group, and $\psi_k$ is the contribution due to length of time in RVN, etc. The statistical approach here is via maximum likelihood.

Logistic models (Walker and Duncan, 1967) have been extensively studied at USAFSAM for application in cardiovascular disease analysis. These models, in the herbicide context would have the form

$$P = \left[1 + \exp(\alpha + \beta_1 A + \beta_2 T + \beta_3 R + \beta_4 E + \beta_5 AE + \ldots)\right]^{-1}$$

where

- $P$ = probability of death
- $A$ = age in years
- $T$ = length of time in RVN
- $R$ = indicator variable for race
- $E$ = exposure variable

and where $\alpha, \beta_i, i=1,2,\ldots$ are coefficients to be estimated from the data. Testing for a group difference can be accomplished by estimating $\beta_4$ and the interaction coefficients such as $\beta_5$. If all interaction coefficients involving the exposure variable $E$ are zero and $E$ is treated as a 0/1 variable, Cox (1958a, 1958b) has shown that the most powerful test for non-zero $\beta_4$, in the setting of matched pairs, is McNemar's test. This latter test makes full use of the paired design of the study. For McNemar's test, the data are cast into a $2 \times 2$ table as shown in Table 8. In this table, "a" is the number of pairs in which both members have died, "b" is the number of pairs in which only the RANCH HAND person has died, etc. Using McNemar's test, the test statistic

$$\chi^2 = \frac{|b - c|^2}{b + c}$$

is calculated and referred to the chi-square distribution with one degree of freedom. Cox (1966) and Meittinen (1969) provided extensions of McNemar's test for R controls per exposed (R-to-1 matching). Of course the above analyses will be accomplished considering all deaths, and deaths by specific cause.

As previously discussed, RANCH HAND personnel may be characterized as risk takers. This risk taking behavior may be associated with increased mortality from a variety of causes. On the other hand, herbicide exposure has caused neuropathy in the RANCH HAND personnel, one could anticipate that this disability would increase the probability of accidental death. Therefore,
accidental death rates among RANCH HAND participants will be corrected for risk taking. This can be accomplished by including assessment of risk taking behavior in the questionnaire, indepth interview, and psychological evaluation. Both control and RANCH HAND mortality could be corrected using these measures, with the resultant rates being less biased and, therefore, a better indicator of exposed versus control effect.

(b) Mortality analysis without covariates.

The first step in the statistical analysis of survival data is descriptive, i.e., the construction of summary measures which provide a basis for comparing different exposure groups without any allowance for the effects of possibly confounding variables (e.g., age) except perhaps for some limited stratification. Since one must expect many "losses to follow-up", only methods which take full cognizance of this complication will be considered. It should be pointed out that all the methods described below assume independence between censoring (e.g., loss to follow-up) and death or morbid event, although some techniques permit different patterns of censoring in different exposure groups.

The life table method can be adapted to obtain a step-function approximation to survival distributions in the presence of censoring (Chiang, 1968, Gross and Clark, 1975). The failure time distribution is the function $F_0(t)$ which provides the probability of death at or before time $t$ in the study. The Kaplan-Meier estimator of $F_0(t)$ is $\widehat{F}_0(t)$ where

$$\widehat{F}_0(t) = 1 - \prod_{i \in D(t)} [1 - 1/R(T_i)]$$

In this equation, $D(t)$ is the "death set" at time $t$, i.e., the set of all indices $i$ of individuals who were observed to fail before time $t$. $R(T_i)$ is the number of individuals who were at risk just before time $T_i$, the time of death (or morbid event) of the $i$th study individual in $D(t)$. A nonparametric approach to testing the equality of survival distributions in a matched
pair study has been developed by Wei (1980). His statistic is a generalization of the Gehan (1965a) statistic. A second test for homogeneity of survival distributions for discretized failure data is the test for marginal homogeneity in a KxK table due to Stuart (1955). Thirdly, the McCullough Model and test may be used on the KxK array to test for marginal homogeneity and stochastic ordering.

(c) Mortality analysis with covariates.

These methods allow adjustment of mortality rates or morbidity rates using covariates such as age, race, length of time in RVN, AFSC, risk taking score, etc. For the purposes of this discussion it will be assumed that the covariates are categorical, that there are only two such covariates and the covariates do not interact in affecting the hazard of death or morbidity. These assumptions can all be relaxed using available methods.

The hazard function $h_i(t)$ for the $i$th individual in the study is the function which provides the conditional probability of death or morbid event in the time interval $(t, t+dt)$ given his survival up to time $t$. The function $H_i(t)$ where

$$H_i(t) = \int_0^t h_i(\tau)d\tau$$

is called the cumulative hazard for the $i$th individual. It is readily shown that the failure time distribution $F_i^0(t)$ is given by:

$$F_i^0(t) = 1 - \exp(-H_i(t))$$

From this last equation it follows that $h_i$ and $F_i^0$ are transforms of each other, hence the dependence of $F_i^0$ on covariates may be modeled via $h_i$. This may be accomplished as follows. Let $X_i(t)$ and $Y_i(t)$ denote discrete valued stochastic processes pertaining to the $i$th individual and describing two covariates of interest (e.g., one may be an exposure variable and the other may be covariate such as age or crew position). A basic model for hazard is:

$$h_i(t) = \exp [\xi X_i(t) + \eta Y_i(t)]$$

where $\xi$ and $\eta$ are "log-relative risks". This model may be extended to allow for any number of possibly interacting factors. Inference about log-relative risks may be drawn using either an approach derived from D. R. Cox (1972) by E. Peritz and R. Ray (1978) or using an approach described by Frank (1977). Another model, termed the proportional hazards model, is given by
The proportional hazards model has been discussed, for the special case that $X_i(t)$ does not change with time, by Cox (1972). A test for the equality of survival distributions in a matched pair study which incorporates the proportional hazard model has been given by Breslow (1975). A test of fit for the proportional hazards model is given by Schoenfeld (1980).

E. Morbidity Study

(1) General Considerations

A vigorous attempt to determine the morbidity experience of all exposed subjects and their primary controls will be undertaken using questionnaires, indepth personal interviews, and physical examinations. A waiver will be requested from the U.S. Attorney General so that medical information collected during the conduct of this study may be exempted from subpoena into Federal Court. Total confidentiality of medical information will be granted to subjects who are not on active duty, and partial confidentiality will be given to active duty subjects with release of information to the DOD only in instances where there is a public safety or national security risk. The schedule and method of contact with the study subjects is depicted in the Appendix Table A-7.

(2) Questionnaire Methods

All living exposed subjects and their primary controls will be offered a comprehensive personal and family health questionnaire administered in the subject's home by a civilian contractor.

In addition to subject interviews, a face-to-face interview will be conducted with the current spouses of the subjects to obtain a more accurate and complete assessment of fertility and reproductive function. Reproductive information that will be collected includes but is not limited to the number of live births, the number of still births, the number of miscarriages, the number of children conceived, the number of abnormal offspring, and the total years of marriage. Previous spouses of divorced or remarried subjects will also be interviewed to obtain similar data. Interviews with the first order next-of-kin of deceased subjects will provide morbidity data on the subject prior to his death. Whenever subjects, their spouses or next-of-kin will not consent to participate in a face-to-face interview, attempts will be made to elicit the information by telephone.

The questionnaire is an important part of this study because non-compliance rates for the physical examination and its face-to-face medical interview are expected to be substantially greater than non-compliance with the initial questionnaire. The questionnaire serves a four-fold purpose: (1) to capture baseline personal and medical data on subjects who might be noncompliant for subsequent physical examinations, (2) to serve as a cross-reference
source for objective data obtained at the time of physical examination, (3) to 
obtain a targeted medical inventory, independent of the physical examination 
process, and (4) to obtain health perception data to serve as a foundation for 
the replacement strategy. As depicted in the Appendix, Figure A-2, only an 
estimated 40% of the RANCH HAND population will participate in the examina-
tion, while at least 65% will respond to the questionnaire. The information 
collected by questionnaire from these additional 309 individuals and their 
controls will provide valuable morbidity data which would otherwise be lost. 
The questionnaire (see Section XI) will emphasize identification data, RVN 
tour history, dermatologic conditions, neuropsychiatric conditions, fertility 
aberrations, genetic defects in offspring, sensory defects, and personality 
factors. A targeted medical inventory will be included in the questionnaire, 
and will inventory symptoms prior to, during, and after duty in RVN as well as 
those currently manifested. It will take approximately six months to complete 
all initial questionnaires on both groups. The questionnaire will be "field-
tested" by the contractor on former Air Force personnel with RVN experience. 
Specific questions on the questionnaire will be directed to verifiable 
information, wherever possible. Questionnaire development and refinement, 
including specific response verification procedures have been pursued through 
civilian contract. Questionnaire data will be cross-linked and integrated 
with medical record information and physical examination findings. Question-
naire data from individuals not completing all phases of the study will not be 
discarded, but will be incorporated within the entire data base where statisti-
cally appropriate. Each participant will be asked to sign release forms so 
that all civilian health records, including those of dependents, can be 
obtained and reviewed as necessary. Attempts will be made to obtain patholog-
ical reports and specimens following surgical procedures. Federal health 
records on all family members on file in the NPRC will be retrieved. For 
retired members, and separated members with VA privileges, all available VA 
medical records will be obtained. All retrieved medical records will be 
reviewed, scored, compared to questionnaire data for reliability, and then be 
entered into a repository system. Identified participants who are non-
responsive to questionnaire will be pursued to determine status, disinterest, 
moribund state or death, etc. These individuals will be cross-referenced in 
other federal record systems in an attempt to achieve total ascertainment. 
Death certificates and autopsy reports will be retrieved on all dead exposed 
and matched control subjects for the mortality analysis. Birth/death certifi-
cates will be sought for all offspring.

(3) Physical Examination

A voluntary comprehensive physical examination will be offered to 
all individuals in both the exposed and primary control groups within one year 
of questionnaire administration. The condition for entry into the examination 
phase of the study will be the completion of the baseline questionnaire. In 
the event that the primary control does not complete both the questionnaire 
and the physical examination, a replacement will be selected from the control 
set [See Figure 3 and Section F(3)]. Statistical testing will be conducted by 
a variety of techniques on both questionnaire and examination findings. At 
the time of physical examination, an extensive physical examination, medical
history, and review of symptoms will be conducted. A standardized protocol will be used to insure comparability of data. This will provide cross-reference data to the initial questionnaire and to medical record data, if retrievable. Specific response verification and bias indicator questions will be included in this interview as well.

(a) Examination Parameters

A comprehensive physical examination will be conducted on all willing participants. The examination will be structured as outlined below and in Section XII and will be performed at the earliest practical time following the completion of the questionnaire. The close sequencing of these study components will limit the development of major symptoms in the interval between the questionnaire and the examination. Examinations will be performed under contract at a single civilian medical center having dermatologic, neurologic and electromyogram/nerve conduction capabilities. Informed consent forms will be obtained for all procedures. Physicians and technicians will handle all participants without a knowledge of exposed or control status, and will conduct the examinations by standardized protocols to minimize variability. Medical students, interns, and residents will not be allowed to perform these examinations, and specialty trained neurologists and dermatologists will perform the appropriate portions of the examination. An onsite monitor will insure that the examination protocol is followed. All laboratory tests will be subject to rigid quality control. Laboratory and physical examination data will be measured on a continuous scale whenever possible in order to improve statistical power in the analysis.

Under special circumstances, additional testing will be accomplished. Karyotyping of the individual and his family members will be considered if clinical history or physical examination findings are suggestive of this need. Most well conducted studies have shown that, when present, chromosomal abnormalities due to TCDD are transient. If on detailed analysis of the baseline examination and questionnaire, reproductive areas are heavily affected, routine karyotyping may be included in the test battery for the followup phases of the study. TCDD analysis on blood and urine will be considered in the future provided that (1) strong cause and effect relationships can be ascribed to Herbicide Orange and (2) high resolution mass spectrometry technology achieves 10 femtogram sensitivity with high isomeric specificity. Serum, urine, and semen specimens will be obtained from all participants, aliquoted, and preserved at -70°C for possible analysis in the future. These serum and/or urine specimens will also be used for analysis of porphyrin metabolites if analytic techniques make this a feasible diagnostic procedure. Extensive immunologic function analyses will be conducted on a randomly selected group of subjects.

Physical examination and laboratory data will be placed in the member's coded master file for detailed cross-analysis to questionnaire data. Information identifiable to the subject will not be released without his consent in accordance with the Privacy Act. However, in accordance with Air Force regulations, active duty flying personnel and active duty air traffic controllers found to have conditions which are disqualifying for flying duty will be temporarily "grounded" pending resolution of the medical condition.
**Physical Examination Profile**

<table>
<thead>
<tr>
<th>General Physical Examination</th>
<th>Hemoglobin</th>
<th>CPK</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBS, 2 Hr Post Prandial</td>
<td>Hematocrit</td>
<td>ECG</td>
</tr>
<tr>
<td>Urinalysis</td>
<td>White Blood Cell Count</td>
<td>Chest X-Ray</td>
</tr>
<tr>
<td>BUN/Creatinine</td>
<td>and Differential</td>
<td>VDRL/FTA</td>
</tr>
<tr>
<td>Cholesterol/HDL</td>
<td>Platelet Count</td>
<td>Cortisol Differential</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>RBC Indices</td>
<td>Thyroid Profile (RIA)</td>
</tr>
<tr>
<td>Serum Protein</td>
<td>Sedimentation Rate</td>
<td>Pulmonary Function Studies</td>
</tr>
<tr>
<td>Electrophoresis</td>
<td>Prothrombin Time</td>
<td>Blood Alcohol</td>
</tr>
</tbody>
</table>

**Dermatologic Examination**
- Urine Porphyrins
- Urine Porphobilinogen
- Delta-aminolevulenic Acid

**Neuro-Psychiatric Examination**
- General Neurologic Examination
- Nerve Conduction
- Psychological Battery: Nerve Conduction
  - MMPI
  - WAIS
  - WRAT
  - Halstead-Reitan
  - Wechsler Memory Scale Subtests
  - Cornell Index

**Reproductive Examination**
- LH, FSH, Testosterone
- Semen Analysis

**Neoplastic/Hepatic Examination**
- SGOT
- Alkaline Phosphatase
- SGPT
- LDH (Isoenzymes if elevated)
- GGTP
- Hepatitis B Antigens/Antibodies
- Bilirubin, Total and Direct

**Additional Studies (Individuals with abnormal history or examination)**
- Karyotyping
- Immunelectrophoresis
- Hepatitis A Antigens/Antibodies
- Bilateral profile and full-face photographs
- Anti-Nuclear Antibody
- Skin Biopsy
- Quantitative Immunoglobulins
- Additional Consultations as Required
- Immunologic studies (conducted on a randomly selected group of subjects)
  - Enumeration of B and T Cells
  - B and T Cell Function
  - Enumeration of Monocytes
(4) Analysis of Questionnaire and Physical Examination Data

The Questionnaire and Physical Examination will produce data of three types: (1) dichotomous, (2) polytomous and (3) continuous.

Dichotomous (e.g., present/absent) rates will be evaluated using the tools described above for mortality analysis. For example, the questionnaire will provide data concerning the first occurrence of disease states by age, and standardized rates and relative risks may be calculated. The occurrence of such findings can be related to age, time spent in RVN, exposure, and other variables using logistic models followed by McNemar's test where appropriate. These tests will examine the presence or absence of group effect and allow assessment of the statistical significance on non-unity relative risks.

Polytomous findings will occur in both questionnaire and physical examination responses. As an example consider retinal findings categorized into four grades, and studied as a function of age and exposure group as represented in Table 9. In this table the \( x_{ijk} \)'s are counts of occurrence. In analyzing tables such as these, techniques as described by Bishop, Fienberg, and Holland (1975) will be used. Specifically, if \( m_{ijk} \) is the expected value of \( x_{ijk} \), general log-linear models of the form

\[
\ln m_{ijk} = u + u_1(i) + u_2(j) + u_3(k) + u_{12}(ij) + u_{13}(ik) + u_{23}(jk) + u_{123}(ijk)
\]

will be used, where \( u_1(i) \) is the effect of RANCH HAND membership alone on cell frequency, \( u_{12}(ij) \) is the effect of an interaction on RANCH HAND membership with retinal grade, etc. This model can work with dichotomous as well as polytomous data. Under appropriate conditions on expected values of entries in Table 9, the pairing in the study design can be used with the data being organized as shown in Table 10. In Table 10, \( N_{ij} \) is the number of pairs such that the exposed person has retinal grade \( i \), and the control person has retinal grade \( j \). Appropriate tests for this setting are indicated by Fleiss (1973) and McCullough (1978).

With regard to continuous variables, the intended method follows Carpenter (1977) who found substantial gains in analysis efficiency by matching cases, subsequently employing covariance analysis to remove non-controlled effects. The conditional logistic regression model for relative risk, Holford, White and Kelsey (1978), is also applicable and will be used.
(5) Analysis of Fertility/Reproduction Data. The herbicides under consideration in this study have been alleged to effect fertility and/or reproductive functioning. An attempt will be made to address these allegations by analyzing at least three primary variables: the total number of conceptions since exposure in RVN, the number of miscarriages in spouses since exposure in RVN and, the number of abnormal offspring since exposure in RVN. The interview with current and former spouses will provide much more accurate information on fertility and reproductive functioning than if similar data were obtained from the male subjects themselves. The study questionnaire will provide the numbers of miscarriages, abnormal offspring and of live births. The sum of the number of miscarriages, still births, and live births will provide an estimate of the total number of conceptions. If differing divorce rates are found in the RANCH HAND and control groups, this may render the average number of years of marriage and the distribution of the years of marriage different in the two groups. This will be investigated and adjusted.
for if need be, either by analyzing total number of conceptions divided by (or normalized by) the number of years of marriage, or by using a more detailed covariance analysis. Further, the ratio of the number of miscarriages to adjusted total conceptions will be calculated and compared, as will be the ratio of the number of abnormal births and adjusted total conceptions.

In summary, the following statistics relating to fertility will be calculated and analyzed at the very least:

\[
\text{TOTAL CONCEPTIONS} = \# \text{Live Births} + \# \text{Still Births} + \# \text{Miscarriages}
\]

\[
\text{NORMALIZED FERTILITY INDEX} = \frac{\text{TOTAL CONCEPTIONS}}{\text{YEARS OF MARRIAGE}}
\]

\[
\text{MISCARRIAGE FRACTION} = \frac{\# \text{MISCARRIAGES}}{\text{TOTAL CONCEPTIONS}}
\]

\[
\text{ABNORMALITY FRACTION} = \frac{\# \text{ABNORMAL OFFSPRING}}{\text{TOTAL CONCEPTIONS}}
\]

F. Follow-up Study

(1) Study Adaptations

Following complete data analysis of the initial mortality and morbidity studies, adaptive or restrictive health surveys will be developed and administered to all follow-up study subjects three, five, ten, fifteen and twenty years after the initial questionnaire. Similarly, a condensed physical examination profile that will achieve adequate sensitivity and specificity for prospective diagnosis will be developed. The adaptive physical examination will be offered to all follow-up participants, and will also be conducted in years three, five, ten, fifteen, and twenty (see Appendix, Table A-5). An interim examination in year three is essential in this study because the age group under study is approaching that portion of the mortality/illness incidence curve with the steepest slope. A lapse of five years between the first two examinations could easily miss significant development of disease in the intervening years. Ample precedent for interim examinations can be found in the Framingham cardiovascular disease study, and in the follow-up evaluation of West Point graduates being conducted by the Air Force.
(2) Entry Criteria

All exposed or control individuals completing the baseline questionnaire and physical examination will be entered into the follow-up; further continuation will depend upon the member's willingness/ability to participate in additional health surveys and condensed examinations. Specific study entry rules are detailed in Table A-6 and Figure A-3 of the Appendix.

(3) Loss to Study; Key Issues

Loss of participants over time adversely affects any epidemiologic study in two ways. As the sizes of the study groups decrease, statistical power also declines, and bias is injected into the study if losses are not randomly distributed in the study populations. It is reasonable to assume that in this study, losses will be non-random with greater non-compliance among individuals who perceive their health as "well," since there is less incentive for this group to continue participation. As shown in Figure 5, such a differential pattern of loss will alter the population, and skew the frequency distribution curve.

Most previous epidemiologic studies have approached the problem of declining statistical power by beginning the study with multiple controls per exposed subject, and passively allowing attrition to occur throughout the study period. However, this approach does not address the problem of bias. This study will take an active approach to both of these problems by using a replacement concept. As a control is lost to study, a replacement will be chosen from the original set of ten matched controls. The replacement will be selected from the control set, and will have a perception of health similar to that of the lost control (Figure 6). The replacement strategy will maintain statistical power and the integrity of the matched design despite loss to study in the control group, and will correct anticipated bias while minimizing the number of required physical examinations.

At the initiation of the follow-up study, loss of an exposed member will not be cause to cease surveillance of his primary matched control. In the event of a control loss (for reasons other than death), another control from the set will be brought to study (Figure 7), the comprehensive questionnaire will be administered, and a baseline physical examination performed.

If a control is noncompliant for one portion of the study and is replaced by another control, the noncompliant individual will be approached at the time of subsequent questionnaires and examinations, and encouraged to reenter the study. If he reenters, both he and the replacement will be included in the evaluation. Similarly, noncompliant exposed subjects will also be aggressively recruited for all subsequent study phases.
EFFECT OF NON-RANDOM LOSS TO STUDY IN THE
CONTROL POPULATION

- If control losses are ill, a spurious effect is attributed to herbicide exposure.
- If control losses are well, a true/valid health effect is diluted.
THE REPLACEMENT STRATEGY

EXPOSED

PRIMARY CONTROLS

LOSSES

REPLACEMENTS

MATCHED FOR HEALTH STATUS

(DEAD CONTROLS NOT REPLACED)
CONTROL REPLACEMENT FOR THE MORBIDITY AND FOLLOW UP STUDIES

EXPOSED

1000

CONTROL

1000

YEAR 0

YEAR 1

YEAR 2

YEAR 3

YEAR 4

YEAR 5

YEAR 6

● QUESTIONNAIRE DATA

○ RECONSTRUCTED DATA

* LOSS TO STUDY

☆ PHYSICAL EXAMINATION DATA
For exposed and control individuals who drop out of the study but subsequently re-enter, medical data for the intervening years will be reconstructed from questionnaire and interview responses. IN ALL CASES OF LOSS-TO-STUDY, INTENSIVE EFFORTS WILL BE MADE TO DETERMINE THE SPECIFIC REASONS FOR NON-COMPLIANCE, AND DATA FROM REPLACEMENT CONTROLS WILL BE REVIEWED TO ASSESS COMPARABILITY WITH THE LOST INDIVIDUALS. Medical record reviews of new entrants will continue throughout the follow-up period.

(4) Study Length

The follow-up study is initially planned for 20 consecutive years. Procedures, progress, and interim results of the study will be monitored by an independent scientific review group, responsible to the Office of Science and Technology Policy in the White House.

G. Determination of "Disease"

(1) Introduction

Since this study is dealing with an unknown clinical endpoint with unknown latency, determination of a disease state by statistical methodology is a prime scientific thrust of the investigation. From the literature, chloracne is the only generally accepted chronic disease associated with high exposure to dioxin. The questions of primary interest are: (1) Does a history of chloracne invariably lead to future disease? and (2) In the absence of chloracne, is there emergence of other attributable diseases? Under a broad concept of "spectrum of illness", either or both of these conditions are possible. The clarification of their respective contributions to the natural history of past or of subsequent "disease" is of significant interest.

(2) Discussion

Inferences about a disease state from this study can be derived from several logical approaches. These approaches can be grouped into two categories: (1) those dealing with symptoms which can be used to construct a symptom complex that may represent disease, and (2) those dealing with physical signs which in themselves represent disease. In the former, one can form a subset of individuals that have symptoms (e.g., infertility) and study them during the morbidity and follow-up studies. Focusing on the overall patterns of alleged symptoms and categorizing them into a symptom complex may identify those individuals with a disease syndrome, or those at higher risk of developing disease (e.g., genetic disorders, cancer). In the latter approach, data on abnormal physical signs (e.g., genetic defects in offspring) and laboratory results can be compared between exposed and non-exposed groups in an attempt to again establish the presence or absence of disease. By putting this array of data into a logical decision-making scheme, specific relative risks can be calculated in the follow-up study, and specific response patterns can be inferred as shown in Figure 8.
### Interpretation of Horizontal Comparisons

<table>
<thead>
<tr>
<th>OVERT EFFECT</th>
<th>SUBCLINICAL</th>
<th>OVER-REPORTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_R &gt; M_C$</td>
<td>$M_R = M_C$</td>
<td>$M_R = M_C$</td>
</tr>
<tr>
<td>$S_R &gt; S_C$</td>
<td>$S_R = S_C$</td>
<td>$S_R &gt; S_C$</td>
</tr>
<tr>
<td>$F_R &gt; F_C$</td>
<td>$F_R &gt; F_C$</td>
<td>$F_R = F_C$</td>
</tr>
<tr>
<td>$F_{RS} &gt; F_{CS}$</td>
<td>$F_{RS} &gt; F_{CS}$</td>
<td>$F_{RS} &lt; F_{CS}$</td>
</tr>
<tr>
<td>$F_{RS} &lt; F_{CS}$</td>
<td>$F_{RS} = F_{CS}$</td>
<td>$F_{RS} = F_{CS}$</td>
</tr>
</tbody>
</table>

Mortality/Symptom/Sign regression on exposure

<table>
<thead>
<tr>
<th>OVER-RAPPORTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO REGRESSION ON EXPOSURE</td>
</tr>
</tbody>
</table>

\[
F_R = F_{RS} \cdot S_R + F_{RS} \cdot (1 - S_R)
\]
Again referring to Figure 8, at least three clinical patterns can be defined. These patterns are delineated using relative risks (mr/mc, sr/sc, fr/fc etc., between group or "vertical" comparisons, referencing Figure 4) and using within group ("horizontal" study) comparisons such as regressing symptoms and findings rates against an index of herbicide exposure, and other comparisons. Specifically, an overt clinical effect would be marked by: an increased mortality rate in the RANCH HAND group (mr > mc), an increased rate of symptom formation in the RANCH HAND group (sr > sc), and an increased rate of objective medical findings in the RANCH HAND group as compared to the control group (fr > fc). Further, the occurrence of physical or objective medical findings would consistently relate to symptoms in the overt case (that is, frs > fcs and fr$ > fcs$), and finally, in the classic instance, mortality, and symptom and sign formation would be seen to be increased with increasing herbicide exposure.

A subclinical pattern is indicated in the central column of Figure 8. In this setting, one expects no statistically significant differences in mortality or symptom reporting between the two groups, exposed versus control. However, one expects a consistent predominance of medical signs in the RANCH HAND group with regression of the signs on increasing herbicide exposure.

A pattern strongly suggesting over-reporting is presented as the right column of Figure 8. In this setting, there is no difference between the groups as regards mortality or medical sign incidence; however, more symptoms are reported by the RANCH HAND group. While in this pattern the RANCH HAND subjects are reporting more symptoms, objective medical finding rates are not consistent with symptom reporting. When no regression of symptoms on exposure level is found, over-reporting is clearly and strongly suggested.

This discussion of response patterns has used regression on an exposure index in a central way. Development of such an index is discussed below. It is noted, however, that a direct index of exposure can be confounded by other factors such as cellular repair mechanisms or bioaccumulation in adipose tissue with release over time upon weight loss. Use of other factors, such as time since exposure, should help to overcome these confounders.

The strength of any inferences made from these analyses is dependent upon the statistical power inherent in the study. In addition, due to the possibility of latency being a factor in this study, a negative analysis at any time within the study does not categorically imply lack of disease, since sufficient time for emergence may not have passed.

H. Exposure Indices

(1) Exposure Concepts

A major concern in conducting this study is the lack of accurate exposure data. Although most personnel assigned to RANCH HAND squadrons were undoubtedly exposed to Herbicide Orange and TCDD, the exposures within the
group must have varied widely. Exposure to herbicides and TCDD by RANCH HAND personnel occurred almost daily. Anecdotal information suggests that many had direct skin contact which was repetitive over a long period of time (one-year tour for most individuals). Further, it is also suggested that most RANCH HAND personnel felt that the herbicides employed in the operations were not toxic to animals and man, and hence, they did not exercise the caution in handling these chemicals that is recommended today.

From a historical review of RANCH HAND operations, it appears most individuals can be classified into one of three groups based on their likely potential for exposure to the herbicides:

1. Pilots, Co-pilots and Navigators: low potential
2. Crew Chiefs, Aircraft Mechanic, and other Support Personnel: moderate potential
3. Console Operators and Flight Engineers: high potential

The "pilot" group received most of their exposure during pre-flight checks as well as during the actual dissemination missions. The crew chief group experienced contact with herbicides during de-drumming and aircraft loading operations, as well as during on-site repair of the aircraft and spray equipment. The console operator group was exposed while supervising the loading of the aircraft, during ground testing of equipment, and by tank leakage during dissemination missions.

The available historical records on Operation RANCH HAND indicate that personnel assigned to the project seldom had a "routine" work schedule or environment, thus complicating estimates of the level of herbicide and dioxin exposure. Since actual exposure data (e.g., mg of herbicide/kg body wt) are not available, an exposure index will be used. The exposure indices will be calculated for each RANCH HAND individual to obtain frequency distribution, and will be calculated by evaluating the known factors that would have influenced exposure. These will include such factors as:

1. Date of tour with RANCH HAND in Vietnam.
2. Number and lengths of tours in Vietnam with RANCH HAND.
3. Number of herbicide dissemination missions (as reflected by flying hours and air medals).
4. Herbicides employed (records are available that reflect the amount of each herbicide sprayed each month and year).
5. Crew position.
6. Routes of exposure (the major route of exposure for most RANCH HAND personnel was probably percutaneous, although exposure through inhalation may have also been significant).
A crude exposure index which is applicable to the entire RANCH HAND cohort is expressed with the following formula:

\[ E_i = q_i \times T_i \]

In this formula, \( E_i \) is the calculated exposure for the \( i \)th RANCH HAND member, \( q_i \) is the quantity of TCDD-containing herbicide sprayed from aircraft assigned to the \( i \)th subject's base during his assignment, and \( T_i \) is the length of the \( i \)th subject's assignment (tour length). However, great care must be exercised when applying the above index. For example, the index should be used as an independent regression variable against clinical findings only within occupational strata, to avoid confounding occupational effects with exposure effects. Different degrees of regression between clinical findings and the exposure index can be expected in differing occupational groups since: (a) modes of exposure are likely to be different in different occupational categories, (b) socioeconomic correlates within occupational category could confound an herbicide effect, and (c) other exposures which could synergistically or antagonistically interact with TCDD-containing herbicide may be correlated with occupational category.

Another factor which must be considered when applying this crude exposure index is the problem on confounding a possible herbicide effect with an effect associated with tour length. Being in a combat zone is a major psychophysical stress, and time spent in such an area may be significantly associated with changes in long term morbidity and/or mortality. This crude exposure index, when used alone, could result in a positive regression with disease incidence or prevalence which is not due to the herbicide exposure. An approach that will correct for this potential confounding is to regress observed medical findings on both \( E_i \) and \( T_i \) to differentiate the independent effects of herbicide exposure and combat zone experience.

The values of \( q_i \) and \( T_i \) needed to calculate \( E_i \) are generally available from government records. Specifically, tour dates are available from military personnel records, and the quantity of herbicide sprayed is available for the period January 1965 through April 1970 from the "HERBS TAPES." These tapes are comprised of computerized data obtained from actual spray mission reports. This material provides the date, base of mission origin, amount and type of material sprayed (Herbicides Orange, Blue, or White) and location of the intended spray target. Estimates of the amount of herbicide sprayed prior to 1965 may be available from procurement records for Herbicides Purple, Pink, and Green, which were sprayed exclusively from Tan San Nhut Air Base from 1962 through 1964.

Animal data imply that TCDD is the most toxic component in the herbicides used in RVN. By using \( q_i \), the amount of herbicide sprayed, one is using a variable that roughly correlates with TCDD exposure. However, it would be highly desirable to be able to analyze observed health effects in terms of specific TCDD exposure. The material sprayed from 1965-1970 had significantly lower
TCDD contamination then did those herbicides manufactured and purchased prior to 1962 and used from 1962 through 1964, but due to data limitations from a scarcity of Herbicide Purple, Pink, and Green samples, TCDD concentration profiles for those chemicals cannot be quantitatively determined. However, it may be feasible to develop estimates of the degree of contamination based upon the TCDD concentration from military and manufacturers' data.

As another approach to examining the effect of TCDD itself, one might consider stratifying the exposed cohort by date of assignment in Vietnam, expecting that those assigned earlier were more heavily exposed to TCDD. While it may well be true that earlier assignees were exposed to higher TCDD concentrations, it is unlikely that differences between "early" and "late" assignees, if they occur, can be reliably attributed to TCDD concentration changes, since several potentially confounding variables exist: (a) volunteerism among early assignees, (b) differing assignment patterns between early and late RANCH HANDers (TDY vs long term pattern) and (c) different RVN living conditions.

It is preferable to use an exposure index which is more closely tailored to the specific individual than the crude index discussed above. While \( T_i \) is subject specific, \( q_i \) is a value which refers to all individuals on the base during the period of time represented by \( T_i \). A refined index for ground crew can be expressed as:

\[
E_i = F_i \times q_i \times C \times T_i
\]

where,

- \( F_i \) = Average flights per day served by the \( i^{th} \) ground crew member.
- \( q_i \) = Average quantity of herbicide dispensed by flights served by the \( i^{th} \) ground crew member.
- \( C \) = Estimated TCDD concentration of the herbicides in use during the \( i^{th} \) subject's tour of duty.
- \( T_i \) = Time spent in TVN in days for the \( i^{th} \) ground crew member.

The variable \( F_i \) can be estimated by dividing the number of RANCH HAND flights per day by the number of crew chiefs during the time period \( T_i \). All other variables are estimated as with the crude index.

A refined index is also possible for aircrew members and is expressed as follows:

\[
E_i = M_i \times D_i \times q_i \times C \times P_i
\]

where,

- \( M_i \) = Total number of missions flown by the \( i^{th} \) air crew member.
- \( D_i \) = Average duration of missions flown by the \( i^{th} \) air crew member.
- \( q_i \) = Average quantity of herbicide dispensed per flight served by the \( i^{th} \) air crew member.
- \( C \) = Estimated TCDD concentration of the herbicides in use.
- \( P_i \) = A crew position weighting factor.
As with the refined ground crew index, this refined aircrew index cannot be directly calculated in a strictly quantitative sense using available government records, since records to specifically link missions with particular individuals are not available to objectively determine \( M_i \) and \( D_i \). However, reasonably accurate estimates of these parameters may be feasible using questionnaire data. Also air medal awards may allow an indirect estimate of \( M_i \).

The crew position parameter \( P_i \) must also rely upon estimations. While the specific crew duties of each subject are known, the differential exposures associated with the crew positions within the C-123 aircraft were not determined during RVN spray missions. The 355th TAS/Spray Branch, Rickenbacher AFB OH is presently using the C-123 aircraft, configured with the A/A 45 Y-1 Internal Dispenser and attempts to assess \( P_i \) can be made. Air flow measurement and herbicide simulant deposition studies conducted by Meek are performed during the course of four C-123 flights. However, difficulties with the measurement equipment limit the validity of the value of the data in an exposure index. Further work along these lines could yield a more quantitative position weighting factor, \( P_i \), for each individual.

Refined ground crew and air crew exposure indices can be used singly or in combination with the crude exposure index first presented; however, as with the crude index, confounding must be avoided when the refined indices are used in statistical analyses.

The exposure indices listed above are, of course, only applicable to the Ranch Hand cohort. As mentioned, a positive regression of disease incidence or prevalence with increasing exposure index will strongly support herbicide causation. We do not wish to minimize however the role of RANCH HAND versus control group disease incidence/prevalence differences as indicators of a herbicide effect. A major component differentiating the RANCH HANDers from the controls is the increased residence of RANCH HANDers in the RVN itself. If within country time does not correlate with disease incidence, RANCH HAND versus control disease incidence differences may be strongly related to herbicide. If in-country time is significant as a disease correlate, this in itself will be valuable information with regard to assessment of the RVN experience.
VI. Special Statistical Considerations

The previous discussion has outlined the general statistical approach followed by this protocol, and has outlined planned analytical methods and inferential strategies for the mortality, questionnaire and physical examination study phases. This section provides an indepth consideration of some special statistical study aspects.

A. False Reporting/Misrepresentation

Since concern for compensation could unconsciously or consciously influence symptom reporting, and since press reporting itself can stimulate anxiety-based symptom formation, a discussion of false reporting is indicated. A data pattern indicating overreporting has already been discussed in Section V. The goal here is to understand the effect of misrepresentation on estimates of relative risk and the odds ratio. Let S stand for presence of a symptom, and S' denote its absence. This false reporting may be represented as in Figure 9.

![Figure 9]

**FALSE REPORTING/MISREPRESENTATION**

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>S'</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>A</td>
<td>B</td>
<td>A+B</td>
</tr>
<tr>
<td>S'</td>
<td>C</td>
<td>D</td>
<td>C+D</td>
</tr>
<tr>
<td></td>
<td>A+C</td>
<td>B+D</td>
<td></td>
</tr>
</tbody>
</table>

The proportion of correctly classified positives is defined by A/(A+C) and is called the sensitivity of the classification scheme; the proportion of correctly classified negatives D/(B+D) is called the specificity.

When there is non-differential misrepresentation, that is, when the sensitivity and the specificity are the same among the exposed and nonexposed, the bias induced in the estimate of relative risk will be toward the null value. The situation is summarized by Figure 10.
Using this representation, the true relative risk is \( \frac{a+c}{n} \) + \( \frac{e+g}{n} \), and the apparent relative risk is \( \frac{a+b}{n} \) + \( \frac{e+f}{n} \). Figure 11 provides a graphic representation of how apparent relative risk varies as a function of specificity. For this curve, the true relative risk is 2 with the exposed population having a symptom incidence of 0.1 and the nonexposed population having a symptom incidence of 0.05 (Copeland et al. 1977). The effect of nondifferential false reporting on the odds ratio is nearly as severe as that shown in Figure 11 for relative risk. A technique does exist for correcting the estimate of relative risk to account for false reporting, but the technique requires knowledge of the sensitivity and specificity of the classification scheme; knowledge that may not exist in this study. It should be noted that since the above remarks are concerned with relative risk, the number \( n \) of subjects in each group is irrelevant, as the results shown are independent of \( n \).

If the false reporting is differential, an estimate of relative risk that is biased away from the null value can result. This will occur in situations in which the RANCH HAND personnel and controls do not misrepresent their symptoms in the same manner (Copeland et al. 1977). Thus the "true" outcomes of herbicide exposure may be distorted depending upon the degree and direction of misrepresentation.

B. Adequacy of Sample Sizes

(1) Overview

The size of the RANCH HAND cohort is approximately 1000 individuals. It is clear that a lethal effect of herbicide which occurs in only 1 out of 2000 controls will be quite difficult to detect unless the herbicide effect is very strong. For example, at a rate of 1 in 2000, 0.5 affected controls are expected. If the basic rate is doubled by herbicide to 2 per 2000, one affected RANCH HAND individual would be expected. At a rate of 1 per 2000 for
controls and a rate of 2 per 2000 for RANCH HAND personnel, the probability of observing no affected individuals in both groups is

\[(1 - 1/2000)^{1000} (1 - 2/2000)^{1000} = .22\]

or, in other words, "there is a 22% chance" that no affected individuals will be found in this study. In a population of 100,000 exposed individuals, 100 cases would be expected, 50 of which would be due to herbicide. In short, since the size of the RANCH HAND group is fixed, this study has limited statistical power to define the relationship of herbicide to the rarer diseases.

The power \((1-\beta)\) of a study design is the probability that a specified difference between populations will be detected if it in fact exists. In general, power is a direct function of sample size; that is, for a particular study design, the more subjects measured the larger the study power. It is understood that this protocol makes use of the entire known RANCH HAND population (and excludes ancillary exposed groups for reasons previously cited); the exposed sample size cannot be increased. Power augmentation, therefore, can only be accomplished by the less efficient procedure of increasing the control group size which has statistical limitations as well as staggering financial and logistic considerations. Hence, considerable effort has been made to correct loss to study issues (by replacement and other techniques to induce participation) and to use the most powerful statistical design concepts.
Essentially all previous animal and human studies concerning herbicide suffer from a lack of adequate consideration of study power. The following presents a preliminary analysis of study power for the case of continuous and dichotomous variables expected from the study. Also reviewed are alternative studies involving Marine samples.

(2) Power in Continuous Variable Case

Assume that blood cholesterol levels are being compared between RANCH HAND and control groups, and that the coefficient of variation for cholesterol in the control group is 0.1, where the coefficient of variation is the ratio $\sigma_c/\mu_c$. Assume $\sigma_{RH} = \sigma_c$. The symbol $\alpha$ is the probability that the study will indicate an effect where none exists, and $1-\beta$ is the power as defined before. Consider that the RANCH HAND mean cholesterol $\mu_{RH}$ is shifted from the control mean $\mu_c$. A natural question is to inquire about the study power as a function of available pairs ($n$) and mean ratio $\gamma = \mu_{RH}/\mu_c$.

<table>
<thead>
<tr>
<th>$r$</th>
<th>$\gamma$</th>
<th>n=180</th>
<th>n=450</th>
</tr>
</thead>
<tbody>
<tr>
<td>.20</td>
<td>1.01</td>
<td>.20</td>
<td>.38</td>
</tr>
<tr>
<td>.20</td>
<td>1.02</td>
<td>.55</td>
<td>.88</td>
</tr>
<tr>
<td>.20</td>
<td>1.05</td>
<td>&gt;.995</td>
<td>&gt;.995</td>
</tr>
<tr>
<td>.70</td>
<td>1.01</td>
<td>.86</td>
<td>.995</td>
</tr>
<tr>
<td>.70</td>
<td>1.02</td>
<td>&gt;.995</td>
<td>&gt;.995</td>
</tr>
<tr>
<td>.70</td>
<td>1.05</td>
<td>&gt;.995</td>
<td>&gt;.995</td>
</tr>
</tbody>
</table>

Power calculations are displayed in Table 11. Study power in the case of a matched pair design is strongly dependent on the degree of positive correlation produced between the involved groups by the matching procedure. Of course, the degree of correlation can be expressed by the correlation coefficient $r$ which can take values between -1 (negative correlation) and +1 (positive correlation), and two values of $r$ have been employed in Table 11. From this table it is seen that if only 450 pairs are studied a 1% shift in mean ($= 1.01$) will not be reliably detected, but a 2% shift will be detected with a probability of 0.88 if $r = 0.2$ at least. From this calculation one can infer the need to examine at least 450 pairs to obtain the 2% shift, and to strive for more if possible.
(3) Power in the Dichotomous Variable Case

There is significant discussion in the mathematical statistics literature concerning the efficacy of paired designs in the setting of dichotomous responses (Billewicz, 1974; Ury, 1975; Miettinen, 1970; and several others). Table 12 shows a set of calculations which are applicable to the present study.

Table 12

POWER CALCULATIONS FOR THE DICHOTOMOUS VARIABLE CASE AS A FUNCTION OF EFFICACY OF PAIRED DESIGNS

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>Rel. Risk</th>
<th>r</th>
<th>n=250</th>
<th>n=350</th>
<th>n=450</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>.01</td>
<td>5</td>
<td>0</td>
<td>.77</td>
<td>.82</td>
<td>.92</td>
</tr>
<tr>
<td>.04</td>
<td>.01</td>
<td>4</td>
<td>0</td>
<td>.61</td>
<td>.75</td>
<td>.85</td>
</tr>
<tr>
<td>.03</td>
<td>.01</td>
<td>3</td>
<td>0</td>
<td>.40</td>
<td>.51</td>
<td>.59</td>
</tr>
<tr>
<td>.10</td>
<td>.05</td>
<td>2</td>
<td>0</td>
<td>.61</td>
<td>.75</td>
<td>.85</td>
</tr>
<tr>
<td>.20</td>
<td>.10</td>
<td>2</td>
<td>0</td>
<td>.87</td>
<td>.94</td>
<td>.97</td>
</tr>
<tr>
<td>.05</td>
<td>.01</td>
<td>5</td>
<td>.1</td>
<td>.89/.029</td>
<td>.94/.032</td>
<td>.98/.064</td>
</tr>
<tr>
<td>.04</td>
<td>.01</td>
<td>4</td>
<td>.1</td>
<td>.72/.033</td>
<td>.87/.038</td>
<td>.88/.041</td>
</tr>
<tr>
<td>.03</td>
<td>.01</td>
<td>3</td>
<td>.1</td>
<td>.38/.020</td>
<td>.68/.046</td>
<td>.71/.077</td>
</tr>
<tr>
<td>.10</td>
<td>.05</td>
<td>2</td>
<td>.1</td>
<td>.76/.055</td>
<td>.85/.048</td>
<td>.88/.048</td>
</tr>
<tr>
<td>.20</td>
<td>.10</td>
<td>2</td>
<td>.1</td>
<td>.94/.043</td>
<td>.98/.046</td>
<td>.99/.057</td>
</tr>
</tbody>
</table>

*α = .050
**α = as indicated

In this table, r is again the correlation coefficient indicating the degree of correlation induced between the involved groups by the matching procedure. The probability of the disease among RANCH HAND personnel is symbolized as P1, while P2 is the probability of the disease among the controls. Relative risk is the ratio P1/P2. With r = 0.1, sign test power tables were used as an exact version of McNemar's test, and therefore different α levels are shown under each power number. Table 12 shows the positive influence of effective
pairing in the higher power levels noted. Also, it appears that for \( p_2 = 0.01 \) and \( p_1 = 0.03 \), physical examination of 450 pairs (900 examinations) will disclose the three-fold relative risk with probability less than the minimum target .80. In other words, there is a greater than "20% chance" that a three-fold relative risk on a 1/100 disease state will go undetected in this study if only 350 pairs are examined and if low correlations occur. Once again the need to examine the maximum numbers of pairs in the study is seen.

To present these dichotomous power calculations more clearly, calculations in the context of actual disease states have been accomplished. The diseases considered are cardiovascular disease and cancer, corresponding to high and low rate illnesses for the age groups presently under investigation.

(a) Cardiovascular Disease

A logistic risk function was fitted to data from 17,455 autopsies gathered in a WHO collaborative study in Czechoslovakia, Sweden and the USSR. The function fitted has the form

\[
P = \left[1 + \exp(a + \beta(x-.5) + \gamma(y-.5))\right]^{-1}
\]

where

- \( p \) = the probability of a complicated coronary lesion
- \( x \) = age scaled linearly so that \( x = 0 \) is equivalent to 30 years, and \( x = 1 \) is equivalent to 58 years (the age span of the current study)
- \( y = 1 \) or \( 0 \) if the subject is exposed or not

and \( \alpha \) and \( \beta \) were obtained from the data. The function represents a fairly high rate disease in that at 40 years of age 7% of the group had the lesion, and at 60 years of age 20% had the lesion. The coefficient \( \gamma \), represents the exposure effect. Power calculations for \( \gamma = \beta \) and \( \gamma = .8\beta \) are shown in Table 13. This table suggests that if, as a cell toxin, herbicide exposure accelerates cardiovascular disease, this study has a good chance of detecting that acceleration if the herbicide effect is comparable to the age effect. A slight beneficial effect of pairing is seen in this hypothetical example.

(b) Cancer

A logistic risk function was fitted to breast cancer data presented by Breslow and Day (1975). The function fitted represents a low rate disease in that at 35 years of age only .000336 of the group had the lesion while at 70 years of age .00676 of the group will have the lesion.
Using pairing to achieve a power of 0.80 in this setting, 1312 pairs would be needed, when the exposure effect is equal to the age effect. This exceeds the size of our RANCH HAND cohort, and reinforces the fact that herbicide exposure effects on rarer diseases will not have a high likelihood of being detected by this study, and again supports an attempt to examine as many pairs as possible.

Table 13

POWER CALCULATIONS AS A FUNCTION OF HERBICIDE EFFECT

**ASSUMPTION:** $\alpha = 0.05$

<table>
<thead>
<tr>
<th>Number of Pairs</th>
<th>Power With Neglecting Pairing</th>
<th>Power With Pairing</th>
<th>Power With Neglecting Pairing</th>
<th>Power With Pairing</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>.93</td>
<td>.93</td>
<td>.64</td>
<td>.53 ($\alpha = .036$)</td>
</tr>
<tr>
<td>160</td>
<td>&gt;.97</td>
<td>.98</td>
<td>.81</td>
<td>.82</td>
</tr>
<tr>
<td>200</td>
<td>&gt;.99</td>
<td>&gt;.995</td>
<td>.86</td>
<td>.87</td>
</tr>
<tr>
<td>250</td>
<td>&gt;.99</td>
<td>&gt;.995</td>
<td>.93</td>
<td>.95</td>
</tr>
<tr>
<td>300</td>
<td>&gt;.99</td>
<td>&gt;.995</td>
<td>.96</td>
<td>.97</td>
</tr>
<tr>
<td>350</td>
<td>&gt;.99</td>
<td>&gt;.995</td>
<td>.97</td>
<td>.98</td>
</tr>
</tbody>
</table>

(3) Alternative Studies Using Marine Cohorts

The GAO and the National Academy of Sciences have referred to specific Marine cohorts as candidates for a Herbicide Orange epidemiological study. In one suggested study configuration, 5900 marines who were within one half kilometer of a herbicide spray track on the day of spraying are called the exposed group, while 212,100 marines are considered unexposed. In a second suggested study configuration, 21,900 marines within one half kilometer of a spray path within 4 weeks of spraying are considered exposed, while a remaining 196,100 marines are considered unexposed. A mortality study was proposed in both of these study configurations. The mortality phase of this protocol involves approximately 1200 exposed and 6000 control individuals, so that, on the surface, the Marine studies would appear to be more powerful in a statistical sense due to larger numbers. However, in fact, two factors couple to render the marine studies less powerful than the RANCH HAND study detailed in this protocol. First, calculations show that a soldier standing directly...
under a spray track at the exact time of spraying receives approximately 1/1000 the dose received by RANCH HAND individuals repeatedly disseminating the mixture throughout the usual RVN tour. Thus even if the unlikely event of being directly under a spray path were repeated 10 times during a marine's RVN tour, the marine's dose would still be only 1/100 that of the RANCH HANDERS. The second factor impacting the Marine study power is the difficulty imposed by the fact that troop positions are only very inexactly known. The available data provide only the battalion headquarter's position relative to herbicide spray paths. Thus troops considered to be exposed could be very far from spray paths, and in fact, be unexposed. On the other hand, troops deemed unexposed in terms of their battalion headquarter's position could in fact have been near spray paths on the day of spraying. Thus, the Marine studies are limited by the problem of misclassification in addition to the fact that the marines received a lesser herbicide exposure than RANCH HAND personnel.

It is possible to compare the RANCH HAND study described in this protocol with the Marine studies in a quantitative way. Results of such an analysis are set out in Tables 14 thru 17. In Table 14, the Marine study using 5900 exposed soldiers is contrasted with the RANCH HAND study considering a disease with an incidence of 0.001 in the control groups, and 0.004 in the RANCH HAND exposed cohort. With a relative risk of 4 against a control disease incidence of 0.001, RANCH HAND power is 0.87 while the Marine study power is much less for several combinations of Marine exposure and misclassification. The misclassification figures shown refers to the percentage inclusion of unexposed individuals into the exposed Marine group. For the calculations, disease incidence in the marine exposed group was assumed to be linearly related to exposure. Table 15 is strictly analogous to Table 14 except that the disease state studied has an incidence of 0.01 in the control groups and 0.02 in the RANCH HAND exposed cohort. Again the RANCH HAND study is seen to be significantly more powerful than the Marine study. Tables 16 and 17 directly parallel Tables 14 and 15, respectively, except that the Marine exposed group is considered to consist of 21,900 soldiers. Here again RANCH HAND study power is seen to be significantly superior.

Figure 12 shows the RANCH HAND mortality study power as a function of relative risk, and disease incidence in the control group. Figure 13 shows marine study power versus marine exposure for zero to 25% misclassification and a control disease incidence of 0.001 and RANCH HAND relative risk of 4. For this circumstance it is clear that the marine study becomes competitive with the RANCH HAND power only if one assumes that the marines received approximately one half of the RANCH HAND exposure dose. Figure 14 is the same as Figure 13 except that 21,900 marines are considered exposed. Again the Marine study becomes competitive with the RANCH HAND study only if one can assume the exposed marines received 0.2 or more of the RANCH HAND exposure, an assumption which is not supported by the available data.

C. The Replacement Concept

In the mortality analysis, a randomly selected group of control individuals will be compared to the RANCH HAND group, and the data gathered will be analyzed for evidence of herbicide effect. In the questionnaire and
<table>
<thead>
<tr>
<th>Ranch Hand Power 1-B (%)</th>
<th>% Misclassification</th>
<th>Marine Study Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.87</td>
<td>0</td>
<td>0.18</td>
</tr>
<tr>
<td>10</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>25</td>
<td>0.15</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Assumptions: RH study pop. 1,200: 6,000 (1:5)  
Marine study pop. 5,900: 212,100  
Normal incidence of disease 0.001  
Disease incidence in RH 0.004  
Linear dose - response  
Misclassification of marine controls excluded
TABLE 15 MORTALITY ANALYSIS

POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE

POWER TABLE

<table>
<thead>
<tr>
<th>RANCH HAND POWER</th>
<th>% MISCLASSIFICATION</th>
<th>MARINE STUDY POWER EXPOSURE LEVELS RELATIVE TO RANCH HAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-B</td>
<td>0</td>
<td>.19 .10 .06 .05</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.17 .10 .06 .05</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>.14 .09 .06 .05</td>
</tr>
</tbody>
</table>

ASSUMPTIONS: RH STUDY POP. 1,200: 6,000 (1:5)
MARINE STUDY POP. 5,900: 212,100
NORMAL INCIDENCE OF DISEASE = 0.01
DISEASE INCIDENCE IN RH = 0.02
LINEAR DOSE - RESPONSE
MISCLASS. OF MARINE CONTROLS EXCLUDED
**TABLE 16**

MORTALITY ANALYSIS

POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE *

POWER TABLE

<table>
<thead>
<tr>
<th></th>
<th>RANCH HAND POWER 1-B</th>
<th>MARINE STUDY POWER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% MISCLASSIFICATION</td>
<td>EXPOSURE LEVELS Relative to Ranch Hand</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/10   1/20   1/100  1/1000</td>
</tr>
<tr>
<td>.87</td>
<td>0</td>
<td>.38     .17     .07      .05</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.33     .15     .08      .05</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>.26     .13     .06      .05</td>
</tr>
</tbody>
</table>

**ASSUMPTIONS:**
- RH STUDY POP. 1,200; 6,000 (1:5)
- MARINE STUDY POP. 21,900; 196,100
- NORMAL INCIDENCE OF DISEASE = 0.001
- DISEASE INCIDENCE IN RH = 0.004
- LINEAR DOSE - RESPONSE
- MISCLASS. OF MARINE CONTROLS EXCLUDED

* INCORRECT POPULATION NUMERICS BASED ON ENVIRONMENTAL FATE OF TCDD
TABLE 17

MORTALITY ANALYSIS

POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE *

POWER TABLE

<table>
<thead>
<tr>
<th>RANCH HAND POWER 1-B</th>
<th>% MISCLASSIFICATION</th>
<th>MARINE STUDY POWER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1/10</td>
</tr>
<tr>
<td>.92</td>
<td>0</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>.28</td>
</tr>
</tbody>
</table>

ASSUMPTIONS:
- RH STUDY POP. 1,200; 6,000 (1:5)
- MARINE STUDY POP. 21,900: 196,100
- NORMAL INCIDENCE OF DISEASE = 0.01
- DISEASE INCIDENCE IN RH = 0.02
- LINEAR DOSE - RESPONSE
- INCORRECT POPULATION NUMERICS BASED ON ENVIRONMENTAL FATE OF TCDD
- MISCLASS. OF MARINE CONTROLS EXCLUDED
POWER VERSUS RELATIVE RISK,
1:5 MORTALITY STUDY BY THREE
DISEASE INCIDENCES
FIGURE 13

POWER CURVES OF THE MARINE STUDY CONSIDERING RELATIVE EXPOSURE AND MISCLASSIFICATION OF THE STUDY POPULATION

5,900 EXPOSED
212,100 CONTROL
RH RR = 4.0
p 1 = .001
p 2 = .004

MARINE EXPOSURE/RANCH HAND EXPOSURE
FIGURE 14

POWER CURVES OF THE MARINE STUDY CONSIDERING RELATIVE EXPOSURE AND MISCLASSIFICATION OF THE STUDY POPULATION

INCORRECT POPULATION NUMERICS BASED ON ENVIRONMENTAL FATE OF TCDD

NO MISCLASSIFICATION
25% MISCLASSIFICATION

21,900 EXPOSED
196,100 CONTROL
RH RR = 4.0
p₁ = .001
p₂ = .004
physical examination phases of this study, one of the mortality controls will be randomly selected for each RANCH HAND individual. During the physical examination phase, we must anticipate a significant degree of unwillingness to participate, particularly on the part of control personnel. This loss to study can result in significant bias and loss in statistical power; thus the replacement concept has been developed to mitigate these consequences.

In this replacement strategy, we make use of the control individuals matched with each RANCH HAND person. As previously noted, this is accomplished using computerized data files and the matching parameters of age, AFSC, and race. With each RANCH HAND individual R_i there will be associated ten controls C_{i1}, C_{i2}, C_{i3}, ..., C_{i10}. The first of these controls, C_{i1} will be employed in the questionnaire and physical examination phases of the study. If C_{i1} is alive, but unwilling to participate in the study, he will be replaced by another randomly selected participant with similar perception of health status. In order to avoid bias in morbidity analyses, no dead control will be replaced.

It is important to emphasize that all replacement controls will be carefully flagged so that they may be treated separately in the statistical analysis. These replacements will be carefully compared to the lost controls to develop indicators of comparability (e.g., morbidity and mortality experience). The initial analysis will be performed on the intact exposed/control pairs. Additional analysis will be conducted on all pairs, both those intact, and those with replaced controls. If we consider RANCH HAND individual R_i with living control C_{i1}, we can calculate the probability that control C_{i_k} will be available for the 1st, 2nd and 3rd physical examinations. To examine this question, a small computer Monte Carlo simulation was required. A short BASIC language computer program and glossary are included in Appendix Table A-8. This simulation examines the effect of non-participation expressed as two probabilities P_1 and P_2. Figure A-2 displays the expected participation by the RANCH HAND population, and control group participation is expected to be somewhat less. P_1 is the probability that when first asked to attend a physical examination, the control individual will not comply. P_2 is the probability that a control individual who has agreed once to a physical examination, will not comply for a subsequent examination. In general, P_1 may be greater than P_2. Note that the probabilities P_1 and P_2 must reflect all causes of non-compliance including morbidity and mortality. Table 18 displays a representative simulation run, which provides the number of controls required to find willing matches for 1000 RANCH HAND personnel.

The potential bias introduced by non-willingness in controls can be analyzed statistically. If P_C(x) is the probability density function for compliant individuals and P_{NC}(x) is the same function for non-compliant individuals, we have
### Table 18

**CONTROL DISTRIBUTIONS BY EXAMINATION MATCHING 1000 RANCH HAND PERSONNEL**

\((p_1 = .70, p_2 = .25)\)

<table>
<thead>
<tr>
<th>CONTROL COHORT</th>
<th>EXAMINATION NUMBER</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>C_1</td>
<td>318</td>
<td>237</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>C_2</td>
<td>211</td>
<td>188</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>C_3</td>
<td>131</td>
<td>133</td>
<td>136</td>
<td></td>
</tr>
<tr>
<td>C_4</td>
<td>96</td>
<td>101</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>C_5</td>
<td>74</td>
<td>89</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>C_6</td>
<td>49</td>
<td>68</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>C_7</td>
<td>34</td>
<td>43</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>C_8</td>
<td>25</td>
<td>39</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>C_9</td>
<td>16</td>
<td>18</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>C_{10}</td>
<td>13</td>
<td>20</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Number of Matching Failures</td>
<td>33</td>
<td>64</td>
<td>88</td>
<td></td>
</tr>
</tbody>
</table>

\[ p(x) = \alpha p_C(x) + \beta p_{NC}(x) \]

where \(p(x)\) is the probability density function for the entire population and \(x\) is a vector of important health parameters available on each person. Since \(\int p(x)dx = \int p_C(x)dx = \int p_{NC}(x)dx \neq 1\)

it follows that \(\alpha + \beta = 1\)
and $\alpha$ and $\beta$ may be viewed as coefficients which "mix" the two subpopulations.

If $M_C$ and $M_{NC}$ are the means of the compliant and non-compliant subpopulations respectively, it can be shown that

$$M = \alpha M_{11} + \beta M_{NC}$$

where $M$ is the mean of the entire population. From this last equation, it is clear that as noncompliant individuals are lost (i.e., $\beta$ tends to zero, $\alpha$ tends to one), $M$ tends to $M_C$. Thus the maximum bias is the quantity $M_C - M$.

In this study we propose to replace non-compliant control individuals with matched RANCH HAND control individuals, that is with individuals drawn from a population with density equal to or at least similar to $P_{NC}(x)$. The resulting new density is $p''(x)$ such that

$$p''(x) = \alpha'' p_{11}(x) + \beta'' P_{NC}(x)$$

where

$$\alpha'' + \beta'' = 1$$

$$M'' = \alpha'' M_{11} + \beta'' M_{NC}$$

and where $\tilde{P}_{NC}(x)$ approximates $P_{NC}(x)$. If $\beta''$ is chosen to be close to or equal to $\beta$ above, it appears that $M''$ can well approximate $M$, the true population mean. The difficulty in this approach will be to assure that the replacements are representative of the non-compliant individuals in all respects other than logistic factors impacting willingness to participate in the program.

Our proposed approach is to obtain sufficient data on the unwilling personnel so that a discrimination function of the form

$$D = f(h_1, \ldots, h_1; l_1, \ldots, l_1)$$

can be derived. This function is envisioned to have the following properties:
(a) larger values of $D$ correspond to decreasing probabilities of compliance with the physical examination,

(b) the factors $h_i$ relate to the subjects' health status, while the factors $l_i$ relate to logistic difficulties (distance, job) which tend to preclude attendance at the physical. Factors to be considered in the formulation of this function are displayed in Table 19.

(c) $D$ is an increasing function of each $h_i$ and of each $l_i$.

Table 19

<table>
<thead>
<tr>
<th>FACTORS AFFECTING COMPLIANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health Status ($h_i$)</strong></td>
</tr>
<tr>
<td>Subjective Health Assessment (good/poor)</td>
</tr>
<tr>
<td>Current Utilization of Long-Term Health Care (Yes/No)</td>
</tr>
<tr>
<td>Absenteeism Pattern (Greater Than/Less Than Ten Lost Days in Past Six Months)</td>
</tr>
<tr>
<td><strong>Logistic Difficulties ($l_i$)</strong></td>
</tr>
<tr>
<td>Time Away from Family</td>
</tr>
<tr>
<td>Time Away from Job</td>
</tr>
<tr>
<td>Distance to Examination Site</td>
</tr>
<tr>
<td>Active Pilot</td>
</tr>
<tr>
<td>Income (Greater than/Less than $17,000$)</td>
</tr>
</tbody>
</table>

In the replacement scheme, controls substituted for noncompliant controls, should have identical health factors ($h_i$) as those individuals they replace. The only significant differences should be in the logistic factors ($l_i$). The replacement method should permit correction of non-compliance bias given that health factors $h_i$ and logistic factors $l_i$ are actually distinct. The determination of these two classes of factors will be made using data from the study itself. Specifically, the logistic factors $l_i$ will be independent of health status to the degree testable by the quantity of data available in the study. This replacement strategy has two major advantages: selection bias reduction/estimation and cost reduction. Were replacements not employed, one would be compelled to start the morbidity study with a 4 to 1 or 5 to 1 design in order to insure an adequate number of participating controls on the third physical examination (see Table 18). Such a large control group for physical examination is very costly with little
corresponding gain in study power and with no correction of the selection bias.

D. **Statistical Analysis of Large Data Sets**

A large amount of data will be collected on each subject in this study. Testing at the 0.05 \( \alpha \) level means that in 5 out of 100 instances where there has actually been no herbicide effect, a herbicide effect will be falsely inferred. This is the inverse of the power question which concerns the probability of detecting an event when it actually occurs. If 100 independent measures are taken from subjects one should expect, testing at the 0.05 \( \alpha \) level, that five measures will be positive on the average. This awareness itself should help prevent over reaction to isolated findings. Further, the present protocol does not in fact have one hundred independent measures. Rather the data gathered are grouped into correlated batteries or systems of data. Findings with any given measure will be related to the values of other correlated variables to provide substantiation indicating an authentic finding.

E. **Time-In-Study Effects**

The study outlined in this protocol is expected to involve up to six examinations extending over a period of twenty years. It could be anticipated that participation in the study, by increasing the health awareness of the subjects, would tend to improve the health of the cohorts. The possibility of differential participation in the study by the exposed and control groups could bias against finding a herbicide effect if one exists. The control group could be less willing to participate in the study than will the exposed RANCH HAND personnel. Thus, if on the average, controls spend less time in the study than RANCH HANDERS, and under the supposition that increased time in study will correlate with better health, increased RANCH HAND participation would counterbalance any adverse herbicide health effect.

The corrector for this time-in-study effect is simply to study the relationship between health outcome and participation in the RANCH HAND study by regression or other analogous statistical methods. Participation can be quantitated by such metrics as (a) number of physical examinations attended (b) age at physical examinations attended or (c) pattern of physical examination attendance. Special study design features do not need to be incorporated to properly evaluate time-in-study effects on questionnaire and physical examination portions of the study. However, the effects of differential time-in-study on the mortality analysis must be carefully considered. In order to detect time-in-study effects on mortality, individuals whose mortality are being tracked should have been in the study for the same length of time (both exposed and control individuals), or the distribution of time spent in the study should be similar in both groups. Because of anticipated differential participation between the exposed and control groups, one cannot assume that both cohorts will have equal time in study distributions. Steps must be taken to insure that a proper time-in-study distribution occurs in the control mortality group. Control over this distribution is possible through placement of the mortality cohort in the structure of the control group with
respect to the replacement strategy. The following five designs have been considered:

I. mortality subjects randomized over all ten control positions, and therefore called into the study randomly.

II. mortality subjects in the first five control positions, and therefore called into the study first.

III. mortality subjects in positions #1 and #2, with the three remaining subjects randomized into positions #3 through #10.

IV. mortality subjects in positions #1, #2, #9, and #10, with the one remaining subject randomized in positions #3 through #8.

V. mortality subjects in the first four positions and position #10.

For each of these five designs, certain quantities were calculated. For testing a physical examination effect on mortality, one would require adequate numbers of mortality subjects having had all six physical examinations, and adequate numbers having had none. Therefore, assuming 1200 RANCH HAND subjects,

\[ E_1 = \text{expected number of mortality subjects having all six physical examinations.} \]

\[ E_2 = \text{expected number of mortality subjects never asked to take the physical examination.} \]

\[ E_3 = \text{expected number of mortality subjects having taken no physical examinations.} \]

For testing or modeling time-in-study effects, one would want adequate numbers of mortality subjects having only one physical, having exactly two physicals, etc. Hence, we calculate, for \( J = 1, 2, 3, 4, 5, 6 \):

\[ N_J = \text{expected number of mortality subjects taking exactly } J \text{ physicals (for example } N_3 \text{ is the number of mortality subjects who will have taken three physicals by the end of the study).} \]

and

\[ M_J = \text{expected number of mortality subjects which will actually have taken examination } J. \]

The values of \( E_1, E_2, E_3, N_J, \) and \( M_J \) have been calculated for the five study designs outlined above using an adaptation of the Monte Carlo program.
shown in Appendix Table A8. Best case and worst case situations were con-
sidered. In the worst case, it was assumed that when first asked to partici-
pate, 75% of the subjects refused, while when asked after having once partici-
pated, 50% of subjects refused further contact. In the best case, the first
time refusal rate was assumed to be 50%, and the refusal rate for a subject
who had participated in a prior examination was assumed to be only 15%. Table
20 shows the calculated results. In examining this table it is of interest to
note that the calculated values are not strikingly dependent on study design
configuration. However, for both the worst and best cases, design 2 where the
mortality subjects are placed in the first five control positions, appears
superior and will be used in this study.

Table 20. TIME-IN-STUDY EFFECTS

<table>
<thead>
<tr>
<th>DESIGN</th>
<th>WORST CASE</th>
<th>BEST CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>E1</td>
<td>18 29 25 22 27</td>
<td>267 521 454 428 504</td>
</tr>
<tr>
<td>E2</td>
<td>765 194 580 865 493</td>
<td>3953 2409 3137 3478 2690</td>
</tr>
<tr>
<td>E3</td>
<td>4691 4548 4645 4716 4623</td>
<td>4977 4204 4569 4739 4345</td>
</tr>
<tr>
<td>N1</td>
<td>700 751 713 681 714</td>
<td>227 370 284 239 331</td>
</tr>
<tr>
<td>N2</td>
<td>340 374 347 328 355</td>
<td>185 307 284 239 331</td>
</tr>
<tr>
<td>N3</td>
<td>163 188 168 157 177</td>
<td>146 249 189 161 225</td>
</tr>
<tr>
<td>N4</td>
<td>71 84 76 72 79</td>
<td>116 203 157 136 188</td>
</tr>
<tr>
<td>N5</td>
<td>33 43 36 31 39</td>
<td>94 171 129 110 160</td>
</tr>
<tr>
<td>N6</td>
<td>18 29 25 22 27</td>
<td>267 521 454 428 504</td>
</tr>
<tr>
<td>M1</td>
<td>18 29</td>
<td>267 521</td>
</tr>
<tr>
<td>M2</td>
<td>14 14</td>
<td>46 77</td>
</tr>
<tr>
<td>M3</td>
<td>24 17</td>
<td>53 83</td>
</tr>
<tr>
<td>M4</td>
<td>40 25</td>
<td>62 94</td>
</tr>
<tr>
<td>M5</td>
<td>54 28</td>
<td>76 107</td>
</tr>
<tr>
<td>M6</td>
<td>80 34</td>
<td>85 116</td>
</tr>
</tbody>
</table>
VII. Data Repository

Throughout the period of this investigation, data collection methods will be integrated by use of computer systems. A data repository will be established at the USAFSAM. Master files will be formed for each exposed member and for his matched control/controls. The individual master files will be keyed to one or more identifiers. Confidentiality of data will be maintained by the use of computer generated code numbers. Addresses and telephone numbers of all study subjects will be continually updated to insure proper follow-up.

Individual data items and their sources are as follows:

1. Questionnaire
   a. Initial
   b. In-depth interview (during physical examination)
   c. Follow-up

2. Psychological Battery
   a. Initial
   b. Follow-up

3. Physical Examination
   a. Initial
   b. Follow-up

4. Medical Records
   a. Active duty
   b. VA
   c. Civilian
   d. Dependent

5. Historical Data
   a. Military personnel files
   b. Flight records
   c. Military unit

6. Death Certificates and Autopsy Reports
   a. Study members
   b. Dependents

7. Birth Certificates
   a. Dependents

Mortality data will be obtained from individual medical records, VA records, the screening of personnel records, contact with family or personal physicians, and other available information sources. Date of death (verified by death certificate and available autopsy reports) will be obtained. Cause of death will be expressed as an ICDA number or numbers. The reliability of the mortality data coding will be evaluated by using a dual coding system based on underlying cause of death criteria in use by the National Center for Health Statistics. This will assure that the results of this study are compatible with data based on US mortality statistics. In addition to standard coding for the underlying cause of death, all diagnoses entered on the death certificates will be coded so that multiple cause of death analyses can be conducted.
The computer software for the data analysis phase will be prepared to assure proper data conversion, quality control and standardization of test measurements. Quality control areas will include verification of identification data, range checks, and identification/correction of ambiguous or conflicting data.
VIII. Recognized Study Difficulties and Corrective Measures

A. Medical Precedence

(1) Problem

A departure from the usual methodological approach characterizes this particular epidemiological investigation. Clearly there is no historical "roadmap of methodology" to conduct this study. Most occupational exposure studies use the presentation of an unusual disease to justify the initiation of a comprehensive study. A rare disease or a common disease in an uncommon site, or one with an unusual presentation appearing in space-time clusters, (often in an unusual population or age group) usually generates the requirement for a new study. In the case of Herbicide Orange, the evidence for long-term human effects is tenuous and controversial. Despite the unique problems that this study possesses, such as the lack of clinically defined endpoints, there are many problems that it shares with other occupationally related exposure studies. For example, the question of a latent period in the development of symptoms/signs, the lack of accurate dose-response relationships, and the possibility of a synergistic effect with other toxins/carcinogens are all operating in this study. Since most cohort studies of occupational mortality use the general population as a standard for deriving the expected number of deaths, preemployment selection ("healthy worker" bias) affects the comparative experience. Age-standardized mortality ratios (SMR's) in general are 60-90 percent of the standard in the working population. Similar conflicting results can occur using the matched cohort method proposed in this study design. Statistical verification of the validity of utilizing such a control for a summary mortality index (e.g., SMR) has been infrequently attempted in the past. Inability to verify the validity of the more classical methods of comparing mortality will necessitate the use of multiplicative and/or logistic models to obtain a valid standardized mortality ratio.

(2) Corrective Measures

Study approaches generated by unprecedented occurrences of occupationally related medical complaints require novel approaches, and reorientation beyond standard methods. The success key to this study design is a series of effective, progressive, and helpful peer reviews (all of which have occurred to date and have been incorporated herein). Beyond even the immediacy of the current study, is the growing problem of a myriad of occupationally-related exposures, both in the military and civilian sector, which will require similar epidemiological studies in the future in order to make some judgment as to whether or not an association is of causal significance.

B. Group Accountability Bias

(1) Problem

The numerous media presentations on "Herbicide Orange" issues have focused attention on the RANCH HAND group. Several attempts have been
made to construct lists of former members of this group, and thus, the RANCH HAND population should be somewhat easier to locate and contact than the control population. This difference will be particularly evident with respect to reported mortality experience. The incentives for cooperation and study participation are likely to be greater in the exposed group than in the controls. Also, the close knit reunion association of former RANCH HAND personnel will lead to a more precise reporting of morbidity and mortality in that group. Such group identity tends to decrease the degree of unaccountability in the exposed group while its absence in the controls may lead to underascertainment of mortality. This could then lead to the attribution of excess mortality in the exposed population.

(2) Corrective Measures

Unaccountability bias will be minimized by keeping the percentages of unaccounted for study subjects below 1% in both exposed and control groups. The morbidity and mortality status of all individuals selected for the study will be strongly pursued utilizing a variety of techniques previously described in this document.

C. "Risk Taking" Behavior Bias

(1) Problem

The early RANCH HAND aircrew population was an exclusively volunteer group; the C-130 control population, while volunteers in the Air Force, were not volunteers for special hazardous missions. RANCH HAND mission conditions were considered to be more dangerous than those encountered in the normal combat environment. This suggests that some differences may exist in the psychological profiles of the two groups. A sensation seeking or risk taking psychological orientation may have altered the accident mortality or morbidity patterns of the exposed group. In addition, an accident rate affected by peripheral neuropathy could be masked by undetected risk taking behavior bias.

(2) Corrective Measures

In an attempt to correct for the unique psychological factors that affect the choice of an aeronautical career, and to adjust for the effects of combat stress, transport aircrew members were matched with crewmembers of similar transport aircraft. However, the volunteer nature of the pre-1965 RANCH HAND operation suggests that this basic matching (as an attempt to control for the psychological effects of combat stress) is not totally ideal. The factors of volunteerism and risk-taking behavior must be considered from both the individual and group perspectives. The assessment of individual risk-taking behavior has been quantified by psychological instruments such as the Sensation Seeking Scale (SSS) of Zuckerman, et al. and the Life Experience Inventory (Torrance). The SSS has been demonstrated to have considerable validity in measuring a variety of phenomena including volunteerism and participation in risky activities and has been applied to naval
aviation trainees. This study was unable to demonstrate an increased accident-related mortality in this group of individuals.

D. Response Bias

(1) Problem

False positive response is anticipated as the primary bias operating in this study. Compensation issues arising from individual claims to the VA or from class action suits, heightened health concern generated by extensive publicity, disenchantment with military service, and the simple desire to please the interviewer may introduce positive responses that exceed the study's ability to correct or adjust. False negative response will also operate, and such bias is even more difficult to assess than the spurious response in a positive direction. Significant factors in this direction include: issues of patriotism and loyalty, personal conviction as to the propriety of the defoliation program and their participation in it, the strong virility orientation of the pilot/aircrew population (particularly with reference to questions of libido and fertility), personal inconvenience caused by study participation, errors of memory, and fear of the adverse effects on career goals that abnormal physical examination results could produce (a significant problem for active civilian and military pilots).

(2) Pending Retirement Bias

The military retirement system also creates a potential source of bias when personnel who are approaching the end of their careers exaggerate their symptoms so that they may become eligible for disability benefits.

(3) Corrective Measures

The primary correction technique for questionnaire response bias will be a carefully constructed and standardized physical examination. Multiple verification and bias indicator questions will be designed and included in the initial questionnaire. Memory verification will be conducted by cross-referencing responses to medical and personnel records. Detailed statistical correlations between the questionnaire responses and the physical examination results will be conducted. All interviews and physical examinations will be conducted on a "blind" basis to the maximum extent possible. Self-administered and group-administered questionnaires, which would allow for uncontrolled response changes, will not be conducted. The payment of a $100 per day stipend to all eligible participants will be arranged to increase participation rates. Medical data will not be released to agencies such as the Federal Aviation Administration, and therefore civilian flying activities will not be adversely affected by participation in this study. Models of anticipated biases and their estimated impact on the study will be attempted prior to the final analysis of any phase in order to justify the analytic methods used. Conclusions drawn from this study will be predicted and coupled to a bias estimate.
E. Interviewer Bias

(1) Problem

Voice inflection, speed of interview, intonation and ethnicity are recognized factors which can affect positive or negative interview response. These factors will definitely operate in this study.

(2) Corrective Measures

The questionnaire itself will be developed and refined by a civilian contractor. This contractor will assure that the instrument will elicit sensitive personal and medical information in an accurate and efficient manner, while minimizing discomfort to both the subject and the interviewer. All questionnaires will be administered by well-trained and experienced personnel employed by an opinion research organization under contract to conduct this aspect of the study.

F. Changes to the Protocol

(1) Problem

The question of adverse health effects due to Herbicide Orange exposure in Vietnam has evoked many strong emotions. The actions of consumer groups, environmentalists, and other special interest groups have generated defensive responses on the part of some governmental agencies, and reactive decisions by others. Frequently, these responses have been based on unsubstantiated claims and/or scientific evidence of questionable validity. As a result of these governmental actions, the impact on the planning of this study has been substantial. Suggestions to increase the scope of the effort to include other "exposed" individuals or poorly defined ancillary groups continue to surface. However, problems of group ascertainment, exposure validation, control group selection, and control of additional bias make the inclusion of such individuals undesirable from a sound scientific perspective. If such decisions are made without regard for their scientific impact, compromise of study validity is assured.

(2) Corrective Measures

The scientific groups participating in the extensive peer review process agreed with these concerns. The formation of an effective scientific monitoring group will insure that scientific issues will take precedence over emotional pressures to alter the study design when such changes will limit the scientific validity of the study. The dilution of the scientific credibility of this effort by unscientific decisions will be diplomatically resisted. While all suggested improvements will be considered, any alterations or corrections to the study protocol will be based on sound scientific assessments of the proposed changes. Alterations of the protocol will be made only after careful review and analysis by the principal investigators and the monitoring group.
G. Loss to Study/Statistical and Bias Considerations

(1) Problem

Losses to study in the RANCH HAND group pose a major problem to the validity of the inferences that can be made from any subsequent comparisons between or within groups. The avenues of loss will conceivably arise from individual apathy (volunteer bias), lack of appropriate financial reimbursement for loss of salary, the presence or absence of illness (perception of health), and the lack of a desire for "treatment". Losses of matched controls during the questionnaire and physical examination phases of the study, though predictably greater than in the exposed group, may be managed by replacement from the predetermined set of controls. The estimated participation of individuals is shown in Section XV, Figure A-2. It is estimated that the overall response rate of the exposed group will be 65% in the initial questionnaire and 40% in the physical examination phase of the study. These high non-compliance estimates are expected to occur despite great efforts to keep the questionnaire at an acceptable length, and to coordinate questionnaire administration and physical examination with the subject's personal schedule. Losses to study in either the exposed or control groups will obviously lead to decrements in statistical power, and will raise the possibility of severe bias. Losses from the control group are expected to be greater than losses from the exposed set. Such losses would skew the distribution of controls, (Figure 5) and thus alter the characteristics of the population available for study. If differential losses in the control group occur (i.e., "well" controls dropout more frequently than "ill" controls), a "true" herbicide effect would be diluted (Figure 15). Conversely, if "ill" controls are differentially lost, a spurious effect would be attributed to herbicide exposure. To a lesser extent, losses in the exposed group could create similar effects; however, loss to study in the RANCH HAND population should be much less of a problem then in controls, due to their vested interest.

(2) Corrective Measures

The USAF is committed to expending maximal effort to encourage participation. Loss to study problems in the study participants will be avoided as much as possible by detailed and exhaustive efforts to contact and followup each identified participant. NON-PARTICIPANTS WILL BE STRONGLY ENCOURAGED TO RECONSIDER THEIR INITIAL DECISIONS. Design considerations have been made to minimize loss to study in both the exposed and control populations. Although the USAF can not fully compensate study subjects for lost wages during the physical examination, transportation costs, per diem, and lodging costs will be reimbursed, and a $100 per day stipend will be paid to all eligible participants. The replacement concept will help to counteract the decrement in statistical power, and offset the bias created by differential patterns of loss. The exposed group is already of maximum size and cannot be increased, but non-compliant controls can be replaced. This will maximize the degree of pairing between the two study groups. If a non-compliant control is replaced by a control with a similar perception of