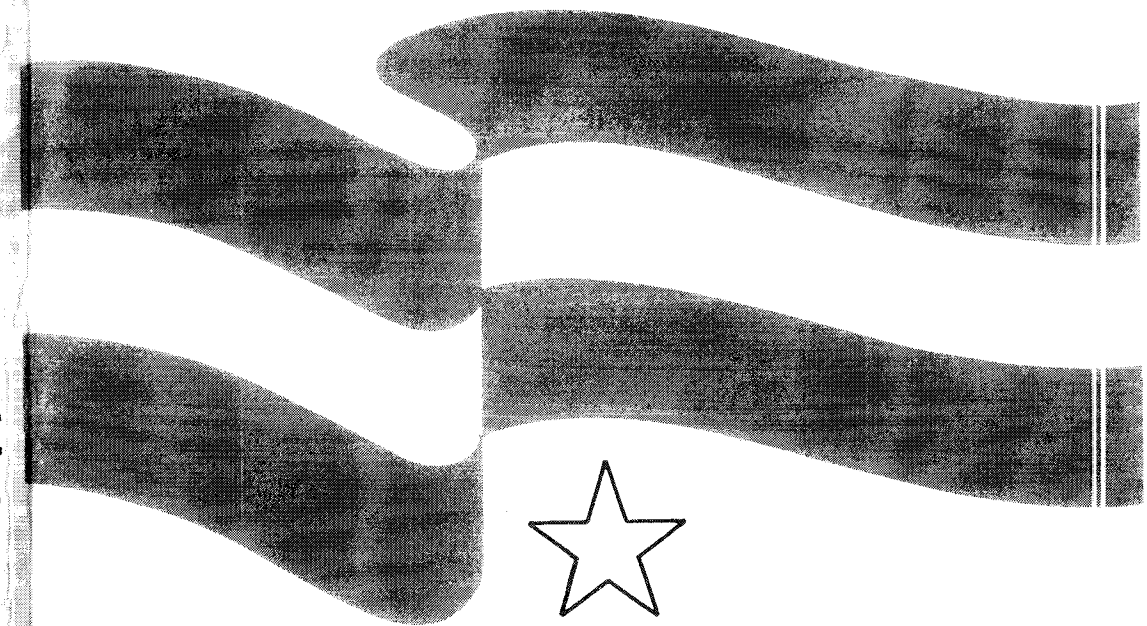

Vietnam Veterans' Risks for Fathering Babies with Birth Defects

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
CENTERS FOR DISEASE CONTROL

Vietnam Veterans' Risks for Fathering Babies with Birth Defects

VIETNAM VETERANS' RISKS FOR FATHERING BABIES WITH BIRTH DEFECTS

J. David Erickson
Joseph Mulinare
Philip W. McClain
Terry G. Fitch
Levy M. James
Anne B. McClearn
Myron J. Adams, Jr.

August 1984

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
Centers for Disease Control
Center for Environmental Health
Atlanta, GA 30333

Use of trade names is for identification only and does not constitute endorsement by the Putlc Health Service or by the U.S. Department of Health and Human Services.

ACKNOWLEDGMENTS

The authors thank the many individuals of several organizations who helped with this project. They include:

Lynn Bartlett, Sue Browning, Carlton Chapman, Richard Christian, Bruce Dupee, Donald Hakenson, Cmdr. Hugh Highland, Anna Mammone, Gerald Rice, Maj. Lorraine Shively, Maj. John White, and Joan Wilson of the U.S. Army Agent Orange Task Force.

Hugh Angle, Martha Berlin, Deborah Cesare, Robert Crandal, Philip Dow, Stephen Durako, Charles Eastlack, Susan Englehardt, Alan Jaworski, Helena Light, Thomas McKenna, Donna Matos, Stanley Mead, Doris Northrup, Virginia Osterlund, Sandra Sperry, Fauzia Temnazi, Mary Usher, and the Interviewing and Tracing staffs of Westat, Inc., of Rockville, Md.

John Adams, U.S. Postal Service; Margaret Aldredge, Georgia Department of Human Resources; Jerome Bricker, U.S. Department of Defense; David Brown, National Institute for Occupational Safety and Health, CDC; Samuel Hider, Atlanta Housing Authority; Cpt. Herschel Hyde, Drivers Support Division, Georgia Department of Public Safety; Margaret Karlsen, Internal Revenue Service; Michael Lavoie, Georgia Vital Records Unit; Barclay Shepard, Veterans Administration; and Jacquelyn Vernon, Southern Bell.

We also thank the many members of CDC's staff who contributed to this project over the past 4 years. We especially thank Janice Calloway, Marguerite Canfield, Henry Cassell, José Cordero, Jacquelyn Crumpton, Edwina Davis, Dorothy Dougherty, Larry Edmonds, Barbara Ervin, Evelyn Finch, Sharon Gentry, Richard Goodman, Helen Hankins, Clark Heath, Martin Hirsch, Vernon Houk, James Hudson, Martha Hunter, Bobbie Lev, Brenda Mitchell, Meade Morgan, Godfrey Oakley, Anne Olin, Brian Plykatis, Cynthia Rogers, Joyce Smith, Norma Strawn, Carolyn Tucker, Patricia Weis, Paul Wiesner, and Matthew Zack.

TABLE OF CONTENTS

SUMMARY

Vietnam veterans' risks for fathering babies born with major structural birth defects were assessed by using a case-control study design. Case group index babies were registered by the Centers for Disease Control's (CDC's) Metropolitan Atlanta Congenital Defects Program during the years 1968-1980. With data from multiple sources, this program ascertains live and stillborn babies with structural defects born to women who are residents of the five-county metropolitan Atlanta area. To be included in the registry, a baby's birth defects diagnosis must have been made during the first year of life and recorded in a hospital chart by a physician. Control group index babies—babies born without defects—were selected from among the 323,421 babies who were live born in the area to resident mothers during 1968 through 1980. The control group index babies were frequency matched to the case group index babies by race, year of birth, and hospital of birth ("sampling design variables"). The number of case group families eligible for the study was 7,133, as were 4,246 control group families.

Information about paternal military service in Vietnam was obtained during 1982 and 1983 through interviews with the index babies' parents; information about a wide variety of other factors that might be associated with the occurrence of defects was also gathered during the interviews. Separate interviews were done with babies' mothers and fathers, by using a computer assisted telephone interviewing system. Interviewed Vietnam veteran fathers were asked if they felt that they had been exposed to Agent Orange while in Vietnam. In addition, they were given scores (ranging from 1 to 5) reflecting subjectively estimated opportunities for exposure to Agent Orange (that is, a score on the Agent Orange Exposure Opportunity Index). These scores were based on location in Vietnam, period of service, and occupational duties, and were given without knowledge of the fathers' case/control group status. All men who had served in Vietnam were also asked if they had contracted malaria while there, and whether they had used malaria chemoprophylactic medicines.

Before parents could be interviewed, they had to be located. Because many of the index babies were born many years before the interviews were done, locating the parents was often difficult. In all, 4,929 case group mothers and 3,029 control group mothers completed interviews (70% participation rate). Fewer fathers were interviewed: 3,977 from the case group and 2,426 from the control group (56% participation rate). The participation rates were lower for "Other race" than for "White race" parents, particularly for fathers; this may somewhat limit our ability to generalize about the results of the study. The major reason for nonparticipation was inability to locate parents.

For the purposes of data analysis, the birth defects that affected the case group index babies were classified into 96 groups, including 1 group comprising all types of defects combined. For each of these 96 groups four hypotheses were tested: 1) Whether veterans (excluding Vietnam veterans) were at a different risk for fathering babies with defects than nonveterans. The purpose of this test was to decide if Vietnam veterans should be compared with all other fathers, or only with veteran fathers for the remaining three hypotheses. 2) Whether Vietnam veterans were at a different risk for fathering babies with birth defects. This hypothesis was the major focus of this study. 3) Whether Vietnam veterans who received higher Agent Orange Exposure Opportunity Index scores were at a different risk than other men. 4) Whether Vietnam veterans who said they believed that they had been exposed to Agent Orange were at a different risk than other men.

Each of the hypotheses for each of the 96 defect groups was evaluated three times. First, the hypotheses were evaluated with only the potentially confounding sampling design variables being considered. At this stage of the analysis, the possibilities that the risks varied with

time of birth and with race were assessed. Second, the hypotheses were evaluated with the sampling design variables and four other potentially confounding covariables being considered. These four were 1) maternal age, 2) maternal education, 3) maternal alcohol consumption, and 4) birth defects in first-degree relatives of the index babies. These four covariables were identified prior to data analysis as being particularly important. Third, the hypotheses were evaluated with 108 potentially confounding maternal and paternal covariables being considered. These covariables were considered singly, each along with the sampling design variables. The main analytical tool used was conditional logistic regression; the Mantel-Haenszel procedure was, however, used for some analyses.

Four hundred and twenty-eight fathers of case group index babies were (said by mothers to be) Vietnam veterans and 268 fathers of control group babies were Vietnam veterans; the non-Vietnam veteran case and control group fathers numbered 4,387 and 2,699, respectively. For the analysis that gave no consideration to potentially confounding variables (except for the sampling design variables), the logistic regression-derived odds ratio for all types of defects combined was 0.97, with 95% confidence limits of 0.83 to 1.14. This estimate provides no support to the notion that Vietnam veterans, in general, have been at an increased risk for fathering babies with birth defects than other men.

With few exceptions, the same type of finding applied to Vietnam veterans' risks for the remaining 95 defect groups. The same overall pattern also applied to the tests of hypotheses regarding the Agent Orange Exposure Opportunity Index and those regarding Vietnam veterans' self-reports of Agent Orange exposure. Three exceptions to this general pattern are noted below; those instances in which Vietnam veterans, or subsets of them, appear to have been at lower risk than other men are not presented here.

1) The estimated risks for fathering babies with spina bifida were higher for Vietnam veteran fathers who received the higher Agent Orange Exposure Opportunity Index scores. No similar association was found with the related defect, anencephalus.

2) Vietnam veterans who had higher Exposure Opportunity Index scores had higher estimated risks for fathering babies with cleft lip with or without cleft palate.

3) The estimated risks for fathering babies with defects classed in the group "Other Neoplasms" was higher for Vietnam veterans who had higher Exposure Opportunity Index scores. The neoplasms classified in this group include teratomas, neuroblastomas, hamartomas, and dermoid cysts and similar problems.

When potentially confounding covariables (other than the sampling design variables) were considered, results of analyses were similar to those obtained when they were not considered. This indicates that the results of this study were not a consequence of confounding by these variables, or factors.

Vietnam veterans who stated that they had contracted malaria while in Vietnam had a higher estimated risk for fathering babies born with hypospadias. No associations were found between the use of malaria chemoprophylactics and defects risks.

This study provides strong evidence that Vietnam veterans, in general, have not been at increased risk of fathering babies with the aggregate of the types of defects studied here. The use of CDC's birth defects registry precluded the study of other reproductive issues of concern to Vietnam veterans, for example, infertility, or purely functional deficits in Vietnam veterans' offspring.

Assessing Vietnam veterans' risks associated with Agent Orange exposure is difficult. The measures available today for estimating exposure are, at best, imperfect. Thus, we do not know whether the few positive associations found in this study reflect true effects of exposure or merely represent chance occurrences. The fact that Vietnam veterans in general do not appear to have been at increased risk suggests, however, that if effects have been caused by exposure, those effects are small, are limited to select groups of veterans, and/or are limited to rare types of defects.

1. BACKGROUND

This report describes the methods and results of a case-control study designed to determine if Vietnam veterans are at an increased risk of fathering babies born with major structural congenital malformations (serious birth defects). A major concern is whether those Vietnam veterans who were potentially exposed to the dioxin-containing herbicide "Agent Orange" are particularly at risk. The study was based on the experiences of parents of babies born with and without serious birth defects in the metropolitan Atlanta, Georgia, area during the years 1968 through 1980. Some information about babies and their parents was derived from vital, hospital, and military records; other relevant data were obtained through telephone interviews with parents.

1.1 RATIONALE FOR CONDUCTING STUDY

Over the past few years public concern has mounted regarding possible adverse health effects experienced by United States and Australian military personnel who served in Vietnam. These concerns have been widely reported by the news media and have generated U.S. Congressional interest and support. Among a wide range of alleged adverse health effects, the possibility that the incidence of birth defects is increased among the offspring of Vietnam veterans has been of particular concern. When the present study was designed (1980), there was no scientifically acceptable human evidence directly relevant to the claim that Vietnam veterans might be at increased risk. Anecdotes had, however, been reported in letters to the editor in scientific journals (Bogen, 1979), and these anecdotal reports have continued (e.g., LaVecchio et al., 1983). Moreover, reports emanating from Vietnam since the end of the Vietnam war have put forth the claim that Vietnamese men who served in the war in the south were at increased risk, although these reports have not been published in scientific journals. A study of the risk for Australian Vietnam veterans of fathering babies with birth defects was recently completed. The study did not confirm the suspicion of an increased risk (Donovan, 1983).

The alleged increase in risk for fathering babies with birth defects is thought by some to be the consequence of exposure to herbicides (especially the herbicide known as Agent Orange, see below) which were used extensively by U.S. troops in Vietnam. Herbicides were used for three principal purposes: (1) defoliation—to cause trees and plants to lose their leaves in order to improve observation; (2) crop destruction—to destroy the food value of certain crops; and (3) on a smaller scale, to clear vegetation around fire bases and other installations, around landing zones, and along lines of communication. The use of herbicides during the Vietnam war began in 1962, was greatly expanded during 1965-1966, and peaked from 1967-1969. In 1969 birth defects were reported in offspring of female mice exposed to certain herbicide components (Hayes, 1982). Between 1970 and 1971 the use of herbicides was phased out in Vietnam.

The tactical military project for the aerial spraying of herbicides in South Vietnam was named "Operation Ranch Hand"; this program used fixed-wing aircraft and disseminated the bulk of the herbicides used in Vietnam (Young et al., 1978). Smaller quantities of herbicides were applied from helicopters, trucks, and riverboats and by hand applicators. At least two groups of U.S. personnel appear to have been at risk for exposure to herbicides—(1) those involved in the transport and dissemination of the chemicals and (2) those who may have been exposed at the time of spraying, such as troops on the ground. Although exposures may have occurred during transportation (e.g., due to damage of containers), aircraft crew—particularly flight mechanics and crew chiefs—were thought to be at greatest risk of exposure. Even though the major portion of herbicides used was disseminated by Operation Ranch Hand, a significant and even a major source of exposure of ground troops may have been from non-Ranch Hand applications.

Herbicides used for military purposes during the war were identified by color bands on their containers (e.g., orange, white, and purple). The herbicide now known as "Agent Orange" was the most widely used in Vietnam. It was a 50:50 mixture by weight of the butyl esters of two phenoxy acid herbicides, 2,4-dichlorophenoxy acetic acid (2,4-D) and 2,4,5-trichlorophenoxy acetic acid (2,4,5-T). The mixture also contained TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin, "dioxin"), a synthetic contaminant of 2,4,5-T (Young et al., 1978).

TCDD is one of the most toxic compounds known, at least for some animal species (Hayes, 1982). In experiments with animals, a wide variety of adverse health effects have been observed after TCDD administration. In humans, toxic effects have been reported after occupational exposure during the industrial synthesis of 2,4,5-trichlorophenol (TCP) and 2,4,5-T, after exposure in factories and in the surrounding environment following explosions that occurred during the synthesis of TCP, and after exposure to herbicides and other materials containing TCDD. Reproductive effects in animals appear to be limited to maternal (fetal) exposures (Courtney et al., 1970; Courtney and Moore, 1971; Gaines et al., 1975; Smith et al., 1976). The few studies that have addressed the possibility of paternally mediated effects have not shown differences in the rates of poor reproductive outcomes between the exposed and nonexposed (Murray et al., 1979; Lamb et al., 1980). The human data on reproductive outcomes after exposure are also generally negative, but most specific poor reproductive outcomes are rare and the studies of men exposed in industrial settings have been of relatively small numbers (Townsend et al., 1982).

When this study was designed, there was little evidence in scientific reports to suggest that paternal environmental exposures could result in the birth of babies with defects. This is still the case, and the situation is in marked contrast to the well-recognized potential for the embryotoxic effects of maternal exposures, for example, the defects caused by rubella or thalidomide. However, the causes of the overwhelming majority of birth defects are unknown and there is undoubtedly a possibility that some could result from some form of paternal effect. Furthermore, the possibility of paternal contributions has not been widely investigated.

Based on reasonable biological models, paternal contributions to the occurrence of birth defects deriving from environmental exposures could arise in two general ways. First, an environmental insult could cause a change in a father's sperm cells, either as a single gene ("point") mutation or as a chromosomal alteration. Older fathers are known to have a greater risk of having babies with defect syndromes due to point mutations (e.g. Apert syndrome, some achondroplasias; Blank, 1960; Penrose, 1957); paternal chromosomal nondisjunction is now thought to be the origin of perhaps 25% of babies born with trisomy 21 (Down's Disease; Hook and Regal, 1984). Whether these paternal effects are the result of environmental insults is not known. Second, embryos and fetuses might be exposed to toxic substances through the agency of the father. For example, the mother (and hence the fetus) could be exposed to substances on the father's clothes. Even more direct paternally derived embryonic or fetal exposures are possible, as, for example, by contamination of the semen with a toxic chemical. The notions that such exposures occur and that they might cause birth defects are, however, purely speculative.

The concern that Vietnam veterans are at an increased risk of fathering babies with birth defects deserves to be viewed from another perspective. During the past decade or so, U.S.

Vietnam veterans have probably fathered somewhere between 50,000 and 160,000 babies born with serious birth defects! These figures may seem alarming, but they are merely estimates based on some assumptions about the fertility of Vietnam veterans and the usual frequency of serious birth defects. About 2.6 million American men served in Vietnam (Veterans Administration, 1981). The vast majority of them were in their late teens or early 20's when they were stationed in Vietnam; therefore, since the war, they have gone through the ages that are generally the most fertile. Suppose that since these men left Vietnam, they have had an average of one baby each. Further, take the usual frequency of serious birth defects to be 2% to 3%. This would have resulted in the birth of 52,000 to 78,000 babies with defects whose fathers were Vietnam veterans. Similarly, if Vietnam veterans have had an average of two babies each, the numbers of affected babies would be double those suggested above.

The points to be taken from this discussion are that birth defects are relatively common, that even in the absence of an increased risk, many Vietnam veterans will have had affected babies, and that anecdotal reports of relatively common events do not provide a basis for making inferences about excess risks among Vietnam veterans, particularly in the absence of well-documented evidence of paternal effects due to environmental exposures. Even so, it is possible that men who served in Vietnam could have been exposed to toxic substances that conceivably could have increased their risk of fathering babies with birth defects. These facts and the intense public interest in this issue prompted the study described here.

1.2 STUDY DEVELOPMENT AND SUPPORT

In 1979 the Centers for Disease Control (CDC) became aware of the concern about Vietnam veterans' risks for birth defects among their offspring. Members of the staff of CDC's Birth Defects Branch realized that its Metropolitan Atlanta Congenital Defects Program registry was a unique resource for developing a study to address this issue of great public import. The same year an internal memorandum describing the basics of such a study was prepared. In 1980 then President Jimmy Carter formed an interagency working group to foster and oversee federal activities relevant to veterans' and others' concerns about the possible adverse health effects of phenoxyacetic acid herbicides and TCDD. In the spring of 1980, CDC staffers presented the concept of the present study to this working group. The group encouraged further development, and in the summer of 1980 a full draft protocol was completed.

The draft protocol was given "human subjects" and peer scientific review by members of the CDC staff employed outside of the operating component responsible for conducting the study. CDC invited four nongovernment scientists to come to Atlanta and provide further scientific review, and a scientific subcommittee of the interagency working group mentioned above also reviewed the draft protocol. Several veterans' service organizations were also invited to comment on the draft protocol. In the light of suggestions from these reviewers, the draft protocol was revised, and the revision was submitted to the Office of Management and Budget for final approval.

In September 1980, CDC, the Veterans Administration (VA), and the Department of Defense (DoD) signed an Interagency Agreement to provide support for the study (throughout the study CDC provided its own funds and personnel), and the VA provided funds to CDC. The DoD provided funds to CDC and also provided support for work done for the study by the Army Agent Orange Task Force. CDC contracted with Westat, Inc., of Rockville, Md., to locate and interview the parents of the study babies. In December 1981 a field test, or pilot study, of 113 case and 107 control group families was begun. This was completed in May 1982, and data collection for the main study began in May 1982 and was completed in the fall of 1983.

1.3 STUDY OBJECTIVES

The primary objective of this study was to determine whether Vietnam veterans have been at an increased risk for fathering babies with serious structural birth defects. Even though many veterans have focused their concern on Agent Orange and TCDD exposure, this study was deliberately designed to take the broader approach of assessing risks among Vietnam veterans in general. One reason for this is that because accurate measurements of exposure to Agent Orange are difficult to obtain, it is difficult to group veterans on the basis of exposure. Moreover, many observers apparently believe that service in Vietnam is (in and of itself) an indication of exposure and, hence, of risk. Despite this broad approach and the difficulties of accurately assessing Agent Orange exposure, we attempted to assess the possible effects of exposure by using information gathered during interviews with Vietnam veteran fathers and data compiled by the Army Agent Orange Task Force (AAOTF).

The objective of assessing the birth defects risk can be considered in specific and general terms. In specific terms, the risks can be assessed for particular types of defects. The causes of most types of defects are unknown, but it is generally believed that different types of defects are the result of different pathogenetic processes. Considered from this point of view, some environmental factor or particular exposure encountered by men serving in Vietnam would not be likely to result in an increased risk for many types of defects. Therefore, risks can be assessed separately for different groups of defects. In general terms, the risk can be assessed for the totality of defects. This assessment might have little meaning from a causal perspective, but it will be of substantial interest to veterans who want to know if they are at an increased risk of having a baby with any type of serious birth defect and how much greater that risk is than the usual 2% to 3% background risk.

2. STUDY METHODS

This study was conducted in the "case-control" format (MacMahon and Pugh, 1970). That is, one determines which fathers of case group index babies (those with specified birth defects) and which fathers of control group index babies (those without defects) served in the military in Vietnam. The frequency of such service in the fathers of case and control babies can then be compared. If Vietnam veterans are at increased risk of fathering babies with birth defects, then a larger proportion of Vietnam veteran fathers would be expected to be in the case group than in the control group. On the other hand, if Vietnam veterans have the same risk as all other men, equal proportions of Vietnam veterans would be expected to be among the fathers of case and control group babies. The same types of comparisons can be made for measures of possible exposure to Agent Orange.

2.1 SOURCE OF CASE GROUP INDEX BABIES

CDC's ability to conduct this study was greatly facilitated by its Metropolitan Atlanta Congenital Defects Program (MACDP) registry (Edmonds et al., 1981). Since October 1967, the MACDP has attempted to ascertain all live-born and stillborn babies with structural and biochemical congenital defects whose mothers were residents of the five-county metropolitan Atlanta area at the time their babies were born. In 1970 the area population was about 1.4 million, and in 1980, 1.6 million, and during the past decade, it has had about 25,000 births per year; about two-thirds of these babies have been born to White* mothers and one-third to mothers of Other* races. The program, sponsored by Emory University School of Medicine, the Georgia Mental Health Institute, and CDC, has the support of the local medical community. CDC is responsible for the day-to-day program operation. Program staffers aggressively search for cases during regular visits to all area hospitals that have obstetric or pediatric services. Search methods include contacting nurses on duty in the pediatric and obstetric units and reviewing pathology department records, obstetric logs, and hospital disease indices. In circumstances suggesting that a baby may have a congenital defect, the hospital charts of both baby and mother are thoroughly reviewed. If the presence of a malformation is confirmed (defined as a diagnosis made by a physician and recorded in a hospital chart or a record of a laboratory test confirming certain types of anomalies) in a baby who weighs $\geq$500 gm and is of $\geq$20 weeks' gestation, an abstract is made on an MACDP case-record form (Figure 1). A baby is included in the registry only if the defect is diagnosed before the baby reaches the age of 1 year. Other sources of initial case ascertainment include local cytogenetic laboratories and the State of Georgia's live birth, fetal death, and infant death certificates; abstracts of hospital charts are made regardless of which type of record prompted ascertainment.

The study case group index babies were selected from babies born during the years 1968 through 1980 and who were registered by the MACDP. Those chosen are only a subset of all those registered—about 13,000 babies born during these years were registered. In general, only those babies with serious or major birth defects were selected for the study. Families of babies born with relatively minor anomalies only were not selected, primarily because of resource constraints. The study of each family required considerable effort, and we decided that the effort should be devoted to families of babies born with "major" or "serious"

*In this report, the term "White" refers to persons of Caucasoid origin and the term "Other" to persons of all other racial origins. For births in the Atlanta area, the term "Other" is nearly synonymous with the term "Black," since there have been very few births to mothers of Oriental, American Indian, etc. racial origins.

defects, i.e., those that affect survival, require substantial medical care, result in marked physical or psychological handicaps, or interfere with a baby's prospects for a productive and fulfilling life. The study of families of babies who had defects that were not considered "major" would necessarily dilute the effort available for families of babies with significant problems and the study would have taken considerably longer to complete.

In the course of MACDP's operation, the English descriptions of birth defects diagnoses abstracted from hospital charts have been transformed into codes specified by a modification of the International Classification of Diseases, 8th Revision ("ICD-8"; World Health Organization, 1967). The code numbers used are, for the most part, those from the Congenital Malformations chapter of the ICD-8 (numbers 740.0 - 759.9). Some defects registerable by the MACDP have ICD-8 codes outside the 740.0 - 759.9 range; they have been given special codes, with the prefix "S." Table 1 lists these diagnostic rubrics, along with the numbers of babies in each category who were potential candidates* for inclusion in the study (babies born with more than one type of defect are counted in each relevant category). The rubrics in the table are divided into three categories:

Category 1. Any baby who had one or more of the defects signified by these rubrics was automatically deemed eligible to be a study case group index baby. Most, but not all, babies with these codes have "major" or "serious" defects. For example, hypospadias can be a serious or relatively trivial problem; ventricular septal defects can vary from rather inconsequential to life-threatening. Unfortunately, the records available to the MACDP in the hospitals do not usually permit the "seriousness" of such cases to be assessed. Those babies who had patent ductus arteriosus and no other major defect in this category were treated specially. A patent ductus arteriosus is a normal part of the fetal circulatory system. Normally, the ductus closes shortly after birth. In some instances, a persistently patent ductus represents a "true" defect, but in others the persistent patency is in some way related to premature birth, birth before the ductus' "time to close." Because of this, babies with diagnoses of patent ductus arteriosus (and no other major defect) were chosen for study only if they weighed 2,500 gm or more at birth. Other methods used in an attempt to separate relatively minor defects from more serious defects grouped together under single codes will be described later.

Category 2. Babies with defects signified by the rubrics in this category and none in Category 1 required special review. These categories are generally titled "Other Specified..." or "Anomalies of..." and code for a variety of malformation types, some major and some minor. The case-history form of each of these babies was reviewed simultaneously by two of us (JM and JDE), and separate decisions were made about the including or excluding of specific babies.

Category 3. Any baby who had only the anomalies indicated by these rubrics was automatically excluded from the study. For the most part, these anomalies are relatively minor. Others in the class are "unspecified" defects of certain organs where a decision regarding seriousness could not be made from the available records and those whose etiology is clearly known and probably not associated with paternal effects (e.g., congenital rubella syndrome). Also included in this category are a number of inborn errors of metabolism

*This table was created in late 1983 from the master MACDP computer data tape that was then current, not from the master tape current in September 1981, when the final choice of study cases was made. From time to time, as errors are identified, data on the master tape are corrected. Corrections of major consequence are very infrequent, and thus the late 1983 master tape is almost identical to the September 1981 tape with respect to the numbers of cases available for the study.

Most of these result from the inheritance of recessive genes. Even though the occurrence of these disorders might provide a marker for new mutations, such an inference in the context of this study would require that we show which fathers of these babies are not normal heterozygote carriers of the recessive allele (the vast majority of these babies will be born to normal carrier parents who themselves have inherited the allele). Technically, such a demonstration is not possible for all these disorders, and where it is possible, it is far beyond the scope of this study.

In the routine operation of the MACDP registry, certain types of defects that can be considered as "secondary" to other defects are excluded from tabulations, and these conventions have been followed in this report. Here are three of the important conventions: 1) babies with anencephalus and spina bifida without hydrocephalus are counted only under the anencephalus rubric; 2) babies with a neural tube defect and clubfoot or dislocated hip are only counted under the neural tube defect rubric; and 3) babies with upper and lower limb reduction deformities are counted only under the upper limb defect rubric (see Table 1).

This system of categorization and special review resulted in the initial acceptance of 7,530 registrants as study case group index babies. Three hundred and ninety-seven of these were later excluded for the following reasons:

1) evidence on the case-history form indicating that the baby had been, or would be, given up for adoption (111 babies).

2) a physician's indication that he or she preferred that the parents not be interviewed. Letters were sent to the physicians of record informing them of CDC's intention to interview the babies' parents. The physicians were asked to let CDC know if they preferred that CDC not make a contact. If the physician objected, the baby was excluded (15 babies). The procedure of contacting physicians before contacting families of babies with birth defects is part of the understanding between CDC and the local medical community that allows the MACDP to operate.

3) having one or more siblings as a case group index baby. A list of all study case group index babies was prepared and arranged by mothers' names. This list was reviewed manually, and when two or more babies appeared to be siblings (based on parental names and other available demographic data), one was randomly chosen for inclusion and the remaining siblings were excluded (113 babies).

4) duplicate registration. Review of the list of case group index babies arranged by mothers' names identified 22 babies registered twice by the MACDP.

5) erroneous registration. Eight babies initially selected were found to have been erroneously registered by the MACDP.

6) inclusion in the pilot study. One hundred and thirteen babies chosen for the pilot study were excluded.

7) having siblings in the control group. Some case babies had siblings who had been chosen as controls; their mothers' names were different when the case and control index babies were born. In three of these instances, the mothers were reached first regarding the control group index babies and only later were they discovered to also have case group index babies. These three babies were retained in the control group, and their siblings were excluded from the case group.

8) error in the selection process. Parents of 12 babies who had congenital rubella or toxoplasmosis and defects in Category 1 were made eligible for location and interview. These babies are not included in any of the data presented here.

Thus, 7,133 case babies were deemed eligible for the main study (Table 2); of these, 6,815 had one or more Category 1 defects, and 318 had only Category 2 defects.

As will be described, the hospital-of-birth distribution of case group index babies was considered in the choice of control group babies. Case group babies who were not born in one of the area hospitals (e.g., born in an ambulance, born to Atlanta residents in other geographical areas) were attributed to one of the area hospitals for the purposes of analyzing data for this report. Most were attributed to the Atlanta area hospital from which the MACDP ascertainment was made, but 32 were attributed by another method. These 32 babies were ascertained by the MACDP from one of the two area hospitals that specialize in the care of children. Neither hospital has obstetric services, and the 32 babies were therefore attributed to a hospital that does provide obstetrical services. This was done on the basis of the overall race-specific distributions of case babies by hospital of birth. For example, if 10% of White race case babies were born in a particular hospital, then a White race baby requiring attribution had a 10% chance of being allocated to that hospital.

The focus on babies registered by the MACDP precluded the study of other reproductive issues which may be of concern to Vietnam veterans, such as sterility, infertility, early fetal death, late fetal death, infant mortality, and behavioral derangements uncomplicated by malformations.

2.2 STUDY POWER

The ability of this study to detect an increased risk among Vietnam veterans, if indeed there is an increased risk, depends on several factors. Most prominently, it depends on the numbers of parents of case and control group babies interviewed, the proportion of fathers of control group babies who served in Vietnam, the magnitude of the increased risk, and the level of significance chosen. At the time this study was designed, CDC could only roughly estimate the proportion of Vietnam veterans among fathers of babies without defects. At that time the Atlanta office of the VA estimated that about 25,000 Vietnam veterans lived in the Atlanta area. Although the fertility of these men was unknown, it was thought that it would be reasonable to estimate that their babies would represent between 10% and 20% of all babies born in the area during the years 1968 through 1980.

Figures 2 and 3 show power curves similar to those used to decide about the number of control babies to include in the study (the computational procedure used to derive these curves was described by Oliphant and McHugh, 1981). The curves are drawn as a function of the number of cases for various increases in risk, as estimated by the odds ratio (2 X 2 table cross-product ratio). The maximum number of cases available is fixed (by virtue of using the MACDP registry as the source), but because babies born without defects are plentiful, the number of controls could be varied at will. Therefore, curves for varying numbers of controls were assessed for each risk category. The curve in Figure 2 was based on the assumption that 10% of fathers of control group babies would be Vietnam veterans, and Figure 3 was based on the assumption that 20% would be. The figures show that for types of defects that affected small numbers of babies the study was expected to have low power, even with large numbers of controls and a 1.5-fold increase in risk. Even for the types of defects with larger numbers affected, power was expected to be low for smaller increases in risk (e.g., 1.2-fold). On the other hand, for defects affecting a substantial number of babies, the power was expected to be relatively good for detecting increases of 1.5-fold or greater.

On the basis of the information contained in these figures, we decided to include 3,000 controls in the study. We chose this number on the basis of the power to be achieved for detecting increases in risks for all defects combined, and the power to be obtained for specific types of defects. Three thousand controls were expected to give fairly good power to detect small increases in the risk for all defects combined (about 5,000 cases) for Vietnam veterans in general; for example, this was expected to result in a power of between 0.7 and 0.9 to detect an increased risk of 1.2 (Figures 2 and 3). Higher power could be achieved for all defects combined by including more controls. However, we did not plan to subdivide the control group into smaller groups for the purpose of evaluating risks for specific types of defects. Instead, we planned to use the whole control group to evaluate each specific type of defect of interest as well as all defects combined. For specific types that have affected small numbers of babies, adding more numbers to the control group would do little to improve the power (Figures 2 and 3). The specification of fewer controls than cases is unusual—in most case-control studies the number of controls equals or exceeds the number of cases.

2.3 SOURCE OF CONTROL GROUP INDEX BABIES

Control group index babies (babies without birth defects) were chosen from among the 323,421 live births that occurred in area hospitals during the years 1968 through 1980 to women who were residents of the five-county metropolitan Atlanta area. Selections were made from computer tapes containing data coded from the birth certificate records of the State of Georgia (Figure 4). At the time the controls were chosen, CDC was uncertain about how successful it would be in securing the cooperation of control families in the data gathering effort. In addition, a substantial amount of time was required for the Georgia Vital Records Unit to provide CDC with copies of certificates. Initially, therefore, it was decided that about 6,000 potential controls should be selected, since, if the participation rate turned out to be low, this number would provide an ample number readily at hand.

The initial control group series was chosen by a stratified random sampling procedure so that the series would be approximately frequency matched to the final case series for race, year, quarter of birth, and hospital of birth (sampling design variables). Frequency matching implies only that the case and control groups will have similar distributions with respect to the sampling design variables; it differs from the usual matching strategy in which cases and controls are individually matched. The groups were frequency matched because we were concerned that the ability to locate babies' parents and conduct an interview would depend on the accuracy and completeness of information gathered at the time of the baby's birth and that this might vary from hospital to hospital, over time, and among the two major race groups. In addition, we were concerned that the frequency of prior Vietnam service among the babies' fathers might also vary with these factors—in other words, the sampling design variables were considered to be potentially confounding variables. These variables would have to be controlled for in the analyses whether or not they were used as sampling design variables. Since the usual procedure for controlling for such variables involves stratification, strata with no controls would have to be dropped from the analyses. Therefore, frequency matching was also done to minimize the number of cases that would be lost to the analyses because they belonged to strata with absent controls.

Although the frequency matching made the proportional distributions of the case and control series similar with respect to race and time and hospital of birth, the distributions were not exactly proportional for two reasons. First, at the time that most of the controls were chosen the final composition of the case series was unknown because review of Category 2 defects (see section 2.1) had not been completed. Second, the number of potential controls to be chosen was smaller than the number of potential cases. For strata with small numbers of cases, a specified control-to-case ratio could not be realized. In the extreme situation where a race/date of birth/hospital stratum contained one case, the control number for that stratum was forced to be 1. Otherwise, the number of controls selected for a particular stratum was based on a fraction of the number of cases in that stratum, rounded to the nearest whole number. Because of misestimation, this sampling fraction was set too high, and because a minimum of one control was chosen for any stratum containing one or more cases, the final number of initially selected controls was 6,677. This caused no problems, since 6,000 was chosen as the desired number of initially selected controls merely to have an adequate number in the event the participation rate turned out to be low and regardless of the number of Category 2 defects ultimately included in the study.

After the pilot study of 113 case group families and 107 control group families was completed, we estimated that about 70% of case and control families would cooperate in the data gathering process. Therefore, a subsample of about 4,400 was selected from the 6,677 initially chosen controls. This was done by using a stratified random sampling technique that maintained the approximate frequency match of the final control series to the final case series for race, year of birth, and hospital of birth (but with the provision that any stratum containing a case must have a minimum of one control). Although the control series was frequency matched to the complete case series (i.e., all defects combined), it was not necessarily balanced with the distributions of race and time and hospital of birth for specific types of defects.

As with the case series, certain initially eligible controls were excluded. Unlike the case series, however, no letters were sent to physicians, since the State of Georgia provided CDC with permission to contact control families ascertained from its records. Before the control series was reduced to about 4,400, babies were excluded when the copy of the birth certificate CDC received indicated that the certificate had been replaced (this was almost always done because of adoption). In addition, a list of all cases and controls (after the number of controls was reduced to 4,400) was prepared and arranged alphabetically by mothers' names. This list was reviewed manually to identify mothers who had two or more control babies, one of whom was randomly selected for the study. This same method was used to identify control babies born to mothers who also had case babies. In these instances, the case baby was included in the study and the control baby (babies) was (were) excluded. After the subsample was derived and exclusions made, 4,258 control families were "released" for tracing and interview. Later, we discovered that 12 of the 4,258 control index babies were registered by the MACDP as having Category 3 (nonstudy-eligible) defects. The 12 families of these babies were excluded, making the final number of study-eligible control group families 4,246 (Table 2).

2.4 LOCATING PARENTS OF INDEX BABIES

Because interviews were conducted by telephone and since the telephone was a very important locating tool, the processes of locating and interviewing were interrelated.

2.4.1 Locating Information

The process of locating parents of index babies so that they might be interviewed began with information sources created at the time the index babies were born—MACDP case records (Figure 1) and State of Georgia birth certificates (Figure 4). The fact that we had to use forms that had been created when the case and control group babies were born (1968-1980) caused a great deal of difficulty: the parents were located in 1981-1983, and the data on the forms were, except for names, generally out of date. In most instances, these data merely provided a starting point in the locating process. This process often consisted of a sort of detective work, a piecing together of clues gathered from a variety of sources, a task plagued by false starts, dead ends, and blind alleys. The process was formally systematized and simple, but, in reality, success all too often required insight, imagination, and intuition on the part of the tracing workers. Readers who are not interested in the details of the locating process can proceed to Section 2.5.

Great care was taken to insure that the information used for location was similar in type and quality for parents of case and control babies. All information about controls had to be derived from the live-birth certificates. Initially, we had planned to obtain most data for case babies from the MACDP case-record forms, because locating their birth certificates is time-consuming. The initial plans called for obtaining live-birth certificates only for those case group babies born in the years 1974 through 1980 because provision for recording mothers' Social Security numbers (SSN's) on the certificates began in 1974. After a certain amount of information had been abstracted (from the MACDP forms) for case group babies and (from birth certificates) for control group babies, we found that fathers' names were more commonly available on birth certificates than on the MACDP forms, particularly for the early study years. Therefore, we attempted to obtain copies of the birth certificates of all live-born case group babies, and those certificates were used as the source of certain data items, including fathers' names. Certificates of fetal death for stillborn case group babies (about 5% of all case babies) were not obtained because they contained no data of value for locating parents. When live-birth certificates could not be found for case babies (about 5%), the MACDP case-record forms constituted the sole information source. Certificates of live birth were the sole source of information about parents of control babies. Data abstracted from the birth certificates or MACDP case-record forms included:

1) Baby's birth date, mother's given name and surname, and mother's address at the time of the baby's birth. This information was available for all babies; for all case group babies it was taken from the case-history forms and for control group babies it came from birth certificates. The information on the case-history forms and the certificates was of the same quality, and the two sources can be considered interchangeable with respect to the accuracy of these data items.

2) Mother's maiden name. When available, taken from birth certificates for control group and live-born case group babies.

3) Father's given name and surname. When available, taken from birth certificates for control group babies, and for live-born case group babies when a certificate was available; taken from MACDP case-history forms for all stillborn case group babies, and for those live-born when a birth certificate could not be obtained.

4) Mother's and father's dates of birth (or age at time of birth, if dates not available). When available, taken from certificates for control group babies and live-born case group babies, from case-history forms for stillborn cases only.

5) Father's occupation or employer. When available, taken from birth certificates for control and live-born case group babies.

6) Mother's Social Security number (SSN). When available, taken from live-birth certificates for case and control group babies born 1974-1980. Paternal SSN's were not available.

For case group index babies only, the names of the physicians of record were abstracted from the MACDP forms so that they could be contacted before CDC contacted the babies' parents. We never used physicians as a source of parental locating information, because we wanted the quality of information for cases and controls to be equal. Many surviving case group babies require long-term care, and physicians' records would be expected to provide better information for their families than for families of control group babies.

2.4.2 Locating Procedures

Starting with the information described above, the sometimes lengthy and tedious location process was begun. Plans called for the stepwise use of a variety of sources of locating information. Parents were to be tried at one source at a time, in a "filter down" approach, until they were found. The system of sources through which parents were to be located was ordered on the basis of cost and effectiveness. We frequently found, however, that information obtained from a source suggested that the next source be one other than the one next in that order. In addition, during the study it became apparent that if we were going to collect data in a timely manner, we would have to use several sources simultaneously in locating parents. Overall, few parents were located by their names "filtering down" the preferred order of sources.

A computerized tracing management system was used to keep track of the location status of all study parents. The system included a "journal file" for each parent in which the tracing staff recorded all tracing information and activities. The tracing system was integrated with the telephone interviewing system (see section 2.5.1).

If a parent appeared to be "located" after some source had been used, that parent was "activated" for immediate contact by mail and, if a telephone number was available, for telephone contact about 2 weeks later. Parents were sent an introductory letter indicating the purpose of the study and requesting their help; the purpose of the study was explained in terms of learning about the causes of birth defects, and no specific mention was made of service in Vietnam or of Agent Orange. If a letter was returned as undeliverable, the parent's file would be returned to the tracing system for more tracing and, when a new address was obtained, the letter would be mailed again. If it appeared that a letter had reached the desired parent but telephone contact could not be made, more letters were sent, again requesting assistance. A postage-paid return form accompanied the letters. The form gave parents an opportunity to provide a telephone number at which they could be reached. If a telephone number was available, a contact was made regardless of whether the parent had returned the form. In addition, a telephone contact with parents often preceded the letter, since the telephone was an important tracing tool. In these instances, the parent was given a brief description of the study, an appointment for an interview was set up for a later date, and an introductory letter was sent.

The ultimate test that a parent of an index baby had been located was a telephone conversation during which the person responding verified that he or she was the parent of a baby born on the date of the index baby. If the location lead turned out to be erroneous, the parent's file was returned to the tracing system. When a location lead was found to be erroneous, the parent's file might be returned to the system at the next step or at an earlier step. In most instances, mothers and fathers were traced as a family unit. If the mother and father were not living in the same household, the interviewed parent might provide location information about the other. In addition, if no father's name appeared on a baby's case-history form or birth certificate, the mother, if and when contacted, was asked for information that might help locate the father. The various sources of locating information are outlined below, not necessarily in the preferred order of use.

1) Internal Revenue Service (IRS). Public Law 95-210 allows the IRS to disclose mailing addresses of taxpayers to the National Institute for Occupational Safety and Health (NIOSH) for the purpose of conducting occupational health studies. We arranged with NIOSH to use this capability for this special type of occupational health study. The user provides IRS with an SSN and the first four letters of the surname of a person whose location is desired and, if the SSN and partial surname make a match in an IRS file, IRS will return to the user the person's most recent mailing address. For those mothers whose babies were born in

the period 1974-1980 and for whom an SSN was available from the birth certificate, the IRS system was an excellent and very economical locating source. For a relatively small number of mothers and fathers, SSN's were obtained from other sources (see below), and in these instances the IRS was also useful. In February 1983, the system was made even more useful for the purposes of this study. After this date, if there was an SSN match, addresses were provided for "secondary filers" on joint tax returns, without a simultaneous match of the first four letters of the surname. This less stringent requirement for secondary tax filers was of particular value to this study, since CDC usually had only mothers' SSN's and since many of the mothers, because of separation or divorce, had changed their names since the birth of the index baby.

2) Southern Bell Telephone Company Directory Service. This was a quite useful and inexpensive source of locating information and was generally used as the first step for those parents whose SSN's were not available. The names of the parents and the addresses obtained from the case-history forms and certificates were checked against the current metropolitan Atlanta directories or with the company's directory assistance service. For those parents with common names who had moved since the time of the birth, this source was not particularly useful, but for those families with somewhat unusual names who had not moved from the Atlanta area since the time of the birth, this was a useful and economical source. As new leads became available, this source may have been used several times for the same parents. Southern Bell would not release unlisted telephone numbers, but could sometimes confirm that the sought parent had a telephone at a particular address. In addition, a set of Atlanta area telephone directories for the years 1967-1983 was available to the tracing staff (see 7, "Neighbor/Relative" Contacts).

3) Other Telephone Company Directory Services. If the IRS or some other locating source provided an address outside of the Atlanta area, the telephone company serving the area would be used in the same way that the Southern Bell's directory assistance service was used for the Atlanta area.

4) U.S. Postal Service. Letters describing the purpose of the study were sent to all study parents. These letters were accompanied by a prepaid return mailer on which a parent could provide a telephone number at which he or she could be contacted. The letters were sent at various stages of the location process—at any time the process turned up a new address for a particular parent. Even though the Postal Service can provide updated addresses for persons who have moved, this did not prove to be a particularly useful source of locating information—local post offices generally keep forwarding addresses only for a year or so.

5) Georgia Department of Public Safety. If the Department was provided with a name and a birth date (under certain circumstances, with just an age), it could often provide locating clues. This source was particularly useful for those parents who had moved from the metropolitan Atlanta area but had not left the State of Georgia. Occasionally, the Department could indicate to which State those who had left Georgia had moved, and this provided a new source of locating information. Motor vehicle departments of other States--particularly Florida, Tennessee, Texas, and California—also provided some useful information. Because some Georgia drivers' license numbers are their SSN's, the Georgia Department could occasionally give CDC SSN's, and this provided an opportunity to use the IRS search strategy (1, above) for some mothers and fathers.

6) Credit Bureaus. Parental identifying information was submitted to credit bureaus that would search their files for potentially useful locating information, such as residential addresses and parental employers. The bureaus were limited to an expenditure of $7.50 per

parent. Only locating information was requested and received; no personal or telephone contacts by bureau personnel were allowed.

7) Neighbor/Relative Contacts. Neighbor/relative contacts consisted of a series of operations involving telephone and "criss-cross" directories. The criss-cross directories (also known as "reverse" or "area" directories) generally provide an entry for each residence and place of business in the area and an entry for each resident. The entries are arranged by location (street and number) and by occupant's name. Thus, a person's place of residence in any particular year or the resident at a particular address during a given year can be determined.

A set of Atlanta city and suburban directories was assembled for the years 1967-1982 and proved to be a very good, albeit labor-intensive, source of locating information. Often, directories could be used to locate a parent's address when the telephone directory searches failed. They were most useful, however, when used for neighbor/relative tracing. This type of tracing involved contacting neighbors and potential relatives of the index babies' parents in an attempt to get leads about current location. The "relative" part of this procedure was only useful for those parents who had moderately unusual names. If a small number of persons with a particular surname were listed in the telephone or criss-cross directories, some or all of these persons would be contacted by telephone in the hope that some would be relatives of the parent. They would be asked if they knew the parent whose location was sought and, if so, how they might be contacted. The "neighbor" part of the procedure had the same goal—to contact a second party who might know the whereabouts of the parent being sought.

The most recent criss-cross directory was used to identify the current occupant of the residence at the most recent available address for the sought parent. If contact with the current occupant was not fruitful, then neighbors who were probably living in the area at the time the sought parent last lived there were identified. The tracing staff would begin with the directory for the year of birth of the index baby and confirm that the parent was listed at the address taken from the birth certificate or case-history form. The later directories would then be searched, with each change of residence noted and continuing from year to year until the parent was "lost" from the directory. This "loss" was taken to be an indication that the parent had probably left the area covered by the directories. The neighbors identified in the directory containing the last entry for the parent would then be compared with the neighbors listed in the 1982 directory. Names, addresses, and telephone numbers were abstracted for as many as five current neighbors who had lived in the area at the same time as the sought parent. Since these persons were among those who potentially knew the parents and could possibly know their current whereabouts, they were contacted and asked to provide clues as to the current location of the study parents.

Neighbor/relative tracing was done for almost all parents who had not been reached by the other, more economical sources. It was a crucial part of the overall tracing effort, and, despite a formal system of rules set down to govern its application, the creativity of individual tracers was essential for its success.

8) Marriage Licenses. All marriage licenses of persons who have been married in the State of Georgia since 1962 are recorded on microfiche and cross-referenced by maiden name of bride, name of groom, and year of marriage. It was not known when the parents of a particular index baby had been married, if they had been married in Georgia, or indeed if they had been married at all. Therefore, all years available on microfiche were searched—this involved less labor than trying to guess when the parents could have been married. Once a potential match was found, the certificate number was written down and a copy of the

marriage license application was obtained later. The application often contained the birthplaces of the bride and groom and their occupations and employers, and these items sometimes provided tracing leads. More importantly, the application would often also provide the names, birthplaces, and residences of the parents of the bride and groom (i.e., the grandparents of the index babies). The grandparents were often still living in the same location, and, when contacted, they could usually give information about the present location of the parents of the index babies.

9) Greater Atlanta Housing Authority. As one of the last sources to be tried, a rather large number of unlocated parents' names were matched manually against the Authority's tenant list. This source provided some useful information on a limited number of parents

10) Georgia Human Resources Department. The same list of parents given the Greater Atlanta Housing Authority was also given to the Georgia Human Resources Department, which checked its computer tape files. As with the Housing Authority, this source returned some useful information for a small number of parents.

11) Field Locating Work. In a number of instances the post office did not return introductory letters sent by "Certified Mail—Return Receipt Requested," indicating that the desired parent probably resided at the address, but the parent could not be reached by telephone. Presumably, the parent did not have a telephone or had an unlisted number (an estimated 5% to 7% of operating telephones in the Atlanta area have unlisted numbers (personal communication, Southern Bell Company, 1983)). We believed that the parent either did not want to participate or was simply not motivated enough by the letter to respond and indicate how he or she could be contacted. Near the end of the data collection period, it was becoming progressively more difficult to locate study parents. Therefore, we decided that field tracing should be used to try to reach parents such as those just described and others for whom there was a reasonable degree of confidence that the place of residence was known but who could not be contacted by mail or telephone. In this tracing effort, workers personally contacted the desired parents at their homes, explained the study goals to them, and encouraged them to contact the interviewing staff by telephone. This was a very laborious, but fruitful, procedure.

2.5 INTERVIEWING PARENTS OF INDEX BABIES

2.5.1 Computer Assisted Telephone Interviewing

The pilot study interviews were done with the traditional paper and pencil approach—that is, an interviewer read questions from a questionnaire to a responding parent and, with a pencil, recorded the parent's answers on the questionnaire. After the pilot study was completed, we decided that a more modern technique, Computer Assisted Telephone Interviewing (CATI) (Shanks, 1983), should be used to collect the main study data.

Each interviewer had a computer terminal that displayed each interview question in sequence and through which the interviewer recorded the parent's answer. As indicated earlier, the interviewing system was integrated with the computerized tracing management system. When an interviewer "signed on" the computer, she (all interviewers were women) would be provided with a parent's name and telephone number, which she would then call. At the time of the first telephone contact with a parent, the interviewer needed to verify that the parent had received a letter and that he or she was indeed the parent of the index baby. This verification consisted of the parent's positive response that he or she had a baby born in the Atlanta area on the date of birth for the index baby. If verification was obtained, the interviewer attempted to get the parent to proceed with the interview at that time, but if this was inconvenient, the interviewer made an appointment for a later time. If the number turned out to be erroneous, the interviewer would start a routine by which the CATI system would make an entry in the tracing management file (see above), and the parent would be tagged for more tracing work. The system would then provide the interviewer with a new parent and telephone number.

Some of the advantages and disadvantages of using a CATI system for this study are listed in Table 3. One major advantage was that data were checked for inconsistencies, logical errors, and other problems while the interview was being conducted; this improved the quality of the data and reduced the number of "call-backs" required. As a result of this "real-time" data checking, illegal codes for answers could not be entered. An example of this data checking involves the number of pregnancies reported by the parent. The parent was asked early in the first part of the interview to indicate the total number of pregnancies. The computer kept track of this number and automatically checked to see that the pregnancy-by-pregnancy history given by the parent matched. If there was an inconsistency, the CATI system would start a routine to help the interviewer determine the source of the error.

Another advantage of CATI was that the system automatically led the interviewer through the appropriate "skip patterns" (i.e., questionnaire branching in response to parents' answers). For example, if a responding father did not serve in the military, the questions about military service in Vietnam were not asked. With a paper questionnaire, the interviewer would need to read directions to determine the next appropriate question and flip the pages to get to the proper spot in the paper booklet—a procedure that can result in errors. With the CATI system, the computer automatically branched to the appropriate question and displayed it on the screen of the terminal. Parts of the questionnaire were personalized or modified specifically for individual respondents. These modifications ranged from relatively trivial (such as inserting the parent's name in appropriate places in questions) to very intricate.

Many of the interview questions focused on periods of time surrounding the date of the index birth. During the first part of the interview the parent was questioned as to whether the index baby had been delivered early or late and by how many weeks. From this information and the birth date, the computer calculated estimates of certain key dates, for example, the 6-month period from 3 months before conception to 3 months after. Where appropriate, the dates defining the period (e.g., January 1972 to June 1972) would be inserted in the questions.

The monitoring of interviewers was also substantially enhanced by CATI. In a normal telephone interview, only the interviewer's and respondent's voices can be monitored remotely; with CATI and a remote cathode ray tube, an interviewer's recording the respondent's answers can be monitored as they are given (during this study about 5% of interviews were monitored with both audio and video equipment).

Aside from the development time, the only major disadvantage of the CATI system used for this study was that if the central computer malfunctioned, all interviews in progress might need to be terminated. Fortunately, such incidents were rare.

Paper versions of the CATI questionnaires are provided in Appendix A. The CATI questionnaires cannot be exactly reproduced on paper, except by making a listing of the computer program, but the paper versions are good approximations and show all of the major items.

2.5.2 Interviewers and Interviewer Training

As noted above, all interviewers used for this study were women. Many of the questions in the interview were of a highly personal and sensitive nature, so, initially, we believed that it would be most desirable to have women conduct the mothers' interviews and men conduct the fathers' interviews. A second alternative was to use only women interviewers, since we believed that many mothers would not comfortably respond to male interviewers about some of the more sensitive questions but that fathers would probably not be inhibited by a female interviewer. Since very few male interviewers were available for employment by Westat, Inc., we decided that it was best to minimize the potential for a clouding effect caused by the sex of the interviewer, and the second alternative was followed (the tracing staff comprised males and females).

During this study, three groups of interviewers took the initial training session of about 40 hours' duration. The training covered basic interviewing techniques and specifics about this study's questionnaires and CATI system. Interviewers were retrained on an as-needed basis, generally in response to problems identified during the monitoring of interviews. In all, 69 interviewers worked on this study. At the peak, 29 were working regularly and 4 worked on the study from pilot phase to the time the study interviewing was terminated. These interviewers were all part-time employees of Westat, Inc., who worked 20-25 hours per week; they usually worked an aggregate of about 450 person-hours per week. Most interviewing on weekdays was done during the evening hours, but interviewing on Saturdays and Sundays was done during the daytime hours as well as in the evenings.

2.5.3 Interview Content

The basic structure of the questionnaire used for the interview was designed at the time the protocol was developed. Except for the deletion of some items, this basic structure did not change, although the wording and ordering of the questions was tested and revised several times before the main data collection began. Separate but similar questionnaires were used for mothers and fathers. Mothers were asked questions about themselves and about the fathers, and the fathers were asked about themselves and the mothers. Indeed, the fathers were asked many questions about the mothers for which it was expected that fathers would be relatively poor informants (and to a lesser extent, vice versa). This strategy was followed because we did not know how often the father of an index baby would be interviewed in the absence of a completed interview with the baby's mother, and we wanted to maximize the numbers of families with completed interviews with crucial maternal information. As this report later shows, this strategy turned out to be unnecessary because relatively few families yielded only a father's interview. The main categories of data collected during the interviews are shown in Table 4, and paper versions of the CATI questionnaires are found in Appendix A.

Much of the information collected was for the purpose of enhancing the credibility of the study results. The vast majority of birth defects have no known cause, and the number of known strong risk predictors is very small. We believed, however, that the results of the study would have greater impact if the data collected made it possible to assess and, if necessary, remove the confounding effects of suspected predictors of birth defects. In addition, the contact with the parents of these babies provided a unique opportunity to gather data on many potential causes of birth defects, some of which were considered to be very unlikely confounders of the birth defects-Vietnam veteran association. The cost of gathering this extra information was relatively small.

The interview for each parent was divided into two parts (Table 4; Appendix A). The first part consisted of a pregnancy-by-pregnancy reproductive history, and the second part contained all other information of interest. The time required to complete the first part of both mothers' and fathers' interviews depended on the number of pregnancies on which histories were obtained, but the average was about 10 minutes; the second part of the interview averaged about 45 minutes for mothers and about 40 minutes for fathers. Since the first part of a mother's or father's interview dealt with the outcomes of all of the parent's pregnancies, the interviewer doing this part of the interview could not be kept "blind" as to the index baby's case/control status. Since the second part of the interview did not deal with the outcome of the index pregnancy, but rather with the circumstances of the parents before and during gestation, the interviewer conducting the second part could be kept "blind." This was possible if the parent cooperated with the interviewer's request that the parent not mention whether the index baby had or did not have a birth defect. In general, this procedure worked well and the interviewer doing the second part usually did not know at the time the interview was completed whether the index baby was from the case or control group. Occasionally, this technique would fail for parents of case group index babies who would, because of the difficult emotional situation often associated with the birth of a child with a defect, reveal the baby's status. For interviewers to be kept "blind" as to a family's case/control status, four different interviewers were used for each family—two for the mother and two for the father. Wherever possible, the mother from a particular family was interviewed before the father.

Interviewers were "blinded" to minimize interviewer bias in data collection. The extent to which telephone interviewers' bias might influence the results of a study such as this is unknown. Obviously, most parents would know that the baby had or had not been born with a defect (some did not, see section 3.1.7), and the potential for parental response bias about exposure associated with case/control status influencing study results probably far exceeds the potential for interviewer bias influencing results. Nevertheless, the two-interviewer-per-parent technique was fairly easy to implement and, as far as we can determine, did no harm. In retrospect, the potential but unknown gain in the accuracy of the study results suggests that the procedure was worthwhile.

2.5.4 Premature Termination Interviews

To maximize the numbers of parents providing information about the father's military history, we included a special "premature termination" interview routine as a part of the CATI system. If a parent indicated that he or she wanted to end the interview before the military history section was completed, the interviewer transferred the CATI system to the military history section of the interview. This same procedure was used for parents who, on learning how much time a full interview would take, declined to participate; they were given the opportunity to complete the much-abbreviated "premature termination" interview that consisted of the military history section only.

2.6 GRANDPARENT SURVEY

The study protocol called for a postal survey of the grandparents of the index babies. The purpose of this survey was to supplement the information on the history of defects in the families of the parents of the index babies. This supplementation was desired because parents of the index babies might not know that their own siblings had birth defects if they had been stillborn or had died in infancy; the index babies' parents' knowledge of other relatives, such as their own uncles and aunts, was expected to be even less complete and accurate. The postal survey was attempted for the grandparents of the pilot study index babies, but the proportion returning the questionnaire was disappointing (about 45%). A test was made to determine if a telephone survey would increase participation to more acceptable levels, and the results were encouraging. This approach was, however, expensive, and since the results that would be obtained were only marginally relevant to the main purpose of the study—assessing Vietnam veterans' risks for fathering babies with birth defects—the survey was not completed. (The occurrence of defects in such distant relatives of the index babies was thought very unlikely to confound the birth defects-Vietnam veteran association.)

2.7 AGENT ORANGE EXPOSURE OPPORTUNITY INDEX

Agent Orange exposure opportunities for Vietnam veteran fathers were estimated by the military records specialists of the Army Agent Orange Task Force (AAOTF) without knowledge of case/control status. The "Exposure Opportunity" Index has five ordered categories. These categories represent intervals on a scale of the opportunities that veterans could have had for exposure to Agent Orange, opportunities ranging from nil and minimal to fairly numerous. A panel of specialists familiar with existing records of herbicide use in Vietnam used a mixture of objective and subjective methods to estimate exposure opportunities for individual veterans. The estimates were based on consideration of a veteran's military duties, period of service, and location in Vietnam. Vietnam veterans were given two index scores—one on the basis of information in their military personnel and organizational records and another on the basis of information about period and location of service and duties obtained during their interviews (information obtained during interviews did not always agree with information in the records). The AAOTF needed to have Social Security numbers or military service numbers of veterans in order to retrieve their military records, and mothers were not asked to provide the fathers' numbers. In addition, mothers were not expected to be particularly knowledgeable about where the fathers had been in Vietnam or what they had done there, so interview data from mothers were not used for the scores. Therefore, exposure opportunities were estimated only for veterans who were interviewed and who agreed to provide their numbers.

For the records-based index, information about occupational specialty was taken from veterans' personnel files. Some occupational and service groups were thought to have had more opportunities for exposures than others. For example, ground troops and personnel associated with spray aircraft and helicopters were much more likely to have had opportunities for exposure than, say, shipboard personnel. In addition, veterans of the same branch of service with similar occupational specialties might have had very different opportunities for exposure, depending on when and where they served in Vietnam.

For the records-based index, the locations where veterans had served were inferred from the locations of their military units, as recorded in the Operations Reports-Lessons Learned (ORLL) reports. These reports were prepared quarterly from reports submitted by all battalion-sized units to divisional headquarters for review and condensation. Information about day-to-day changes in a unit's location was generally not contained in the ORLL's, but they did provide an available and workable source of data. The unit daily journals, records that contain day-by-day information about unit locations, are available, but they could not be used for this study, since the veterans who participated were drawn from a very large number of different units. Reviewing the daily journals for all of these units would have required personnel resources far beyond those available to the AAOTF. The time of service information for this index was based on data in veterans' personnel files. Information about the locations where a veteran served in Vietnam and his period of service was crucial to the process of constructing index scores: herbicides were not randomly distributed in time and space in Vietnam.

The records of the Ranch Hand project herbicide applications are contained in the so-called "Herbs" computer tapes and are organized by date of mission, type and amount of herbicide used, and place of application. The proximity of these applications to the places of service described above were considered in constructing the index scores. In addition to the Ranch Hand records, the AAOTF has collected data on other herbicide applications, such as base perimeter sprayings, and these were also used in arriving at scores for individual veterans.

For the interview-based index, information about occupational specialties, periods of service, and location of service was derived from statements made by Vietnam veterans during the interviews. Otherwise, the scoring process was the same as for the records-based index, except as noted in the next paragraph.

Scoring on the basis of information in military records was done over a 16-month period. The scoring based on information obtained during interviews was done in 3 months, after all records-based scoring was complete. Establishing the index criteria was a dynamic process that spanned the 16 months during which the records-based scoring was done, with changes being made as more experience was obtained in scoring individual veterans and as more information about non-Ranch Hand herbicide applications was found. Thus, more variability is probably inherent in the records-based index scores than would have been the case with stable criteria. However, the criteria were stable when the interview-based scores were assigned.

A few general rules for scoring can be stated, but most Vietnam veterans' scores required specific, detailed review and discussion of the particular circumstances. The general rules included: 1) a score of 1 to be assigned to any veteran who served in Vietnam before Agent Orange was used (August 1965, National Academy of Sciences, l974) or after it ceased to be used (December 1970), 2) a score of 1 for any personnel stationed offshore, 3) a score of 5 for personnel connected with Ranch Hand after Agent Orange began to be used. Examples of index scores assigned to veterans with selected duty, location, and time-of-service characteristics are presented in Table 5. These examples will not be particularly instructive to people without some knowledge of the war in Vietnam and, in particular, some familiarity with the use of herbicides there.

It must be emphasized that this index, as it applies to individual veterans, does not necessarily reflect true levels of Agent Orange exposures, and even with respect to opportunities for exposure, its accuracy is unknown. The score assigned to any particular individual is only the panel's considered opinion about the opportunities for exposure an individual may have had. On the basis of records that exist today, it is impossible to assess how well the index reflects true levels of Agent Orange exposure. When this study was designed, we did not think that it was practicable to make estimates of any sort regarding even the opportunities for exposure. Since, however, the veterans' scores were determined without knowledge of the case/control status of their babies, the index scores should be free of the potential bias that perhaps contaminates the veterans' self-reports of exposure, obtained during the interviews.

2.8 ANALYTICAL METHODS

2.8.1 Analytical Tools

In this study, the main analytical tool was logistic regression, but several other techniques were also used. For single 2 X 2 (and various R X C) tables a standard chi-square statistic was used. When the data were considered in stratified analyses, a Mantel-Haenszel statistic (Mantel and Haenszel, 1959) was sometimes employed, and the extended Mantel-Haenszel test for trends in proportions (Mantel, 1963) was also used.

As a nonrigorous background, couched in terms of a medical application, the logistic model assumes that the probability of developing disease is related to various exposures by the following:

$$\ln(P_x/Q_x) = \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j,$$

where P_x is the probability of occurrence of disease, given a vector of exposures $X=(x_1,x_2,\dots,x_j)$; $Q_x=1-P_x$ represents the probability of the nonoccurrence of disease, given the same set of exposures; and $B=(\beta_0,\beta_1,\dots,\beta_j)$ is a vector of unknown parameters. Given a population in which the values of the exposures and the disease status are known, the parameters of the logistic model may be derived by using iterative maximum likelihood techniques (Kleinbaum et al., 1982). In case-control studies, two methods are available for formulating the likelihood to be maximized. In the unconditional likelihood model, the likelihood is given as the product of the unconditional probabilities, with each case contributing a term "P_x" and each control contributing a term "Q_x". This results in the estimation of a term β_0, which is interpretable only in studies in which (1) the proportions of all cases and controls are known and (2) cases and controls are random samples of their respective populations (Mantel, 1973). Alternatively, the likelihood may be formulated by using a conditional probability argument (Breslow and Day, 1978). In the latter case, no estimate of β_0 is derived. In stratified analyses involving a large number of strata with relatively small numbers of cases and controls per stratum, Breslow and Day (1980, chap. 7) have demonstrated that unconditional techniques may produce biased estimates for the logistic parameters. Since the circumstance of a large number of strata with a small number of cases in each stratum very often pertains in this study, conditional logistic analyses were used. Logistic regression has other important advantages. The coefficient β_i associated with a particular exposure may be viewed as a measure of association adjusted for the additional variables in the model (Schlesselman, 1982). In the particular case of a dichotomous exposure which is coded 0 (no exposure) and 1 (exposure), the quantity exp (β_i) is interpretable as an odds ratio adjusted for the additional variables included in the model. Including product terms in the model (e.g., x_1x_2) allows one to test for the presence of multiplicative interaction between two or more exposures (effect modification). Additionally, the analysis can incorporate interactions between the exposure variable and any of the design variables (i.e., race, hospital, and period of birth) in order to estimate, for example, the differential effects for White and Other races.

Until recently the conditional logistic method was restricted to situations such as matched pair studies (one case and one control per stratum) or studies with, at most, a few cases and controls per stratum. Recent development has made the method feasible for larger numbers of observations per stratum (Gail et al., 1981).

The specific models and variable codings used in assessing various hypotheses are presented in Appendix B. This Appendix may be best used as a reference when the "Results" section of this report is being reviewed.

2.8.2 Data Reduction for Analysis

Data obtained from the interviews were stored immediately in computer files by the CATI system. After certain items that were too complicated to automatically code through the CATI system (e.g., parental occupations according to the U.S. Census Bureau code (U.S. Bureau of Census, 1981)) were coded manually, the files were converted into the Statistical Analysis System ("SAS"; SAS Institute, 1982) files. These SAS files contained most of the data collected during the interviews and had to be reduced and recoded for the analysis presented in this report. The data items that were considered of interest for this analysis are presented in Table 6 and were stored in special "analysis" SAS files; the interview-derived data in these files were merged with data from the MACDP case-history files, State of Georgia birth certificates, and the Agent Orange Exposure Opportunity Index data. Four items related to data preparation deserve special mention: (1) definition of a veteran, (2) definition of a Vietnam veteran, (3) criteria for classifying self-reported exposure to Agent Orange, and (4) groupings of types of birth defects for the purposes of analysis.

Definition of a Veteran

An affirmative response to the question about whether the father of the index baby had ever served in the U.S. military qualified a father to be classified as a veteran.

Definition of a Vietnam Veteran

To be classified as a Vietnam veteran, a father had to 1) be classified as a veteran (see above); and stated by himself or the mother to 2) have served in the military in Southeast Asia; 3) have served specifically in Vietnam; and 4) have had that service before the index baby was conceived. For those men who had multiple tours of duty in Vietnam, the earliest tour must have preceded the conception of the index baby. If tour dates were not given during the interview, any father who was said to have been in Vietnam while in the military and whose baby was conceived after March 28, 1973, was considered a Vietnam veteran. The Department of Defense has designated July 3, 1965, to March 28, 1973, as the period for which service in Vietnam could be recognized with the award of a Vietnam campaign medal

Self-Reported Exposure to Agent Orange

For a father to be classified as having been exposed to Agent Orange according to his own report, he first had to meet all the criteria for being classified as a Vietnam veteran (see above) or he must have stated that he had been in Vietnam in an extramilitary capacity and exposed to Agent Orange there (see questionnaire, Appendix A).

Grouping of Birth Defect Types for Analysis

For the purpose of analysis the birth defects were classified into 96 groups (Table 7). The use of conditional logistic regression and the large numbers of controls available make it possible to do some analyses for individual types of defects (i.e., separate ICD-8 codes), even for types that affected very small numbers of babies. Other groups were formed from combinations of the codes in those instances where several codes formed natural groups from an embryological and/or epidemiological perspective (e.g., anencephalus and spina bifida). Note that the defect types to be considered in the separate analyses are comprised of Category 1 defects only. Even though parents of babies affected by Category 3 defects were eligible for interview, they were eligible because their babies had a Category 1 or 2 defect. The types of Category 2 defects that resulted in some babies being selected were diverse and difficult to group, and therefore no separate analyses of these Category 2 types were done. These Category 2 and 3 defects were, however, considered in the analysis of multiple malformations and in the search for a "Vietnam veterans' syndrome" of defects, as described below.

A special group comprising all babies affected by "multiple" defects was formed. This group was assembled on the basis of a slightly modified version of a standard Metropolitan Atlanta Congenital Defects Program (MACDP) definition of a baby with multiple defects--a baby with two or more defects where at least two of the defects affect different parts of the body and which are thought not to have a simple primary defect-secondary defect relationship (an example of a primary-secondary relationship is dislocated hip secondary to spina bifida). The specific rules for classifying babies as having multiple malformations were based on the modified ICD-8 codes and are presented in Table 8. Each of the Category 1 defect types is listed in this table along with the ICD-8 codes for defects excluded from consideration in the process of describing a baby with two or more defects as having multiple defects. For example, a baby with anencephalus (7400) and another defect in the range 7410-7439 would be considered to have a "single" defect. However, a baby with anencephalus, another defect in the 7410-7439 range, and reduction deformity (7552-7554) would be considered to have multiple defects. The specification of the rubric "Multiple Congenital Anomalies, Unspecified" (7599) as a "single" defect needs to be explained. CDC dysmorphologists consulted on how to treat this category suggested that babies coded with this rubric could be treated as having either "multiple" or "single" defects. The rationale for the first alternative is clear from the title of the class. The rationale for the second stems from the feeling that many babies so classified would be considered to have a "single" defect in the sense described above, if a more complete diagnostic workup had been available. Since there will be disagreement about the choice made, that is, to consider babies with defects coded under this rubric as having a "single" defect, it should be noted that only three mothers and three fathers of these babies had fully completed interviews; thus, the choice affects the analyses only to a small extent.

A relatively simple search was made for a possible syndrome of defects unique or more frequent among the babies of Vietnam veterans. The search consisted of comparing the frequency of Vietnam veterans among fathers of case group index babies with specific pairs and triplets of defects with the corresponding frequency among control group fathers. An example of a triplet of defects would be anencephaly, cleft lip, and reduction deformity in the same baby.

Recoding of MACDP ICD-8 birth defect codes using a modification of the 9th Revision of the ICD (British Paediatric Association, 1979) was begun after this study was started. The process was not complete when this report was prepared, but we did use the available new coding to sharpen certain diagnostic groupings. For example, diagnoses of ventricular septal defect (VSD) were separated into two groups: all VSD's and all VSD's minus the "probables." For clubfoot, the metatarsus adductus diagnoses were separated from the other defects in the ICD-8 clubfoot code. The new code also made it possible to form an embryologically indicated group of heart defects termed "conus arteriosus defects." Finally, the new code was used to assemble a group of case babies affected by defect syndromes thought to be caused by fresh dominant mutations.

2.8.3 Analytical Matrix

The overall analytical design can be thought of as a matrix having three axes. One is a hypothesis axis. Another is an axis that describes the degree to which the analysis takes account of covariables—confounders and effect modifiers. This axis is called the "adjustment axis." (For the analyses presented in this report, effect modifiers are defined as covariables that modify the association between case/control status and the exposure variable under consideration, on a multiplicative scale.) The third axis is the one on which the various groupings of defects are described. The intersections of the various levels of the three axes describe an analytical matrix, and each cell of the matrix signifies a separate analysis. This is illustrated in Figure 5 and is described below. Some analyses defined by the matrix were not done, because they were considered inappropriate or contraindicated. Except where specifically noted, all available and suitable controls were used for each analysis. Note that the analytical matrix describes most, but not all, analyses. A few important issues were dealt with outside the context of the matrix; they will be described later.

Hypothesis Axis

Four major hypotheses were evaluated for each of the defect groups at each of the levels of covariable adjustment:

1) Veteran Status. The purpose of considering the risk for veterans of military service versus nonveterans was to search for a possible "healthy worker" (soldier) effect. To avoid the possibility of confounding a general veteran effect with an effect of service in Vietnam, we excluded Vietnam veterans from analyses at this position on the axis. The assessment of the effect of veteran status merely represents an evaluation of the possible need for data limitation; the hypothesis does not derive from the major thrust of this study. If the risk of veterans (excluding Vietnam veterans) was found to be different from that of nonveterans, then all remaining hypotheses (for a specific defect and level of auxiliary factor consideration) would be restricted to veterans. Limiting later stages of the analysis to veterans required that the difference in risk between veterans and nonveterans be at least 20% and statistically significant at the alpha = 0.05 level. This relatively stringent requirement was imposed because a male "healthy worker" effect on reproductive outcomes is, on the basis of scientific reports, neither known nor expected (though veterans are known to have better general health than nonveterans (Seltzer and Jablon, 1974)) and because an inappropriate limitation of the data for analysis would sharply reduce the sample size and therefore the power of the analysis. The sample size should not be reduced in the absence of relatively convincing evidence that the risk in veterans differs from that of nonveterans. On the other hand, if there truly were a different risk among veterans, a limitation could result in a gain in power or validity, or both.

2) Vietnam Veteran Status. When this study was approved and funded, the only possible estimates of herbicide exposure for individuals were thought to be those obtained by questioning Vietnam veteran fathers. The study was therefore designed primarily to test the hypothesis that Vietnam veterans have a higher risk than other men of fathering babies with birth defects. Many scientists familiar with the occurrence and origins of birth defects would probably consider this to be a test of a "biologically implausible" hypothesis. It should be viewed, however, as a most important issue for Vietnam veterans and as the major focus of the analysis of these data.

3) Agent Orange Exposure Opportunity Index. Concern at this level of the hypothesis axis is whether the levels of the opportunities for Agent Orange exposure, as measured by the five-category index described above, are associated with the risk of fathering babies with birth defects.

4) Vietnam Veterans' Self-Report of Agent Orange Exposure. During the interviews, Vietnam veterans were asked if they thought that they had been exposed to Agent Orange. Some military experts believe that it would be very difficult for a veteran to give valid answers to such questions—other chemicals were widely sprayed in Vietnam, notably malathion for mosquito control, and such spraying would have been noticed by men serving in Vietnam. Thus, the potential for response bias is substantial, with fathers of babies with defects possibly having a different level of recall than the fathers of babies without defects. Therefore, in contrast with the other three categories on this axis, fathers of babies with defects were not compared with fathers of normal babies. Instead, fathers of babies with a particular type of defect were compared with fathers of babies with all other types of defects. This approach rests on two assumptions: 1) level of recall is equal for fathers of babies with all types of defects, and 2) exposure does not increase the rate of most or all types of defects equally.

Adjustment Axis

Three general levels of analysis were planned (Figure 5); these three levels were distinguished by the degree to which potentially confounding covariables were to be taken into account (in two of the levels various assessments for possible effect modification were also made). Three levels of analysis were done, since some observers will believe that the best analyses would involve a simple comparison of the frequency of the "exposure" variables in the case and control groups, and others will not be satisfied until a comprehensive search is made of the full range of covariables that conceivably (not to say probably) could influence our study results. The latter approach might, at first, seem to suffice for all, but it will not, since those who believe that the simple approach is best will hold that the comprehensive approach could reduce the possibilities for detecting any real associations that might be present in the data. The concern would arise from the notion that the comprehensive approach might introduce "noise" into the analyses, with a consequent increase in the statistical variance of the measures used to test the various hypotheses. The three levels of auxiliary variable (co-variable) consideration, which range from the simple to the comprehensive, are described below.

1) Basic Analysis. This level of analysis involved only case/control status, an "exposure" variable, and the three sampling design variables—time of birth (divided into three 52-month periods), race (White and Other), and hospital of birth. In other words, the hypotheses were tested without consideration of any covariables except the sampling design variables.

The final frequency matching of cases and controls was done by 12-month, not 52-month, periods. As noted before, the frequency matching on time of birth was done for several reasons. Foremost among them was concern that the tracing success, and hence study participation, would be strongly related to year of birth, with the families of babies born in the early study years expected to be more difficult to locate. In addition, year of birth was expected to be related to the likelihood that babies' fathers were Vietnam veterans, and the reported occurrence rates of several types of defects have changed over the years of birth included in the study. Moreover, parents' memories of important events before and during the index pregnancy were expected to be poorer among those whose babies had been born in the early years. Now that the study data collection has been completed and an excellent balance achieved on the year-by-year participation rates of cases and controls (see section 3.1.1 below), the preeminent reason for year-by-year matching has little relevance. We believe that the other considerations are well addressed at this point in the study by the grouping into three time periods. Furthermore, we considered it desirable to assess the possibility that the risks for Vietnam veterans varied by period of birth (see next paragraph). This would have been difficult to evaluate if stratification had been maintained on the 13 birth years. On the other hand, there is no logical way to group the hospitals of birth, and complete stratification on this variable was maintained.

The Basic analyses required that an interview be completed only to the point that a military history was obtained. Race and time of birth were considered simultaneously in a logistic regression model, including the first order interaction terms (with the "exposure" variable at issue), but hospital of birth was treated merely as a stratifying variable. If significant interactions ($p<0.05$) were found, further analyses were done with the data stratified on the interacting design variable.

2) Primary Adjusted Analysis. This level included those variables used in the Basic analysis and a small number of additional covariables designated as "essential" for a valid analysis. These essential covariables were chosen by a modified "nominal group process" (Delbecq et al., 1975). The group was composed of members of the staff of CDC's Birth Defects Branch, who are specialists in birth defects epidemiology. The charge of the group was to arrive at a set of potential confounders and effect modifiers that would have to be included so that a "reasonable" reviewer would feel that the analysis was valid. A reasonable reviewer might criticize analyses that did not take into account known, strong predictors of case/control status that are also related to an "exposure" variable. For example, if race was not a sampling design variable (and therefore automatically included in the analyses), it would have to be considered an "essential" covariable, since it is strongly predictive of several types of defects and was expected to be related to military service status. The modified nominal group process used for this decision making can be summarized as follows:

- Members of the group silently and separately wrote down their own ideas about what variables needed to be included in the analyses.
- A structured sharing of ideas then took place. Each member of the group, in round-robin fashion, presented one idea from his or her list, and a recorder wrote the idea on a flip

chart. The round-robin listing continued until all group members indicated that they had no further ideas to share.

— A discussion, covering all variables recorded on the flip chart, followed. If necessary the person who originated the idea for the variable or another member of the group clarified the idea. Members could voice support or nonsupport for the variable.

— Individuals then voted privately in writing on the importance of including particular variables. The final group decision was the pooled outcome of the individual votes.

The "essential" covariables defined by the group, for the analysis of all defects combined were socioeconomic status as measured by maternal education divided into three categories: less than 12 years, 12 to 15 years, and 16 and more years; maternal age in years; maternal alcohol intake, measured as estimated number of drinks per week; and birth defects in first-degree relatives (parents and full siblings) of the index baby.

The statistical treatment of the first three of the four "essential" auxiliary variables was straightforward, but treatment of the last, the occurrence of birth defects in first-degree relatives of the index baby, was complicated. Compared with the other "essential" covariables, this factor is unusual in that it can be considered both as a health outcome, which could be related to service in Vietnam, and as a confounding variable, which could cloud assessment of Vietnam service-associated risks. Considered as a health outcome, some experience related to service in Vietnam might confer a lasting risk that expresses itself by the birth of several affected children. When defects in first-degree relatives are considered as possible confounding variables, one entertains the possibility that Vietnam veterans (and/or their spouses) differ from other men (and/or their spouses) in terms of their inherent risk of parenting multiple babies with defects—that is, their risk could have differed before the service in Vietnam.

Statistical control for the possibility that Vietnam veterans had different risks before their service in Vietnam (i.e., a possibly confounding variable that could cloud the association between service and the occurrence of defects) was attempted by removing the following families from the analyses: 1) those in which the father or mother was said to have a defect and 2) those in which affected siblings were said to have been born before the index baby. Because a substantial number of families were removed when those with affected members were excluded, the Primary Adjusted analyses were done twice. Both sets of analyses were done with the other three "essential" covariables being considered, but one excluded the families with a history of birth defects and the other included them.

The degree to which Vietnam veterans were at a different risk of having multiple babies with birth defects after their tours in Vietnam was evaluated as follows: the frequency of birth defects in the siblings of case group index babies born after the fathers' periods of Vietnam service was compared with the frequency in the later-born siblings of case group index babies whose fathers did not serve in Vietnam. This analysis rested on the assumption that the preexisting risk for those compared was equal and that without an effect of Vietnam service the risks among later born siblings should be equal in the two groups. An effect of some Vietnam-associated risk should be manifest as a difference between the two groups. This evaluation took into consideration the numbers of offspring that the mother and father of an index baby had together.

The "nominal" group mentioned above also considered what variables should be considered "essential" for the analyses of specific types of defects. In most cases, the variables identified were the same as, or a subset of, those mentioned above. Many other characteristics were mentioned but rejected for consideration at this level of analysis (see Secondary Adjusted Analysis section below), because they occur so infrequently that they would be quite unlikely to cause any major distortion of the associations of interest. For example, treatment of maternal epilepsy with dilantin is thought to increase the risk of having babies with certain types of defects, including cleft palate (Hanson and Smith, 1975). Few pregnant women, however, use dilantin (about 0.3%; Niswander and Gordon, 1972). To influence the results of the analyses for cleft palate to any appreciable degree, mothers under this treatment would have to be greatly concentrated in one or the other of the "exposure" groups under consideration, which, a priori, seemed very unlikely. Therefore, the auxiliary variables to be considered for specific types of defects at this stage of analysis were limited to those variables chosen for all defects combined.

Except for the special treatment of the variable birth defects in first-degree relatives described above, the sampling design and "essential" covariables were considered simultaneously in logistic regressions, in the same fashion as for the sampling design variables in the Basic analysis. The Primary Adjusted analyses required fully completed interviews, since data on maternal education were among the last to be gathered during the interviews.

3) Secondary Adjusted Analysis. In this analytical stage the 108 auxiliary variables listed in Table 6 were considered for possible confounding of the estimated risks for the main exposure variables. These analyses were performed in a straightforward manner, one covariable at a time, without regard for possible effect modification. The basic design variable stratification on race, hospital, and period of birth was maintained, and an odds ratio estimate was calculated both with and without the covariate being considered. The ratio of the two odds ratios was examined to determine the magnitude of the change associated with the introduction of the particular covariate. Any covariate that changed a disease-"exposure" odds ratio by 50% (i.e., by a factor of 1.5 or its reciprocal, 0.667) was set aside for further consideration. Most covariates were coded as "yes" and "no," indicating 2 levels of stratification (beyond the 120 levels dictated by the stratification on the sampling design variables), but some were coded into 3, 4, or 5 levels (Table 6).

Logistic regression analyses were not feasible because of the very large number of covariables screened in this way and because of the massive amount of computer time that would have been required. Instead, the Mantel-Haenszel method was used. After all variables had been considered individually by the Mantel-Haenszel approach, we planned to assess those that met the criterion just mentioned together, along with the design variables, in a final analysis. This final analysis was to make use of the conditional logistic regression technique. As for the Primary Adjusted analyses, this level required fully completed interviews.

Defect Group Axis

The 96 birth defects groups described above and presented in Table 7 are arranged on this axis (Figure 5). Each case baby constituted an observation for each relevant defect category. In particular, a baby with two (or three, or four, etc.) defects was considered a case for each of the two-defect (or three- or four-defect) categories.

Data Bases

Depending on the position in the analytical matrix, one of three "data bases" was used for analysis (see Figure 5).

1) The "M" base includes data from all mothers' interviews. It is the largest of the three bases and was used preferentially for analytical phases in which detailed information about the father was not required or where the mothers' responses were acceptable substitutes for fathers' responses. For military service and assignment to Vietnam, mothers were expected to provide adequate reports. The M base was used for the Basic analysis of the two veteran status levels on the Hypothesis axis. It was also used for the Primary Adjusted analysis of the two veteran status levels, since the "essential" variables are attributes best obtained from the mothers' interviews.

2) The "F" base comprises information from all fathers' interviews. Its use was limited to instances in which information from the fathers' interviews was required and where detailed information about mothers was not. The F base was therefore used only for the Basic analysis of levels three and four on the Hypothesis axis.

3) The "MF" base includes data from those families in which both mother and father were interviewed. It was used in analyses where detailed information about both mothers and fathers was needed. Since availability of Agent Orange exposure estimations depended on fathers' interviews and the "essential" auxiliary variables are maternal variables, the MF data base was used in Primary Adjusted Agent Orange "exposure" analyses.

All Secondary Adjusted analyses used the MF base. The reasons for using the MF base for the Agent Orange "exposure" analyses are the same as those for using it for the Primary Adjusted analysis (above). Use of the MF base for the two veteran status levels was needed because many of the auxiliary variables to be considered are paternal attributes and are therefore best derived from the fathers' interviews.

Another data base could be formed, one comprising all families in which either a mother or a father, or both, completed an interview. This base was not used, since data items would have to be derived from mothers' interviews for some families and from fathers' interviews for other families.

2.8.4 Other Analytical Issues

Malaria and Malaria Chemoprophylaxis

All men questioned about their service in Vietnam were asked if they had contracted malaria and whether they had taken malaria prophylactic medicines while they were in Vietnam. The disease and the prophylaxis can be considered as exposures which could place a man at increased risk of fathering babies with birth defects. Since men who served on land in Vietnam generally should have taken the preventive medicines, however, any analysis of the prophylaxis should give results very similar to those found for the tests of the Vietnam veteran hypothesis. The analyses of these two "exposures" were relatively less comprehensive than the major analyses described above. The analytic tool used was the Mantel-Haenszel procedure, and the data arrangement most closely resembled that for the tests of hypotheses regarding self-reports of Agent Orange exposure. That is, fathers of babies with one type of defect were compared with the fathers of babies with all other types of defects. The motivation for this approach is the possibility of the same sort of response bias discussed above for the self-reports of Agent Orange exposure.

Female Vietnam Veterans

All mothers who completed interviews were asked if they had ever been in Vietnam, and if so, when and for what purpose. Not many mothers were expected to answer this question in the affirmative because of the relatively small numbers of American women who were in Vietnam during the war. But fitting the question into the flow of the questionnaire was a simple matter, and given the study's focus on reproductive health and the current interest in the health of women Vietnam veterans, the question seemed worth asking. A simple description of the women who said that they had been in Vietnam is the only data analysis that was planned.

3. RESULTS

3.1 DESCRIPTIVE STATISTICS

Before we present the analyses in which various hypotheses are tested (Section 3.2), we will describe several major features of the study data. These features include participation rates, distributions of the major variables connected with the various hypotheses to be tested, and the opinions of interviewed parents regarding the health of the index babies. A summary description is given below. Those who do not want to read the full description can read the summary and then proceed to the section on the tests of hypotheses (Section 3.2).

3.1.1 Summary

Overall, 69.9% of eligible mothers and 56.3% of eligible fathers completed interviews and an additional 1% or so of mothers and fathers completed interviews to the point that a paternal military history was obtained (Table 10*). The participation rate for parents of the White race was substantially higher than that for parents of Other races, particularly for fathers (Table 10). No marked differences in the case/control participation rates were noted for Whites, but for Other races the control group participation rates were about 5% higher than those for the case group (Table 10). This difference was shown to be of little concern insofar as it might affect the inferences to be drawn regarding the risks of Vietnam veterans (Section 3.1.4). With stratification on race, there was equal participation of case and control group families with respect to the other sampling design variables, year of birth and hospital of birth. In addition, there was relatively little variation in the participation rates for case group families associated with the type of defects that affected the index babies.

Roughly 50% of fathers of White index babies were said by mothers to have served in the military in contrast to about 30% of fathers of "Other races" index babies (Table 20). About 9%-10% of White index babies' fathers were said by mothers to be Vietnam veterans as compared with 6%-7% of Other race fathers (Table 29); again, the case and control group rates differed very little. About 2% of White race fathers could not be classified as to whether they were Vietnam veterans as compared with 6%-7% of Other race babies' fathers. Slightly over 10% of interviewed White race fathers said that they were Vietnam veterans and the percentage for interviewed Other race fathers was about the same (Table 29). In families where both the mother and father were interviewed, they almost always provided the same answer to the question of the father's veteran and Vietnam veteran status (Tables 25 and 31). The fact that the proportions of Vietnam veteran fathers among White and Other race fathers are nearly equal may be surprising to some readers—it is widely believed that a larger proportion of Blacks, as compared with Whites, served in Vietnam, but this was not the case (Veterans Administration, 1980; see also section 3.1.4).

About 25% of interviewed Vietnam veterans said that they believed that they were exposed to Agent Orange, and roughly the same percentage said that they did not know whether they had been exposed (Table 35).

The distributions of Vietnam veterans on the two variants of the Agent Orange Exposure Opportunity Index are presented in Tables 38 and 39 (a score of 1 indicated minimal opportunities for exposure; 5, the highest opportunities; and scores of 2, 3, and 4 graduated intermediate opportunities). Among those classified on both variants, 52% received the same score (Table 40).

---

*Only a selection of all descriptive tables are mentioned in this section; the tables are numbered in the sequence in which they are mentioned in the remainder of section 3.1.