposes:

- Birth defects are significantly underreported on birth certificates, and the degree of underreporting varies across different types of defects.
- There is no quality-control mechanism to insure that physicians in different states and counties diagnose and classify defects in a uniform manner.
- Reporting of fetal deaths are affected by differences in state and county regulations.
- There is no coverage of early miscarriages.
- Although there are data on usual residence, there is no indication of duration at that location.
- There is no opportunity for prospective tracking of pregnancies and their outcomes.
- There is usually a lengthy time delay in NCHS reporting of birth defects data.

## APPROACHES SUGGESTED BY THE WORKING GROUP

Four broad approaches to EPA's problem emerged from the Working Group meetings:

(1) Exploit the existence of ongoing federal surveys to develop baseline data on the incidence and prevalence of reproductive problems among women living in different types of communities and among those from different age and ethnic groups.

(2) Encourage the development and improvement of birth defect monitoring systems in the states.

(3) Work with other agencies to sponsor a jointly-funded network of birth hospitals that would provide high-quality continuing data on birth defects in a representative sample of U.S. births.

(4) Develop standard questionnaires and local survey methods that could be used to determine what the incidence and prevalence of negative reproductive outcomes in a given area really are.

A fifth recommendation was that EPA provide support for basic biomedical research that would ultimately contribute to assessing the impact of environmental contamination on human reproductive outcomes.

These recommendations are described in more detail in the Report of Findings and Recommendations issued by the Working Group. A summary of the deliberations that led to the recommendations is contained in the following sections of this report.

SUMMARY OF THE FIRST WORKING GROUP MEETING
HELD JUNE 24, 1985

## Opening Remarks

Nicholas Zill opened the meeting by asking each of the participants to introduce themselves. He then briefly described the NSF project as a whole and the goals of the EPA project in particular. Zill also emphasized that a corporate goal of Child Trends was to improve national statistics on children in general and that the EPA project should also be seen in light of that goal.

## Clarification of Objectives

Several members of the Working Group requested clarification of the Working Group's objectives. Many of the members have participated in other committees that have written reports describing the kinds of data that can be collected. They wanted to know how this project is different and where this project will go. In particular, will a data collection effort be mounted or will the same old ground be covered again to no avail?

Zill attempted to address these concerns and, later in the meeting, Gruber, from EPA, also attempted to respond to them. Zill's and Gruber's remarks are summarized below:

EPA is very concerned about the enormous amount of time and money that is spent responding to outcries of alleged clusters of adverse reproductive outcomes that may be due to environmental contamination. It wants to preclude responding to false alarms. Thus the mission of this project has a strong policy focus. EPA would like the Working Group to recommend ways to address this policy problem more successfully. EPA would like to develop baseline data so that it could assess whether a particular community has an elevated rate of adverse outcomes. Many of the situations that EPA encounters involve fear feeding on fear. If the distribution of such events were known then EPA and communities would have a better idea of when concern was warranted. Gruber has an image of a booklet that could be widely used, with columns of numbers noting the population incidence of various outcomes. The goal, then, is the development of a feasible way to provide national baseline data and perhaps trends over time, a method that provides data on the expected rate of an event, and which can be looked at to see whether rates in specific areas or studies can be said to be high.

Such a baseline would not necessarily be meant to provide a definitive answer on whether or not negative reproductive outcomes are unusually high or whether they may be due to environmental contamination. It is hoped, however, that such a baseline would be a way to separate potentially justified concerns from obviously unjustified ones. Epidemiological studies would still be needed in attempts to sort out particular causes and effects in particular communities.

Because EPA wants a baseline that would apply to the whole United States, a national perspective is sought. Among the questions that EPA and Child Trends would like the Working Group to consider are:

1. What types of information can one get with reasonable validity from survey reports?

2. What is the incremental advantage of increasingly expensive methods of data collection?

3. What are the merits and drawbacks of the various data collection approaches?

4. What are the outcomes that should be measured and how good are they?

5. Is it worthwhile to pursue sophisticated measures such as chemical pregnancy tests?

EPA would like to see a data collection effort funded. Whether a data collection effort will be funded, though, is another matter. EPA recognizes that it will need to have cooperation from other agencies as it has neither the resources nor the expertise to collect such data alone. However, if the Working Group can recommend a range of possible approaches along with estimates of their likely costs then Gruber and others at EPA would have some solid information to use in advocating the collection of this type of data. Gruber suggested that EPA's need for a solution to their policy problem strengthens existing scientific reasons for developing such data and thus improves the likelihood of obtaining the funds. Furthermore, Child Trends, as an organization, has some experience in lobbying to help bring proposed data collection programs into being -- for example, the Child Health Supplement of the National Health Interview Survey.

Outcomes

After the above discussion, the Working Group turned to its first task of the day: reviewing the list of reproductive outcomes contained in the packet of information that had been mailed to each participant prior to the meeting. There were twelve reproductive outcomes included on this list:

1. Infertility (non-voluntary)
2. Subfecundity
3. Very early unrecognized fetal loss
4. Voluntary abortions because of antenatal diagnostic test results
5. Miscarriages/spontaneous abortions (fetal deaths before the 20th week).
6. Stillbirths (later fetal deaths)
7. Birth defects in live or still births
8. Low birthweight and/or prematurity in live births
9. Low apgar scores in live births
10. Unusual sex ratios
11. Incidence of multiple births
12. Infant mortality

Two criteria for the review of these outcomes were: (1) the usefulness of the outcome as a potential indicator of exposure to hazardous materials; and (2) how well the outcome could acutally be measured. As Zill stated, the group was in a sense looking for the best "miner's canary", a measure characterized by sensitivity, specificity, and feasibility. The group, however, was not expected to narrow the list to a single indicator, but to select those that seemed best for the project at hand.

Mattison made the point that before the group looked at adverse reproductive effects, it needed to consider what a normal or successful outcome was and then discuss adverse outcomes against this definition. He suggested that a successful outcome is one where the timing is as desired, and the child is structurally and functionally normal.

Warburton commented that it is impossible for every pregnancy to end in a healthy, live birth. There will always be some level of chromosomal problems that are not due to environmental or any other apparent influence.

The discussion of these outcomes was somewhat diffuse, ranging over several outcomes at the same time, and intermingling questions about the project with discussions about sources of data and the usefulness of the outcome. Furthermore, not all the outcomes were discussed. What follows is a summary of the main points that were made about these outcomes during the course of the meeting.

## Infertility/Subfecundity/Spontaneous Abortions/ Miscarriages/ Stillbirths

### Usefulness

There seemed to be a general consensus that fetal losses were an important outcome. There was no direct discussion of the usefulness of measuring infertility and subfecundity, although Warburton noted that non-fertility and early pregnancy loss are often indistinguishable. There was some disagreement on whether attempts should be made to measure spontaneous abortions or whether the baseline should simply be recognized pregnancy loss. Arguments against including spontaneous abortions as an appropriate outcome to follow were based on the technical difficulties in accurately measuring them. Arguments for their inclusion were based on a study by Jason that had shown patterns of repeated early abortions and other studies that had shown large proportions of early pregnancy wastage.

### Sources/Measures

Several possible sources or methods of measuring these different outcomes were discussed including hospital records, personal interviews with the woman, prenatal monitoring, and chemical pregnancy tests.

Hospital Records: Hospital records were not considered a good source of information on miscarriages in spite of their use of standardized forms. About 40% of women with first trimester miscarriages do not go to hospitals.

Women's Self Report: Several Working Group members recommended using women's own reports of fetal deaths as the best way to gather such information. In a developed country like the U.S., women seem consistent in their reporting and reproducibility across studies is high.

Women can also be asked questions on time-to-pregnancy. Warburton did not think questions such as "How long did it take you to conceive?" were particularly useful without a lot of other control variables and that they might be difficult to use on a large scale. Mattison mentioned a study by Wilcox which demonstrated the adverse effects of smoking on time-to-pregnancy as reported by women. Mattison liked that type of question as a first dirty screen of both male and female effects.

Prenatal Monitoring: Warburton and Holtzman noted that prenatal monitoring is now more common. It is possible to get a view of pregnancies at 8-10 weeks with ultrasound. These new technologies pick up many non-viable pregnancies and create the possibility of looking for abnormalities very early.

Zill asked if it would be possible to sample facilities performing the procedures. Kessel commented that the distribution of these technologies was uneven and that such a sample might well be biased. Warburton remarked that in New

York ultrasound is done among all women in all classes. She thought that some kind of restricted sample collected nationwide might work. At present, the data are collected, but are not being analysed on a broad scale.

Chemical Pregnancy Tests: The questions here are how much chemical testing increases the number of pregnancies found over self-reporting, and is the gain from doing such tests sufficient to justify the additional effort.

The Group had mixed reactions to the use of chemical tests. On the one hand were those who pointed out the limitations of the tests: It is not uncommon to get different results from different labs. There are problems even in the definition of a pregnancy. It is impractical to perform such tests on a large scale. It is not yet certain that we know how to use the methodology.

On the other hand were those who saw merit in the tests: Jason had used a chemical test on a small study among female employees of the New York State Department of Health in Albany. He got good cooperation from the women and had detected patterns of repeat abortion. Holtzman thought such tests would be worthwhile if we had the methodolology. Jason argued that the methodology is there, but it is a question of the adequacy of support.

The final conclusion seemed to be that, although the methodology is rapidly developing and in the future might be fruitfully employed, at the present time, practical, logistical, and technical considerations argue against their use.

Other Sources of Data: Holtzman suggested that important information might be gained from examining aborted fetuses.

## Birth Defects

### Usefulness

Birth defects in stillbirths and live births were recognized as important outcomes. Several members, however, pointed out that gathering accurate, comparable data is difficult. Not all malformations are recognizable at birth. Cardiovascular defects, for example, are hard to ascertain. Children often have more than one problem, but only one may get recorded and not necessarily the most obvious one.

### Sources/Measures

Several possible sources or methods of gathering data on birth defects were discussed including birth certificates, hospital records, DRGs, mothers' reports, and examination of older children.

Birth Certificates: In their present form, birth certificates are not very useful as sources of information on birth defects. Ventura, however, noted that a panel is considering revisions to the standard birth certificate. The open-ended question on birth defects is being replaced by a check list of anomalies which is very similar to the CDC list. he panel is also considering adding the occupation and industry of both parents. The Working Group members were in favor of these changes. With these changes, birth certificates would become a better, but still imperfect, source of information.

Hospital Records: Hospital records were generally agreed to be the best source of information on birth defects. They do have some limitations, however. First, nowadays women with problems diagnosed by ultrasound may be moved to specialized medical centers. This practice could distort the rate of birth defects in specific hospitals. Second, the quality of hospital statistics is not high. There tends to be too much reliance on the physician for data. The quality of hospital statistics, however, could be improved if hospitals and concerned individuals focused on the problem.

Holmes and Holtzman suggested ways to improve the quality of data derived from hospital records; such as having a team of individuals working in a hospital read through and code doctors' and hospital records. For cases that are not clear, more specific information could be obtained from an appropriate hospital staff person. This team could also review and code data on stillbirths, which we know from the Collaborative Perinatal Project are an important outcome. Similar teams could review birth certificates.

Zill asked if it would be appropriate to sample hospitals as a way to gather data on birth defects. The Group seemed to agree that such an approach might be worthwhile.

Diagnosis Related Groups (DRG): DRG codes were not thought to be appropriate for gathering birth defect data even though their use is becoming more prevalent. One problem is that there is an incentive to report certain types of birth defects because reimbursement for them is higher. For example, if a child has Down's Syndrome and a heart defect, the heart problem will get reported because reimbursement is higher.

Mother's Reports: On the question of whether it is worthwhile to gather birth defect information directly from mothers, Adams noted that an examination of CDC data showed that a high proportion of mothers either did not report the defect or were so non-specific that the data were worthless, especially if the baby died. Most of the Working Group seemed to agree.

Conceivably if women were given a checklist of specific outcomes, reporting would be better. Warburton noted one study in which people were able to fill out a checklist history with what appeared to be considerable accuracy. (See attached letter from Graham Kalton. He discusses this issue further in item 3.)

Examination of Older Children: With regard to the possibility of examining older children as a way of picking up birth defects, several problems were noted. First, changing the time frame inevitably changes the numbers of defects that are detected. Second, you would miss the children who did not survive. Third, the older the child the more chance there is that other factors have intervened. Fourth, shifting treatment methods over time affect the rate of survival. And, fifth, diagnoses may change over time.

Holtzman pointed out, though, that the association between smoking and strabismus would not have been discovered if only newborns had been examined.

Other Sources of Data: Teratogen hot lines that women can call if they have some concern are another potential source of data. The sample, however, is almost certain to be biased.

## Birth Weight, Multiple Births

### Usefulness

The Working Group only briefly discussed birth weight and multiple births, primarily because it was felt that these variables were already well measured and extensively analyzed. Multiple births were not considered a very useful indicator because it was not clear exactly what they revealed. Birth weight was seen as a potentially useful indicator. However, it was noted that the majority of children born with defects are of normal birth weight. Moreover, even if one detects an excess of low birth weight in an area, there is the problem of linking this to environmental contamination as opposed to other potential causes.

There was also discussion of the need for associated data on length, head size, and gestational age.

### Sources/Measures

Multiple births and birthweights are both reported accurately in birth certificates. The certificates lack information on some important confounding variables, however.

## Other Outcomes

Apgar scores, unusual sex ratios, and infant mortality were not discussed.

## Using Birth Certificates as Sampling Frame

The discussion of some of the limitations of birth weight data, as well as the poor reporting of birth defects on vital records noted earlier, led to a discussion of the possibility of using birth certificates as a sampling frame. If that were possible, then additional information could be acquired in a follow-back survey. The National Natality Survey (NNS) does use birth certificates as a sampling frame, but it only has about 10,000 cases and unmarried mothers are not interviewed personally because of state objections. Such a survey might be useful if a larger sample could be supported and if permission could be gotten from unmarried mothers for a subsequent interview. It would be possible then to over-sample on birth abnormalities and to collect some of the data that might be confounded with birth weight.

## Reports from Selected Agency Representatives

In the afternoon representatives of several different agencies were asked to describe ongoing data collection efforts and relevant projects sponsored by their respective agencies.

### National Center for Health Statistics (NCHS)

Stephanie Ventura from NCHS described the changes planned for the next revision of the Standard Birth Certificate. These changes include using a checklist of birth anomolies instead of the open ended question on the current certificate. The panel convened to make recommendations to NCHS is also considering adding occupation and industry codes of both parents, obtaining the date of the last menstrual period, and getting the physician's estimate of gestational age.

Ventura also stressed that NCHS is working at getting the annual data out sooner. The results of data collected in 1984 should be out by the end of 1985. There will always be a lag of 8-10 months or more on the release of data. Data on birth defects are further behind because the International Classification of Diseases has been revised, necessitating additional work at NCHS to recode the data using the new classification.

## Centers for Disease Control (CDC)

Melissa Adams from CDC described both the Metropolitan Atlanta Congenital Defect Program and the National Birth Defects Monitoring Program. She also described three other projects in various stages of development at CDC.

The first is a program, in the final stages of clearance through the Office of Management and Budget, that will be a spontaneous abortion surveillance system among CDC employees. It will monitor contraceptive use, sexual activity, pregnancy, as well as pregnancy loss.

The second is a pilot survey of developmental disabilities being done in an Atlanta-area county. The survey covers blindness, deafness, cerebral palsy, mental retardation, and epilepsy. CDC has been working with the school system to develop a model program.

Finally, CDC is looking at low birth weight counties in Georgia to explore the epidemiology of low birth weight.

## Department of Health and Human Services (DHHS)

Clara Schiffer from DHHS described an initiative in the Office of the Assistant Secretary for Planning and Evaluation. A Request for Proposals is being issued seeking a contractor to evaluate existing procedures and systems for the reporting of birth defects in the states. Several states have legislation pending. Her office is hoping to increase uniformity in data collection and coding.

## Missouri Center for Health Statistics

Janice Bakewell from the Missouri Center for Health Statistics described the development of the Missouri Birth Defects Registry. In Missouri, the approach has been to gather all available data and merge the data with birth certificate information. Smoking, zip code, obstetric record, pre-pregnancy weight, WIC data, and a lot of other information is coded. To date, however, emphasis has been on collecting the data. Not much analysis has been done. There are about 77,000 births per year in the system.

## Maryland Birth Defects Reporting and Information System

Neil Holtzman from the Johns Hopkins Hospital described the Maryland program. He was instrumental in its design.

The Maryland program began when a researcher wanted to see whether the incidence of cleft lip in a newly constructed school was unusually high and discovered baseline data did not exist.

This researcher had a friend in the state legislature. Through his efforts, legislation was passsed to fund a program through 1986. In Maryland, 11 sentinel birth defects--selected from the International Classification of Diseases because they are severe and fairly common--have been recorded since 1983. Residences, contraception history, smoking, and occupation are filled out by the mother with clerical staff. The obstetrician fills in the medical part, and the pediatrician fills out the birth defects section. Coding is compatible with the toxic substances registry, the birth defects registry, and the cancer registry. Reporting is completed within 6 months. Staff are checking the extent to which the data are completely and accurately reported. Physican cooperation is one obstacle. Mothers of all children with a birth defect are asked to consent to the release of their names and 90 percent did so the first year. Data for a control sample may also be collected.

## Possible Approaches

Zill summarized four possible approaches that seemed to have received some support during the discussion of outcomes in the morning. These were:

1. Collect data on infertility, miscarriages, and reproductive histories directly from women in population surveys.

2. Sample hospitals with an aim towards collecting high quality data on birth defects.

3. Sample clinics that do ultrasound, amniocentesis, and other antenatal testing.

4. Use birth certificates as a sampling frame and do a follow-back interview similar to the NNS. Oversampling would be necessary, though, to obtain a sufficient number of cases of birth defects.

Two additional approaches were suggested:

5. Collect prospective data in HMOs. For example, women upon their first prenatal visit could be asked to fill out a questionnaire. These women would then be followed through their pregnancies. One problem with this approach, however, is that when women begin prenatal care varies by such characteristics, as whether they have other children or whether they smoke.

6. Set up a network of state registries in which uniform data on a limited set of birth defects, perhaps other outcomes, and important control variables are collected.

No consensus was reached on which approaches might be the most fruitful to pursue.

Several Working Group members expressed concern about the ability of a national data base to pick up changes in small areas. It was argued, however, that information collected at the national level could be duplicated in communities. Some felt that if community anxiety levels were high in those communities surveyed, then the data might be biased. Selevan, however, noted that at least one study had found reports agreed with medical records even in a location of high anxiety.

The Working Group members also discussed how high the quality of the data needed to be. Holmes and others argued that high quality data were needed in order to obtain good estimates. Kalton, however, disagreed. He said that although much of the discussion presumed the need for very good measures, often you could get by with rather poor measures. For example, you may be able to pick up the effect of an environmental hazard on miscarriages, even if some women don't remember or know they had a miscarriage, unless there is some confounding of exposure and underreporting. (See the attached letter from Graham Kalton. He discusses this issue further in item 1.)

Those who argued for high quality data tended to argue against using a national population survey as a way to obtain data. Holmes, for example, thought that the best approach would be to rely on data obtained from a network of hospitals in several cities. Special teams could work in each hospital to abstract data from hospital and doctors records, supplementing these records when necessary.

Others, who were also against a federal approach, thought it would be best to work through the states. After all, states have a better knowledge of conditions and quirks within their boundaries. Questions were raised, however, over how cooperative the states would be. States are lethargic and lack funds. Change is hard, but if funds are provided they might be more cooperative. Kessel thought a well-defined effort might find support in the Public Health Service (PHS), working with the states in well-defined geopolitical areas. The PHS's primary interests are in infant mortality, low birth weight, and prevention.

With regard to how one would go about developing a network of state data systems, Kessel thought it best to keep it simple. Work with a small group of people. Begin with a pilot study and build up from that.

It was noted that no mention had as yet been made on ways to link exposure data to any of the data that might be generated by the suggested approaches.

## Control Variables

The next question discussed was what some of the crucial control variables were that should be included in any proposed

study. Smoking, parental age, alcohol use, and race were agreed to be important. Adams noted that the proper controls really depended on the outcome under consideration. And Holmes said that the issue is how long the questionnaire would be. Any number of controls could be included. Nutritionists, for example, would want to have detailed questions on nutrition. (See attached letter from Graham Kalton. He discussed this issue further in item 2.)

Reasons for Lack of Baseline Data

Gruber, wanted to know why it wasn't already possible to tell if adverse outcomes were high in particular communities. He said it seemed a national scandal that we cannot tell whether the well-being of children is threatened, that we are unable to mount prevention efforts, that people are forced out of their homes.

The Working Group members speculated on some reasons for this lack of knowledge. It is more than just a lack of money, the problem is more subtle than the thalidomide issue suggested Definitions are a big problem. Adams noted that at CDC a lot of time is spent just fighting fires. Zill commented that not enough emphasis has been placed on analysing the data we have. He suggested that perhaps that is an area that needs money.

Cost of Data Collection Efforts

Gruber asked what it would cost to gather the types of data he would like. Zill noted that the National Natality Survey costs about $4 million. Holmes estimated that his plan involving a network of hospitals in several cities would cost about $100,00 per year per city. Holtzman said the Maryland Birth Defects System costs about $50,000. Mattison thought that even a relatively modest investment of between $20,000 and $30,000 per state could bring about substantial improvements in our knowledge.

Plans for Second Meeting

The next meeting of the Working Group is scheduled for Friday, September 27th. It was agreed that Child Trend's staff would gather estimates of the incidence of the different reproductive outcomes from assorted sources. These estimates and an agenda will be mailed to the Working Group members prior to the second meeting. It was also agreed that the Working Group should begin to consider what the different approaches might cost. Child Trend's staff would begin to gather information on this subject, but Working Group members who advocate specific approaches should also try to estimate the costs of those approaches.

SUMMARY OF THE SECOND WORKING GROUP MEETING
HELD SEPTEMBER 27, 1985

Opening Remarks

Nicholas Zill welcomed the participants who had braved Hurricane Gloria in order to come to the meeting. He stated that the meeting would focus on three possible approaches for EPA to consider: (1) encourage the formation of a network of state registries; (2) use hospital based teams to abstract data from medical charts; (3) add appropriate questions to existing surveys or conduct a new survey that could be duplicated in small areas.

Additional Discussion of the Goals of the Project

Warburton said that she was still unclear about EPA's objectives. Does EPA want to develop statistics on the normal frequency of adverse reproductive outcomes? These already exist as Nord has shown in the estimates of adverse birth outcomes that she gathered from different sources. Why do you need more data than you have? New data will yield similar estimates because a large increase in an outcome is needed before any change can be detected.

Zill commented that current data are not very useful to EPA in evaluating situations in particular locations. He also suggested that outcomes with more frequent occurrences, such as subfecundity or miscarriage, might be more useful for EPA's purposes because it would be easier to monitor and to detect changes in the frequency of occurrence. In any case, estimates on the different reproductive outcomes need to be more refined with more information on mediating variables.

Warburton pointed out that we have lots of data on miscarriage, for example, by parity, smoking, and other important control variables.

Sherry Selevan stated that as an epidemiologist, she would not be confident with results unless there were also a comparison group.

Zill added that there is now no standard methodology for going into a community and finding out what their rates are.

Mike Gruber restated the problem faced by EPA. What if there are three birth defects in a school when there had only been one birth defect in that school in the previous ten years? What can EPA conclude? It would be helpful to be able to say there have been forty similar occurrences elsewhere, or to be able to say that this really is unusual. Can you never resolve these types of questions, or is there a way?

## Discussion of Suggested Approaches

The Working Group next turned to a discussion of the suggested approaches. For each approach, the group was asked to consider: (1) its particular strengths and weaknesses; (2) design aspects, including sampling issues; (3) type of outcomes measures; (4) control variables that could be measures; (5) data quality; (6) cost estimates; and (7) feasibility and time frame. Not all these topics were covered during the ensuing discussion.

## I. Network of State Registries

Large scale registration systems were generally agreed to be an effective way to monitor birth defects as the experience in some other countries has shown. Selevan stated that in Finland, for example, every health professional who comes into contact with someone with a birth defect must report it. Of course, Finland is a small country and, as Donald Mattison pointed out, they have a personal identification number for everyone from birth, which enables them to guard against double counting any particular individual.

### Potential Difficulties

Setting up a registration system in the U.S. would be more difficult because the U.S. has no personal identifier number from birth and is unlikely to develop one. In spite of this difficulty, however, some states, such as Missouri, do have working systems that are not prohibitively expensive. In Missouri data from different sources are matched by name whenever possible. If no name is available, matches are attempted by the use of zipcodes, sex, race, or whatever piece of information is available.

Another difficulty with a network of state registries is that individual states would have responsibility for collecting data so that there could be problems with uniformity of collection and recording of data, as well as potential differences in completeness of coverage and data quality. Systems that rely heavily on hospital discharge data would tend to underreport defects unless they had other sources of data that could be linked to the hospital data to pick up undetected

cases. Differences in staff motivation could also affect the completeness of coverage. Some type of quality control system would probably be needed. As far as the Working Group members were aware, no state that did collect data had yet examined the quality of the data that were gathered. Thus, it is not clear yet how good the data are. Warburton suggested that spending money validating existing systems might be a worthwhile venture.

A third problem is that births that occur out of state may not be caught in that state's registration system. This problem could be particularly severe in large metropolitan areas such as near New York City or even here in the District of Columbia. Holtzman said that about 15% of Maryland births occur outside of Maryland. It is important, then, to get the place of usual residence as well as the place of birth.

## Interest in Registries Exists

In spite of these obstacles, however, there is interest in encouraging such a system. The Centers for Disease Control (CDC) has released a new RFP that encourages the collection of better data on birth defects by states. The RFP also wants to encourage the development of new statistical methods for monitoring pregnancy outcomes and to explore the use of Hcg in detecting early pregnancy loss. Clara Schiffer's Office recently awarded a contract with the specific goals of evaluating existing procedures and systems for the reporting of birth defects and other poor reproductive outcomes in the states and describing any relevant legislation that a state may have or may be considering.

Melissa Adams from CDC strongly advocates a state approach. She had worked up a sketch of one possible design for a network of state registries. Unfortunately, she was unable to attend the meeting because of Hurricane Gloria. Nord briefly summarized the main points that Adams had planned to make based on a telephone conversation with her. What follows is the outline Adams sent of what she had intended to say at the meeting.

## Adams's Notes for Working Group Meeting: Network of State-Based Registries of Adverse Reproductive Outcomes (9.25.85)

### Design

1. Passive surveillance by state registry of stillbirths and infants with structural malformations diagnosed during the first year of life. Reporting would be done by hospitals and reports would include personal identifiers, date of birth, hospital of birth, diagnoses, and procedures.

2. State-level data management (editing and coding) and periodic audit of passive hospital reporting.

3. State-level linkage of stillbirths/birth defects with birth certificates. State-level identification of infants with low birth weight, using birth certificates. Based on maternal address listed on birth certificate, linkage of all birth certificates with the state environmental agency's list of geographic locations of toxic hazards.

4. State-level analysis of data, including analysis of associations between preganancy outcomes and toxic hazards; investigations if warranted.

5. All states to collect a uniform minimum set of data and to follow similar protocol for data collection and coding.

6. Submission of state-based data to a national clearing house responsible for publishing amalgamated data and for sharing findings among states.

Advantages

1. Ability to collect data about method of diagnosis; ability to link outcomes and environmental exposures; ability to conduct detailed studies.

2. Ability to investigate suspected increases in adverse reproductive outcomes.

3. Data compiled by a state registry would have multiple uses (e.g., for planning and evaluation of MCH programs).

4. Development of a uniform population-based data base.

5. Relatively modest cost.

Weaknesses

1. Not all states have the technical capacity to conduct surveillance; some states would require technical assistance.

2. Hospital-based surveillance would miss outcomes of low severity; geographic differences in in-patient and out-patient treatment patterns may yield spurious differences in rates.

Sampling (none; population based)

---

1 Passive surveillance refers to case identification by an agency other than the registry that is collecting the case reports. Active surveillance refers to case identification by the registry itself.

Outcomes Measured

1. Stillbirths
2. Major birth defects
3. Low birth weight

Control Variables that Could be Measured

1. Data from birth certificate (i.e., parental occupations, ages, races).

2. Maternal residence (especially as related to areas of toxic exposure).

Data Quality

1. For stillbirths and birth defects -- better than vital records; worse than in-depth exam by pediatric teratologist.

2. Quality probably comparable across states.

Cost Estimates

1. Depending on the size of the population included in surveillance, $100,000 to $200,000 per year per state.

2. $300,000 per year for national clearing house.

Feasibility and Time Frame

1. State-based birth defects surveillance currently implemented or under development in 13 states.

2. Time needed to have national data:
   -for stillbirths, 6 months after the end of the year of interest;
   -for birth defects, 2 years after the end of the year of interest (extra time needed to collect defects during the first year of life).

Reactions to Nord's Summary of Adams's Points; Further Discussion of Network Approach

Mattison thought that the cost estimate for states seemed low as did others. Selevan commented that the estimate may have been based on the vital records part.

Several people made the point that a clearing house on birth defects already exists. It is located in San Francisco and run by Sylvia Haye of the March of Dimes. Of course, the clearing house envisioned by Adams would involve overseeing more sources of data than the current clearing house.

## Estimated Costs of Missouri and Maryland Systems

An attempt was made to estimate what it might cost states. Janice Bakewell from Missouri said that their system cost Missouri about $5,000 a year, but this figure excludes funds that Missouri gets from other sources. Actual costs could be $50,000 or more. She made the point, though, that other states are also eligible for supplemental funds. Warburton asked about the quality of the data and whether Missouri had to provide some sort of incentive to get the hospitals to cooperate. Bakewell said that as long as only hospital discharge data were used, no incentive was needed. If, however, you wanted hospitals to fill out additional forms then an incentive might be needed.

Holtzman said that the Maryland system is budgeted at $50,000 to gather data on 12 sentinel defects. Other anomalies that are noticed are also supposed to be reported. Also, if it is known which chromosomes are affected, these are supposed to be listed. Because the system is mandated by law, Maryland does not provide incentives to the hospitals to cooperate. No analyses have yet been performed to determine the quality of the data. A drawback of the Maryland system is that it only requires the monitoring of 12 sentinel defects, but as Mattison noted this is also a strength because it provides a very focused system which increases the odds of getting the data you are looking for.

## Data Quality

Zill asked if the data in the state systems is of sufficient quality to justify going with a state system? At this point, however, it isn't clear how good the data are. Graham Kalton pointed out that if you get into bigger systems, you might get better quality data if you sample hospitals rather than trying to gather data from all hospitals within a state. Holtzman thought that this might be a viable alternative to complete coverage (See the next section on Hospital-based teams for an expanded discussion of this idea).

## Should EPA Offer States Incentives?

The question arose: What if EPA offered incentives of say $50,000 to work up systems such as the Missouri or Maryland ones? Would such incentives produce change?

Such an approach could not put it into a state's laws and having such laws has some advantages. Gruber stated, however, that EPA could give money to states that have a law. There is a lobby for this. Holtzman said that eleven states have passed laws recently. There is some sensitivity to the issue. Mattison said that states would like to have the data so that that amount of money might be sufficient incentive. Kristin

Moore asked how long such an incentive would be needed? Not permanently she assumed. Mattison thought that the way to go was to do a five-year start-up grant.

The final consensus seemed to be that this approach was worthwhile exploring in more detail in spite of questions about data quality and possible difficulties linking different records. Warburton, however, did note that data collected by this approach would never match data collected in a community on a door-to-door basis. The methodologies are entirely different and would be unlikely to yield equivalent results. The outcome that is best measured by this approach is birth defects although stillbirths and low birth weight could also be monitored.

II. Use of Hospital-Based Teams

Lewis Holmes during the first meeting of the Working Group had suggested the use of hospital-based teams as a way to gather better quality data on poor reproductive outcomes -- particularly on major malformations in abortuses, stillbirths, and livebirths. Prior to the second meeting of the Working Group he drew up a cost estimate of such a team working in two contiguous Boston hospitals. This estimate was distributed to the other Working Group members. Unfortunately, Holmes was kept from the second meeting by Hurricane Gloria and thus he was unable to defend or explain his ideas in more detail to the other Working Group members.

It was quickly pointed out that this approach is very closely related to the one described above. Both are attempts to improve the quality of data on birth defects, in particular. As Kalton noted earlier, in states with a large number of hospitals it might be more practical to take a sample of them rather than gather information from all hospitals. Unlike the state network approach, however, this approach makes no effort to incorporate non-hospital data. Such linking of data, however, is not precluded by this approach.

Moore pointed out that there is a need to approximate the level of motivation exhibited in local areas that fear some sort of contamination. The better quality data that a system can gather, the more nearly comparable that data will be to data that are gathered in small areas. However, it must also be recognized that unless the data are gathered using the same methodologies, they will never be completely comparable.

Advantages

A prime advantage of this approach is that it is likely to provide much higher quality data on malformations than simply using hospital discharge data because there will be a trained specialist who can verify the discharge data and contact relevant hospital personnel to clarify any questions that might arise.

Besides yielding higher quality data than the state network approach, the use of hospital-based teams means that results of antenatal diagnostic screening tests could also be included. Heuser pointed out, however, that if looking for the number of children in a community with a condition, live births is the appropriate comparison group -- not live births plus births that were terminated or ended in miscarriage or stillbirth. Warburton and Mattison noted, however, that screening tests are going to have a big impact on the patterning of birth defects. It is going to be important to have data on results from screening.

A third advantage is the ability to add additional variables. This ability enables more control over the type of subgroup comparisions that could be made.

Another advantage is the fact that there would be a standard methodology used in all the hospitals. The data then are directly comparable.

Sampling Issues

Sampling questions are very important in this approach. In order to have data that are generalizable, a probability sample must be used. However, as Holtzman pointed out, major hospitals tend to be centers for referrals for high risk pregnancies. Thus some type of stratified sample would be needed. If it is decided to include data from screening tests, the sampling problem will be further complicated. It is also important to get a sample of births without defects as a control group. Holtzman asked whether the number of cases of malformations obtained through a sample of hospitals would be large enough.

The question of whether hospitals could volunteer to be in the system if they happened not to be within the sample drawn was raised. It was recognized that hospitals that voluntarily participated might differ systematically from hospitals that were randomly selected. Kalton said that it would be possible to include volunteer hospitals and then take a probability sample from the remainder. Such an approach would be like taking a stratified sample. Weights would have to be used and thus you would get less precision for a given sample size. Large hospitals, however, may all fall into the sample anyway. In spite of all the potential sampling problems, Kalton felt that they could be resolved and a workable system designed.

Comments on Estimate Provided by Lewis Holmes

Kalton stated that Holmes' estimate seems to be geared toward having a particular hospital move ahead and do its own thing rather than be a part of a sample of hospitals each trying to produce similar data.

Gruber said that it would be necessary to centralize such an approach either regionally or nationally. The national or regional center then would do some of the tasks budgeted for each hospital in the Holmes estimate.

Zill asked whether a full-time hospital-based team was necessary. Kalton said that it would depend on the questionnaire. It might be possible to have one hospital person and a team member who visits regularly. Warburton commented that she thought Holmes would argue that the crucial person in this approach is the clinical teratologist so that the data would be better than in the previous system. She went on to say, however, that a clinical teratologist could take care of more than one hospital. Holtzman said that from a practical standpoint it might not be possible to have a part-time staff person or someone come in from the outside. Some hospitals will not tolerate the intrusion. He said the greatest skepticism and hostility in the Maryland system was from pediatricians and obstetricians who resented the intrusion on the care they provided their patients. A trained nurse might be appropriate for a delimited set of malformations. Nurses would not be viewed as intruders and they are generally very informed and cooperative.

The final consensus seemed to be that this approach could add more malformations and generally improve data quality. Subgroup analyses could be done, but not for local areas or even for most states. However, synthetic estimates of expected rates in local areas could be made and sampling errors provided. Additional variables could be added and there would be a standard methodology, which is very important.

## III. SURVEY APPROACHES

The approaches discussed in the morning were data gathering systems that would be ongoing. They would be best at collecting data on birth defects, but would not be very good at measuring miscarriages, infertility, or subfecundity. Survey approaches on the other hand would not be used to monitor the occurrence of particular outcomes and thus would not be ongoing systems. And, unlike the earlier approaches, they would be best at measuring miscarriages, infertility, and subfecundity. They would probably do less well at measuring malformations.

### 1. Use of Existing Data from Federal Surveys

Zill commented that there is a possibility of adding macro variables on levels of exposure to a major national fertility survey -- the National Surveys of Family Growth (NSFG). As its name implies, this survey is actually comprised of a series of surveys conducted at different times. If EPA provides a tape containing county level exposure data to the staff of the NSFG,

then they could link the tape to individual records within the NSFG surveys. At the moment plans for such an addition are in the talking stage. It would be possible to add such macro variables not only to the most recent cycle, but also to previous cycles of the survey as well. If all three cycles of the survey were used, there would be about 25,000 cases that could be linked to county level data.

In the 1982 survey there were 13,000 pregnancies recorded; the majority of these had occurred in the previous three years. Although ecological studies would be a problem because there is not information on any given individual's experience, crosstabulations and correlational studies could be done. Selevan stated that she would like to see some analyses done because the data set is large and, because it is interview based, it is more comparable with what would be done in a community.

Bill Mosher from NCHS was invited to describe the NSFG in more detail to the Working Group members. He described how the NSFG is an outgrowth of the original Family Growth Surveys begun in 1955. The most recent cycle, cycle III, of the current surveys was conducted in 1982. The main purpose of these surveys is to study the factors that affect the U.S. birth rate. They consist of in-depth, personal interviews with women of childbearing age. Cycle III was the first survey to interview never married women as well. Cycle IV of the NSFG is nearing completion of the planning stage. NSFG staff expect it to be in the field by 1987. There are going to be 11,000 women in the 1987 survey and current plans call for re-interviewing half of these women in a telephone follow-up two years later. (For a more complete description of the NSFG refer to the original packet of information that was mailed to Working Group members prior to the first meeting. The NSFG was one of several data sets that were described as being of potential use to EPA.)

Questions About the NSFG by Working Group Members

The Working Group members expressed a great deal of interest in the NSFG and asked Mosher several questions.

Kalton asked whether it would be possible to study birth defects. Mosher responded no -- there were too few cases for that. The NSFG, however, could be used to study fertility, subfecundity, miscarriage, and unusual sex ratios at birth. He defined subfecund as the inability to conceive after one year of unprotected intercourse. Infertility and subfecundity are measured indirectly in the NSFG through a series of questions on date of marriage, contraceptive use, and pregnancy history.

Mattison wanted to know whether it was possible to look at subfecundity by smoking history or other possibly imortant explanatory variables. Mosher commented that the NSFG only has smoking during the most recent pregnancy. It is possible, however, to look at subfecundity as well as other fertilty outcomes by a variety of control variables.

Warburton wanted to know what was asked in the questionnaire. Mosher said he was willing to send copies of the questionnaire to interested Working Group members. It is a long, detailed questionnaire that gathers in-depth data on contraceptive use, pregnancy history, sexual activity, and other aspects of women's lives that are related to their reproductive behavior.

Warburton also wanted to know whether it would be possible to double the sample size of the NSFG follow-up and, if so, at what cost. Mosher said he wasn't sure what the cost would be, but he thought it might be possible to increase the sample size. The cost would depend on how the sample was drawn. Kalton was not sure what purpose would be served by doubling the sample size. Holtzman asked if it would improve the confidence levels of any derived estimates.

Use of NSFG to Provide Baseline Data

Zill asked if it would be possible to use the NSFG to provide baseline data for EPA. Kalton said that you would have to tease out what are the areal effects. He thought there would be problems combining years. Mosher said that it would depend on what you were looking at. If you were looking at miscarriages, no there would not be problems. If, however, you were exploring contraception then there would be.

Zill clarified his question by saying he didn't mean to imply using the NSFG to provide regional data, but as a way to provide means, standard deviations, and adjusted expectations of the different outcomes. Kalton commented that the group had talked about synthetic estimates in the morning. He was not sure that using the NSFG would add much. Of course, the NSFG data are in hand whereas the other data are not.

Mosher added that although blacks were oversampled in the survey, they seem to underreport the number of pregnancies that they have had. Unmarried women also underreport the number of pregnancies.

## 2. Addition of Questions to Future Federal Surveys

Zill wanted to know if it would be possible to add questions to the 1987 questionnaire. Mosher said it would be difficult now because planning is almost completed. A few

questions and control variables could be added. For example, a more detailed question on smoking could be included. And questions could be added to the telephone follow-up that is planned for 1989. Mosher thought it might be possible to add about 10 minutes worth of questions to the follow-up. He was not sure what that would cost.

Someone asked whether it would be possible to ask about 10 or or 11 birth defects in the follow-up. Mosher was not sure whether it would be possible. Heuser did not think that birth defects would be well reported in the survey. Mosher added that if a particular defect occurred in less than 1% of births, then he did not think the NSFG would be an appropriate vehicle.

### 3. Design of Telephone-Based Survey Instrument for EPA's Use

Another approach is a survey that EPA could develop for its own use. At present, EPA lacks a method that can be used in local communities for determining whether particular outcomes are high. Although you get what you pay for, it is possible to get interviews with a national sample for about $30 per interview. EPA might consider conducting a national survey of its own with an instrument that they also intend to use in local areas. The national survey would serve as the benchmark for the local area studies.

#### Topical Outline of Proposed Telephone Survey

Zill and Nord handed out a topical outline (see Attachment 2) of a proposed telephone survey for EPA's use that they had prepared for the meeting. The general reaction to this outline was that it asked too many questions. It was thought best to keep such a survey as short as possible. Specifically, eliminate the search for possible exposure items. These could be explored in much greater detail in the local community if differences in outcomes were discovered between the local area and the national benchmark. Secondly, gather only basic information. Crucial control variables would include prenatal care, smoking history, alcohol use, prior pregnancy loss, age, reason for negative outcome (as told to Respondent by her physician), SES, and race.

In terms of what the appropriate population to sample would be, many felt that women aged 18-49 who had been pregnant or who had tried to become pregnant in the last five years was the correct population. Selecting age 18 as the lower age cut-off circumvents the problem of obtaining parental consent. There may be a problem, however, screening for this population. Others stated that you would want to gather some basic data on other women as well so that you would have a comparison group.

## Plans for the Third Meeting

It was agreed that rather than meet a third time, it would be more productive for members of the Working Group to spend a day of their time instead reviewing a draft of the final report to EPA which is to be prepared by Child Trends staff. This report will contain more detailed descriptions of each of the three approaches discussed above. Each member should add sections or make amendments where they deem it appropriate. Child Trends staff members may be contacting them to elicit information on specific approaches during the preparation of the draft report.

# BIBLIOGRAPHY

## WORKS CITED


Baird, D.D.; A.J. Wilcox; and C.R. Weinberg. 1986. "Use of Time to Pregnancy to Study Environmental Exposures." American Journal of Epidemiology 124:470-480.

Baird, D.D. and A.J. Wilcox. 1985. "Cigarette Smoking Associated with Delayed Conception." Journal of the American Medical Association 253:2979-2983.

Baldwin, W.H. and C.W. Nord. 1984. "Delayed Childbearing in the U.S.: Facts and Fictions." Population Bulletin 39: Population Reference Bureau, Inc.: Washington, D.C.

Bloom, A.D., Ed. 1981. Guidelines for Studies of Human Populations Exposed to Mutagenic and Reproductive Hazards. March of Dimes Birth Defects Foundation: White Plains, N.Y.

Bureau of the Census. 1984. Statistical Abstract of the United States: 1985 (105th Edition) Washington, D.C.

Centers for Disease Control. Birth Defects Branch. Chronic Diseases Division. Center for Environmental Health. 1985. "Reproductive Surveillance Activities at the State, National, and International Level: Brief Description of Program, State Laws, and Contact." Prepared by L.D. Edmonds. Unpublished Manuscript.

Centers for Disease Control. Chronic Diseases Division. Center for Environmental Health. 1983. "Uniform Methods to Collect Information on Mortality, Serious Diseases, and Human Exposure to Toxic Substances: Recommendations from an Epidemiologic Perspective." Unpublished manuscript.

Daniels, P. and K.K. Weingarten. 1979. "A New Look at the Medical Risks in Late Childbearing." Women and Health 4:17-18.

Earhart, E. and M.J. Sporakowski, Eds. "The Family with Handicapped Members." Special Issue Family Relations 38.

Edmonds, L.D.; P.M. Layde; L.M. James; J.W. Flynt; J.D. Erickson; and G.P. Oakley, Jr. 1981. "Congenital Malformations Surveillance: Two American Systems." International Journal of Epidemiology 10:247-252.

Hellman, L.M. and J.A. Pritchard. 1971. Williams Obstetrics, 14th edition. New York: Appleton-Century-Crofts.

Holmes, L.B.; S.E. Vincent; C.Cook; and K.R. Cote. 1981. "Surveillance of Newborn Infants for Malformations Due to Spontaneous Germinal Mutations." in Population and Biological Aspects of Human Mutation, 351-359. New York: Academic Press, Inc.

James W.H. 1970. "The Incidence of Spontaneous Abortion." Population Studies, 24:241-245.

Jason, C. and B.J. Glick. 1981. "Surveillance of Pregnancy Outcomes: A Proposed Program." Unpublished manuscript.

Mosher, W.D. 1985. "Reproductive Impairments in the United States, 1965-1982." Demography, 22:415-430.

National Center for Health Statistics. 1986. "Advance Report of Final Natality Statistics, 1984. Monthly Vital Statistics Report. Vol.35, No.4. Supplement DHHS Pub. No. (PHS)86-1120. Public Health Service. Hyattsville, MD.

National Center for Health Statistics. W.D. Mosher and W.F. Pratt. 1985a. "Fecundity and Infertility in the United States, 1965-82." Advance Data from Vital and Health Statistics. No. 104. DHHS Pub No. (PHS)85-1250. Public Health Service. Hyattsville, Md.

National Center for Health Statistics. 1985b. Vital Statistics of the United States, 1980. Vol. 1 Natality. Public Health Service, Washington. U.S. Government Printing Office.

National Center for Health Statistics. 1984. Health, United States, 1984. DHHS Pub. No. (PHS) 85-1232. Public Health Service. Washington: U.S. Government Printing Office.

National Center for Health Statistics. M.L. Dundon, G.A. Gay, and J.L. George. 1983. "The 1978 Revision of the U.S. Standard Certificates." Vital and Health Statistics Series 4, No. 23. DHHS Pub No (PHS)83-1460. Public Health Service. Washington: U.S. Government Printing Office.

National Institutes of Health. 1979. Antenatal Diagnosis. Report of a Consensus Development Conference sponsored by the National Institute of Child Health and Human Development and assisted by the Office for Medical Applications of Research and the Fogarty International Center, March 5-7, 1979. Bethesda, Md.

Nordstrom, S.; L. Beckman; and I. Nordenson. 1978. "Occupational and Environmental Risks in and Around a Smelter in Northern Sweden. I. Variations in Birth Weight." Herditas 88:43-46.

Nortman, D. 1974. "Parental Age as a Factor in Pregnancy Outcomes and Child Development." Reports on Population/ Family Planning No. 16. The Population Council: New York.

Pratt, W.F., W.D. Mosher, C.A. Bachrach, and M.C. Horn. 1984. "Understanding U.S. Fertility: Findings from the National Survey of Family Growth, Cycle III," Population Bulletin, Vol. 39, No. 5. Population Reference Bureau, Inc.,: Washington, DC.

Selle, H.F.; D.W. Holmes; and M.L. Ingbar. 1979. "The Growing Demand for Midtrimester Amniocentesis: A Systems Approach to Forecasting the Need for Facilities." American Journal of Public Health 69:574-580.

Stein, Z. and D. Warburton. "Report of Panel II. Guidelines for Reproductive Studies in Exposed Human Populations." In Bloom, Arthur D., Ed, Guidelines for Studies of Human Populations Exposed to Mutagenic and Reproductive Hazards. White Plains, NY: March of Dimes Birth Defects Foundation, 1981. Proceedings of Conference held January 26-27, 1981 in Washington, DC.

Ventura, S.J., S. Taffel, and W.D. Mosher. 1985. "Estimates of Pregnancies and Pregnancy Rates for the United States, 1976-81." Public Health Reports, 100:31-34.

Wilcox, A.J. 1983. "Surveillance of Pregnancy Loss in Human Populations." American Journal of Industrial Medicine 4:285-291.

Williams, L; M.A. Spence; and S.C. Tideman. 1977. "Implications of the Observed Effect of Air Pollution on Birth Weight." Social Biology 24:1-9.

SELECTED PERTINENT WORKS NOT CITED

Aral, S.; W. Mosher; and W. Cates, Jr. 1985. "Self-Reported Pelvic Inflammatory Disease in the United States: A Common Occurrence." American Journal of Public Health 75:1216-1219.

Hatch, M.; V.Stefanchik-Scott; and Z.A. Stein. 1983. Surveillance of Reproductive Health in the U.S.: A Survey of Activity Within and Outside Industry." Prepared for the American Petroleum Institute: Washington, D.C.

Hemminki, E.; B.I. Graubard; H.J. Hoffman; W.D. Mosher; and K. Fetterly. 1985. "Cesarean Section and Subsequent Fertility: Results from the 1982 National Survey of Family Growth." Fertility and Sterility 43:520-528.

Holtzman, N.A. and M.J. Khoury. 1986. "Monitoring for Congenital Malformations." Annual Review of Public Health 7:237-266.

Kalton, G. and D.W. Anderson. 1986. "Sampling Rare Populations." Journal of the Royal Statistics Society Series A, 149:65-82.

Pratt, W.W.; W. Mosher; C. Bachrach; and M. Horn. 1985. "Infertility: United States, 1982." Morbidity and Mortality Weekly Report 34:197-199.

Selevan, S.G. 1981. "Design Considerations in Pregnancy Outcome Studies of Occupational Populations." Scandinavian Journal of Work Environment Health 7 (Supplement 4):76-82.

Wilcox, A.J. and L.F. Horney. 1984. "Accuracy of Spontaneous Abortion Recall." American Journal of Epidemiology 120:727-733.

## COMPOSITION OF THE WORKING GROUP

The members of the Working Group on Human Reproductive Outcomes were:

Lewis Holmes, M.D., Massachusetts General Hospital. Dr. Holmes is a specialist in newborn surveillance and congenital malformations.

Neil A. Holtzman, M.D., M.P.H., Health Program, Office of Technology Assessment. Dr. Holtzman's focus is genetics. Furthermore, he was instrumental in setting up the Maryland Birth Defects Reporting and Information Service.

Casey Jason, M.D., is a practicing physician in northern Virginia with particular expertise on the issue of early miscarriage.

Graham Kalton, Ph.D., Survey Research Center, University of Michigan. Dr. Kalton is well-known for his work in survey sampling and statistics.

Donald Mattison, M.D., University of Arkansas School of Medicine. Dr. Mattison's research areas include reproductive toxicology and obstetrics.

Dorothy Warburton, Ph.D., Columbia University. Dr. Warburton is a human geneticist whose major research interests are cytogenetics and the etiology of embryonic and fetal death.

Janice Bakewell, from the Missouri Center for Health Statistics, represented a state that has developed an ongoing system for birth defects monitoring.

### Federal Agency Representatives

The following individuals represented their federal agencies at the Working Group meetings. Their participation in the Working Group should not be interpreted as agency endorsement of the findings and recommendations made in this report.

William Pratt, Ph.D, and William D. Mosher, Ph.D, National Survey of Family Growth, National Center for Health Statistics;

Melissa Adams, Ph.D., Birth Defects Branch, Centers for Disease Control;

Michael Gruber, Ph.D., Susan Perlin, and Sherry Selevan, Ph.D., Environmental Protection Agency;

Woodie Kessel, M.D., Division of Maternal and Child Health, BHCDA, HRSA, U.S. Public Health Service;

Pat Shiono, Ph.D., National Institute of Child Health and Human Development;

Clara G. Schiffer, Department of Health and Human Services;

Robert L. Heuser and Stephanie Ventura, Natality Statistics Branch, National Center for Health Statistics;

Peter Gergen, M.D., Division of Health Examination Statistics, National Center for Health Statistics.

Child Trends' staff also wishes to thank Delbert Dayton, M.D., of the National Institute of Child Health and Human Development and Allen Wilcox, Ph.D. of the National Institute of Environmental Health Sciences, who did not attend the meetings, but provided valuable assistance.

Child Trends Staff

The following members of the Child Trends' staff took part in the Working Group meetings and helped to draft this report:

Christine Winquist Nord
Nicholas Zill, Ph.D.
Kristin A. Moore, Ph.D.