

Chapter XIX

INTERPRETATION OF STUDY RESULTS AND CONCLUSIONS

1. Introduction

This section presents a cautionary note to both scientific and lay readers who may wish to assert that this study, in whole or in part, is supportive or nonsupportive of a causal relationship between exposure to Herbicide Orange (and its dioxin contaminant) and adverse health. It is important to recognize that this observational study cannot prove the "negative," nor can it be construed as "definitive" science. The process of determining causality is complex and must entail a methodical consideration of many factors (Lilienfeld and Lilienfeld, 1980).

2. Causality Factors

In general, the following factors are very important in making an inference of causality: strength of association; dose-response; biologic plausibility; consistency; time relationships; specificity; and coherence. In an epidemiologic study, not all these factors are required to be present in order to make a correct inference, but clearly, substantial conflict between one or more factors casts doubt on an inference of causality.

In this study, numerous group differences (associations) were detected and expressed in terms of probability (P) values. In any given analysis, statistically significant P values (<0.05) represent the strength of the association, but in and of themselves, do not imply an herbicide causation. As expected under the null hypothesis, most a priori hypothesis tests were negative ($P > 0.05$), but the validity of these findings must be assessed by the power of the given test. As expected, many positive associations were found in the clarifying analyses, or as expressions of the influence of specific risk factors (e.g., age, smoking, etc.). Highly significant associations must also be viewed in the context of relative risk. A very significant association with a relative risk of less than two is generally of minor interest from the traditional epidemiologic perspective. In this study, only four objectively determined group differences of $P < 0.05$ had a relative risk of two or greater. Moreover, statistically significant differences in the group means of a laboratory parameter were often detected, but the overall distributions were similar, the values were within normal range, and the clinical relevance of these shifts was not readily apparent (e.g., LDH, testosterone, T_3 , etc.).

A positive linear dose response relationship is a substantial feature in establishing a cause and effect association. A careful counting of the 388 exposure index analyses cited in this report shows that only 11% are statistically significant, and only 2.8% are increasing from low to high

exposure. While these proportions are suggestive of chance associations, this possibility should be modified by the fact that positive exposure analyses, although not totally consistent throughout all occupational categories, tend to aggregate in only several of the organ systems. Additionally, it is recognized that the exposure index has not been fitted to the most specific format, as further experimental studies are still in progress. Thus, the exposure index used herein is a very indirect measure of exposure, making these analyses less certain than the observed group differences. Numerous other subcategorical exposure analyses (also predominantly negative) were accomplished, but were not included at the discretion of the author. Descriptive opinions of the positive exposure associations were often the sole choice of the responsible principal investigator within each chapter.

The time interval from herbicide exposure to onset of subclinical or clinical manifestations is an important concept for proper interpretation of these study findings. The observational period for the detection of possible latent health effects ranges from 10-20 years for all Ranch Handers. While 10-20 years may be insufficient time for the induction of many systemic cancers, and possibly skin cancer, clearly it is of sufficient length to have already "caused" transient biochemical aberrations, birth outcome abnormalities, fertility problems, chloracne, porphyria cutanea tarda, neurologic sequelae, psychological deficits, etc. Thus, if the above acute/subacute conditions are found attributable in these data, it must be acknowledged that the end result of many of the disease processes is being observed. That notion must be reconciled with essentially identical mortality rates in both groups to date, as many of the proposed diseases would most likely have exerted a subtle mortality influence. Alternatively, the suggestion that the release of dioxin from fat may result in slow systemic poisoning, if true, may account for a delay of clinical manifestations beyond classically accepted latent periods. Another influential time-onset relationship is that of "crossover," i.e., a sequential time-disease association based upon a linkage to a pulsed exposure. While many pre/post-SEA analyses have been performed in this study, reapplication of exposure to herbicides (to complete the crossover) via non-SEA vocations or avocations has not, as yet, been exploited.

Other causal factors merit comment. The finding of no cases of soft tissue sarcoma, porphyria cutanea tarda or chloracne in the Ranch Hand group may reflect a lack of specificity and/or a weak toxicity of the received dose of the putative agent (dioxin), or may reflect the low statistical power to detect group differences for these diseases in this study. The absence of these three diseases may also suggest that a synergism with a yet-to-be-discovered factor is required to induce disease. Findings of this study are, as yet, not fully consistent with other human dioxin studies performed in industrial populations. However, this inconsistency may be attributable to different exposure levels. In terms of biological plausibility, there is no discernible syndrome or symptom cluster that has emerged from this study that makes sense, has an

identifiable pathogenesis, or has an analogous animal model. A systemic poisoning theory carries with it the expectation of finding more biochemical abnormalities than were detected in this study.

3. Other Factors

Chloracne has been proposed as a prerequisite to systemic disease. This premise is not wholly consistent with spectrum of illness concepts or other studies which have suggested attributable soft tissue sarcoma in predominantly nonchloracne populations. However, if the premise is true to the extent that the induction of chloracne represents moderate to high exposures to dioxin, then overall, it may be inferred that the Ranch Hand group (with no chloracne) has received relatively low exposure vis-a-vis industrial populations. Assuming a dose-response hierarchy, this inference may be extended to the contemplated studies of U.S. military ground personnel, for if the Ranch Hand study is deemed "negative," so probably will be the other studies of comparable size.

The question of the validity of this study is paramount. Overall, the processes of data collection have been quite good. To the extent possible, biases have been minimized in both the data collection and data analytic phases. Notwithstanding, a general predominance of adverse findings can be noted in the Ranch Hand group. A closer inspection of this aggregation suggests that most statistically significant findings are found in the subjective data sets, as contrasted to the objective measures. Many of these subjective findings in the Ranch Hand group are in various stages of medical record verification at this time. Unfortunately, some areas, e.g., psychological testing by questionnaire, can never be totally verified. Throughout this study, there is a suggestion of differential reporting (MMPI K and Hypochondriasis scales), albeit unanalyzed, that must temper the interpretation of the subjective results. For the objective data, there is good evidence that the laboratory measurements and the clinical assessments were reasonably accurate. This study has duplicated the classical effects of numerous risk factors (age, smoking, alcohol, etc.) on the clinical measurements throughout all organ systems. The detected effects of age and smoking in the functional and count immunologic tests are new observations, to the best of our knowledge. Thus, the effects of these risk factors have been taken into account throughout the study and lend strong credence to the accuracy of the overall group associations, whether statistically significant or not. It is our belief that this physical examination has reflected the true health status of all participants and groups to the maximum extent possible.

4. Conclusions

a. Preface

This section places into context the thousands of statistical tests which have been accomplished on the enormous data bases generated by the population ascertainment efforts, and the administration of the in-home questionnaire and the physical examination. The total baseline study, including all preparatory tasks and the Baseline Mortality Report, has spanned more than 5 years, has required approximately 100 man-years of in-house work, and has cost about \$11M in direct and indirect costs. The Ranch Hand study has been characterized by solid resource support and stringent timetables throughout all levels of government, intense media interest, and outstanding participation of the study subjects. As part of the mosaic of all dioxin research, the Ranch Hand study has been directed to the herbicide-health effect issue in veterans, and specifically, to heavily exposed Air Force personnel.

b. Study Performance Aspects

Of all live Ranch Hand and comparison individuals who were selected for this study, almost all (99.5%) were contacted, eliminating a major element of bias concern. Participation in the in-home questionnaire was 97% and 93% for the Ranch Handers and comparisons, respectively; and similarly 87% and 76% for physical examination. Differential compliance to the examination may have introduced a participation bias, a bias that is potentially related to the true health status of the participant. Age, race, participation in flying, and military status were also significant factors in determining attendance at the examination, but the relative contribution of each factor has not as yet been determined. Traditionally, individuals in either military or civilian commercial flying occupations do not readily volunteer for physical examinations that might disclose even minor ailments that jeopardize their flying careers.

Early in the study, it was discovered that 18% of the comparison group was ineligible for the study because of inappropriate selection due to a computer programming error. Some selected USAF organizational units containing cargo-hauling aircraft were found not to be engaged in RVN duties (a study requirement). Thus, the direction of the error was for overselection and not for underselection of the comparison group. Ineligible individuals were removed from the randomly ordered comparison set. The replacements for the ineligible individuals were the next-in-line proper comparisons. For both these "shifted" comparisons and the next-in-line comparisons who were also used as substitutes for noncompliant individuals, later statistical analyses suggested that they differed from the original eligible comparisons in a variety of subtle and often opposite ways. Because of the possible bias suggested in their use, and because of the time constraints of this report, a conservative management decision was made to base the bulk of statistical tests upon a contrast of the original comparisons to the Ranch Hand group. Several analyses, using the entire

comparison group, were also performed and found not to differ consistently from the analyses based upon the original comparison group. For those analyses which showed differences between the original versus the total comparison group contrast, it is unclear whether these differences are primarily due to true subset variances or to a sample size effect. A full clarification of the complex biases (selection, compliance, overreporting, etc.) must be conducted before the first follow-up phase of the study.

Most of the stringent quality control aspects of the study were monitored and maintained throughout the data collection phases. As a USAF contract requirement, all contractors were required to maintain "blindness" with respect to the exposure status of each individual, thereby reducing examiner bias to an absolute minimum. In addition, by contract all data are the property of the USAF. Study codes were not provided to the contractors.

c. Clinical Aspects

In terms of overall health, the Ranch Handers perceive their state of health to be poorer than that of the comparisons. This finding parallels the examiner's independent assessment. Percent body fat is similar in both groups as are the hematocrit determinations. A higher proportion of abnormal red cell sedimentation rates is found in comparisons under 40 years of age. The proportions are the same in both groups older than 40. The sedimentation rate, hematocrit, percent body fat, self-perception of health, and age are associated pairwise irrespective of group; these relationships are expected as all variables are traditional indicators of nonspecific illness.

There are no significant group differences for malignant or benign systemic tumors. One case of soft tissue sarcoma is noted in a member of the comparison group. A slight nonsignificant aggregation of genitourinary cancers is identified in the Ranch Hand group, and an aggregation of digestive system cancers is observed in the comparison group. Two Ranch Hand bladder cancers are noted at earlier-than-expected ages. A borderline association between systemic cancer and smoking is observed in both groups. Significantly more nonmelanotic skin cancer is observed and verified by medical record review in the Ranch Handers. The predominant cancer, basal cell carcinoma, is the most common skin cancer in the U.S. White male population, and a proper excision is curative. While this finding is of interest, it is emphasized that these data are not adjusted for sunlight exposure, the recognized primary cause of these cancers. This analysis must await more complete data to be collected at the first follow-up examination. Overall there is no consistent data to show that the Ranch Handers are developing uncommon systemic cancers, or cancer in unusual sites, or at a younger age. Both systemic and skin cancers in the Ranch Hand group do not correlate consistently with the herbicide exposure index.

The fertility and reproductive analyses show mixed findings. As these results are largely based upon subjective self-reports, and must be verified by complete medical record and birth certificate reviews, the findings are judged preliminary at this time. A semen analysis on those participants willing and able to provide a specimen shows essentially identical sperm counts and percent abnormal forms between groups. The finding of an increase in sperm count by age is discounted as physiologically significant because of concomitant noncompliance by increasing age. Four measures of fertility show no difference between the Ranch Hand and comparison groups: number of childless marriages; couples with the desired number of children; the fertility index; and the infertility index. There are no significant findings in conception outcomes for miscarriages; stillbirths, induced abortions, or live births. With respect to live birth outcomes, no group differences are observed for prematurity, learning disability, or infant deaths. Birth defects, as cited by parental history, show no group differences for severe or moderate classifications; however, for minor birth defects (simple birth marks, birth rashes, port wine stains, etc.) Ranch Hand offspring show a significant excess. Reported neonatal deaths and physical handicaps significantly predominate in the Ranch Hand group when contrasted to the full comparison group. All analyses are adjusted for as many of the relevant risk factors as possible, e.g., maternal age, maternal smoking, maternal use of alcohol, paternal age, pre/post-RVN service, etc. Herbicide exposure analyses show several findings of statistical significance but the patterns of association are not fully consistent across all occupational categories.

A thorough neurological assessment of the cranial nerves, peripheral nerves, and central nervous system functioning does not disclose any substantive Ranch Hand-comparison group differences. Past history of neurological disorders is similar for both groups. An increased proportion of abnormal Babinski reflexes are noted in the Ranch Handers but this finding is not statistically significant. Detailed nerve conduction velocities are not associated with group membership but are profoundly influenced by alcohol and diabetes. Similarly, abnormalities in sensation to light touch, vibration, and two reflexes are related to abnormal postprandial glucose levels. Exposure index analyses are predominantly negative.

Detailed psychologic evaluations from the in-home questionnaire and physical examination show consistent findings. Educational level of the participant profoundly influences most all of the subjective test results. Due to the inherently high correlation between military rank and educational level, these variables are considered interchangeable. It is emphasized that the majority of psychologic data are based upon highly subjective self-reporting, most of which can never be fully verified by medical record reviews. There are no group differences for reported past emotional or psychological illnesses. However, the high school educated (mostly enlisted) Ranch Handers demonstrate significant findings or deficits in the following categories: fatigue, anger, anxiety, erosion, fear, startle, psychosomatic behavior, hypochondria,

masculinity, and mania/hypomania. It is noted that the high school educated comparisons exhibit a higher degree of denial in most of these categories. These findings are not observed in the college educated Ranch Handers (mostly officers). The Ranch Hand group demonstrates significant hypochondria, depression, hysteria and schizophrenia vis-a-vis the comparison group, after adjustment for education. In sharp contrast, there are no substantial group differences for the more objective functional and performance psychologic tests (e.g., Halstead-Reitan battery, IQ testing). Almost all exposure index analyses are negative. In full context, differential reporting is strongly suggested, albeit unproven. The roles of an overreporting bias and the Post-Vietnam Stress Syndrome will be clarified in subsequent follow-up psychological evaluations.

The hepatic status is assessed by 9 biochemical tests and a variety of questionnaire and medical record data. The results are mixed. Ranch Hand GGPT and LDH levels are slightly higher while cholesterol levels are lower than the comparisons. Alcohol history is associated with most enzymatic elevations in both groups. Ranch Handers report significantly more skin changes compatible with a historical diagnosis of porphyria cutanea tarda (PCT). However, laboratory determinations for delta-aminolevulinic acid, uroporphyrin and coproporphyrin are similar between groups and no cases of PCT were diagnosed at the physical examination. Reported miscellaneous liver disorders, verified by medical record reviews, are found significantly more in the Ranch Handers. The exposure index analyses are generally inconsistent.

A comprehensive dermatologic evaluation reveals no substantial findings in the Ranch Hand group. No cases of chloracne are diagnosed clinically or by biopsy of suspicious lesions. Questionnaire data show that the incidence, severity, duration, and anatomic locations of past acne do not portray a pattern consistent with significant historical chloracne in the Ranch Handers. The classical "eyeglass" distribution of acne (suggesting chloracne) is the same in both groups. Historical acne correlates with the total cumulative acne found at physical examination. All exposure index analyses are negative.

Examination of the central cardiovascular system reveals no remarkable differences between the groups for systolic blood pressure, diastolic blood pressure, abnormal electrocardiograms, past versus present electrocardiograms, or abnormal heart sounds. As expected, abnormalities in most of these parameters are significantly associated with age, smoking, and a past history of heart disease. The three risk factors: age, smoking, and cholesterol level are strongly associated with each other, and HDL cholesterol is significantly influenced by percent body fat and smoking. An analysis of questionnaire data shows that the Ranch Handers are not having premature heart attacks or generalized heart disease, although subset analyses show differing age and smoking effects. As an unexpected finding, two peripheral pulses are significantly diminished or absent in the Ranch Handers, and several other pulses show weak group differences. Clarifying statistical analyses show that

the the aggregate of Ranch Hand peripheral pulses, predominantly leg pulses, are significantly associated with age, past smoking, current smoking, and verified past heart disease. The weak but similar directional findings in the Ranch Hand carotid and femoral pulses are assigned more significance in view of the peripheral pulse observations. State-of-the-art measurement techniques and a specific medical questionnaire will be used to determine the relevance of these pulse deficits at the first follow-up examination. Detailed herbicide exposure analyses show no associations to any of the central or peripheral cardiovascular findings.

Detailed immunological tests, via B and T lymphocyte enumeration and lymphocyte function studies on a randomized subset of all participants, do not demonstrate significant group differences. Because of the high variability of the quality control data, an independent peer review panel evaluated testing methodology and established criteria for analysis. The numbers of T₁₁, T₃, T₄, T₈, B₁, positive cells and total lymphocyte counts are similar in both groups. Smoking history is observed to significantly affect the T₁₁, T₃, T₄, T₈, marker counts and the total lymphocyte count. Age is seen to affect the T₈ count and the total lymphocyte count. No group differences are observed for the functional studies using phytohemagglutinin, concanavallin A, pokeweed mitogen, and tetanus toxoid. Although the baseline proliferation rate (Control #1) was significantly lower in the Ranch Handers, the biologic relevance of this finding is unclear, particularly in the absence of group differences for concanavallin A and phytohemagglutinin stimulation studies. Age is observed to profoundly affect concanavallin A and phytohemagglutinin results while smoking history is seen to significantly influence pokeweed mitogen results. Because of the overall variability of quality control data, interpretation of a specific individual's immunocompetence is not attempted.

Of 8 measured blood elements and parameters, the mean corpuscular volume and the mean corpuscular hemoglobin level are statistically significantly elevated in the Ranch Hand group, but the relative differences are exceptionally minor and are not of clinical relevance or understanding at this time. Seven of the 8 blood measurements are significantly affected by smoking history. Several exposure index analyses demonstrate positive correlations but a consistent pattern by occupational strata is not observed.

There is no group difference in the distribution of reported past pulmonary disease. Forced expiratory volume for one second and forced vital capacity measurements obtained at the physical examination do not reveal group differences that are consistent in character. There are age/smoking/exposure interactions but it is not possible to further delineate these findings at this time. Several statistically significant herbicide exposure index analyses do not conform to classic dose-response relationships.

Ranch Handers report significantly more kidney disease than the comparisons but this history is not corroborated by 6 laboratory measurements obtained at the physical exam. Proteinuria is of borderline significance in the comparison group. Creatinine clearance may be considered of borderline significance in the Ranch Handers, depending on the laboratory value chosen to determine the abnormal category. Because of the substantial problem of compliance to the 24 hour urine collection, little credence is assigned to the creatinine clearance results. Age is observed to influence the blood urea nitrogen and urine specific gravity results while diabetes affected only the specific gravity results. Herbicide exposure analyses are essentially unrevealing.

A comprehensive assessment of thyroid function and insulin and testosterone production show mixed results. Distributional shifts are noted in thyroid function between the Ranch Hand and comparison groups but the test results are generally within the limits of normal values. There are no group differences for diabetes as determined by abnormal 2 hour postprandial glucose levels. Age and percent body fat determinations are associated with abnormalities in T_3 uptake, 2 hour postprandial glucose levels, and testosterone levels. Herbicide exposure analyses show a variety of positive correlations but many are inconsistent across occupational strata.

Evaluations of personal habits and individual health show that Ranch Handers currently smoke cigarettes more than the comparisons, equally participate in high risk sports activities, and have a similar background of traumatic injuries. An unrefined assessment of the total number of abnormalities found at the physical examination show no Ranch Hand aggregations in the high range nor do arbitrary clinically weighted scores. Overall, both groups are comparable in most health respects, and are probably faring better than similarly aged men in the general population.

d. Final Conclusion

This study has disclosed numerous medical findings, mostly of a minor or undetermined nature, that require detailed follow-up. In full context, the baseline study results should be viewed as reassuring to the Ranch Handers and to their families at this time, because this study has not identified statistical group differences for illnesses commonly attributed to dioxin exposure. The data herein suggest that group differences exist which tend to favor the comparisons, but the cause and clinical relevance of these differences is unclear. This baseline report concludes that there is insufficient evidence to support a cause and effect relationship between herbicide exposure and adverse health in the Ranch Hand group at this time.