EXECUTIVE SUMMARY

1987 FOLLOWUP MORBIDITY REPORT

The Air Force Health Study is an epidemiologic investigation to determine whether adverse health effects exist and can be attributed to occupational exposure to Herbicide Orange. The study consists of mortality and morbidity components, based on a matched cohort design in a nonconcurrent prospective setting with followup studies. The Baseline study was conducted in 1982, and the first two followup morbidity studies were performed in 1985 and 1987. The purpose of this report is to present the results of the 1987 followup.

In the Baseline morbidity study, each living Ranch Hand was matched to the first living and compliant member of a randomly selected Comparison set based on age, race, and military occupation, producing an approximate 1:1 contrast. The Comparisons had served in numerous flying organizations that transported cargo to, from, and within Vietnam but were not involved in the aerial spraying of Herbicide Orange. All previous participants and refusals, newly located study members, and replacements (matched on reported health status) were invited. Eighty-four percent (995/1,188) of the eligible Ranch Hands and 77 percent (939/1,224) of the eligible Original Comparisons participated in the 1987 followup examination and questionnaire process. Participation among those who were fully compliant at Baseline was very high. Ninety-two percent of the Ranch Hands and 93 percent of the Comparisons who were fully compliant at Baseline also participated in the 1987 followup. In total, 2,294 study subjects, 995 Ranch Hands and 1,299 Comparisons, participated in the 1987 followup.

The followup study was conducted under contract to the Air Force by Science Applications International Corporation, in conjunction with the Scripps Clinic and Research Foundation and the National Opinion Research Center. Most of the data were collected through face-to-face interviews and physical examinations conducted at the Scripps Clinic in La Jolla, California. Other data sources included medical and military records and the 1982 and 1985 data bases. As a contract requirement, all data collection personnel were unaware of each participant's exposure status, and all phases of the study were monitored by stringent quality control. The statistical analyses were based on analysis of variance and covariance, chi-square tests, Fisher's exact tests, general linear models, logistic regression, proportional odds models, t-tests, and log-linear models, all of which were specified in an analytical plan written prior to data analysis.

The questionnaire and physical examination data were analyzed by major organ system. The primary focus was on the assessment of differences between the Ranch Hand and Comparison groups based on data from the 1987 followup. Additionally, dose-response relationships within the Ranch Hand group were examined, and longitudinal assessments of differences in the changes of the two groups between the examinations were conducted for selected variables.

In the analyses in this report, Ranch Hand exposure to dioxin was quantified by use of a calculated index based on the quantity of herbicides containing dioxin sprayed each month and the number of Ranch Hands assigned to each

occupational category in those months. The statistical relationships between the evaluated conditions and the calculated index were assessed for significance and patterns suggestive of dose-response. However, early results of serum dioxin studies in Ranch Hand personnel conducted at the Centers for Disease Control indicate the calculated index is not a good measure of actual dioxin exposure. Therefore, the results of analyses using the calculated exposure index should be interpreted with caution. A full report relating the serum assay results to the medical data contained in this report is expected in 1991.

The fixed size of the Ranch Hand cohort limits the ability of the study to detect group differences, particularly for the rare occurrences of soft tissue sarcoma and non-Hodgkin's lymphoma. The study has virtually no statistical power to detect low to moderate group differences for these malignancies. The study has good power to detect relative risks of 2.0 or more with respect to disease occurring at prevalences of at least 5 percent in the Comparison group, such as basal cell carcinoma.

Self-perception of health, appearance of illness or distress, relative age, and percent body fat were similar in the two groups. There has been a decline in the percentage of individuals reporting their health as fair or poor in both groups since the Baseline examination. A significantly greater percentage of Ranch Hands than Comparisons, however, had abnormal erythrocyte sedimentation rates. Only three participants (two Ranch Hands and one Comparison) had rates in excess of 100 mm/hr. The Comparison had lung cancer and died in early 1989. In neither of the Ranch Hands was a diagnosis established during the course of the 1987 followup. A significant difference was also detected at the 1985 followup examination, and it will be important to monitor the sedimentation rates in subsequent examinations.

For all verified neoplasms combined, Ranch Hands had a significantly greater frequency than the Comparisons. Ranch Hands also had a marginally significant greater frequency than the Comparisons when suspected neoplasms were included in the analysis. Because cancers fall into systemic or skin categories, group contrasts were performed within each category. Analyses restricted to systemic neoplasms revealed no significant differences between the Ranch Hands and Comparison groups. Focusing only on skin neoplasms, Ranch Hands had significantly or marginally significant higher frequencies for the following categories: all verified skin neoplasms, all verified and suspected skin neoplasms, all verified malignant skin neoplasms, and sun exposurerelated malignant skin neoplasms. Significant group differences for the sun exposure-related malignant skin neoplasms are not surprising because approximately 90 percent of the participants with those neoplasms had verified basal cell carcinomas, and Ranch Hands had significant or marginally significant higher frequencies of verified basal cell carcinoma than the Comparisons.

The neurological assessment did not disclose significant findings detrimental to the health of the Ranch Hands, although several differences were noted. Of the six reported and verified neurological diseases and disorders, the only significant finding was that Ranch Hands had a higher incidence of hereditary and degenerative neurological diseases. Unadjusted analyses for the 30 physical examination variables showed marginally more balance/Romberg sign and coordination abnormalities in the Ranch Hand group than in the Comparison group. In the adjusted analyses, a significant difference in the

relative risk for the cranial nerve index (without range of motion) occurred with insecticide exposure. Stratified results showed that among those who had never been exposed to insecticides, significantly more Ranch Hands than Comparisons were abnormal on this index. Of those who had been exposed to insecticides, the percentage of abnormalities on this index was marginally higher in the Comparisons. The adjusted analysis for coordination detected two significant group-by-covariate interactions (group-by-occupation and group-by-insecticide exposure). Stratified analyses found a significant group difference for enlisted groundcrew after excluding the group-by-insecticide exposure interaction, and a significant adjusted group difference overall after excluding both group-by-covariate interactions. Ranch Hands had significantly more coordination abnormalities than Comparisons for each analysis. The trend of increasing abnormality in the enlisted groundcrew for coordination will be more fully evaluated in the analyses of serum 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) levels.

The psychological assessment was based on the analysis of 52 variables, which included reported illnesses verified by medical record review, reported sleep disorders, and scores from two clinical psychological tests. The results showed that significant or marginally significant differences between the Ranch Hands and the Comparisons were found for some verified psychological disorders, reported sleep disorders, and the self-administered Symptom Checklist-90-Revised and Millon Clinical Multiaxial Inventory psychological examinations. For these differences, the Ranch Hands generally manifested higher percentages of abnormalities or higher mean scores than the Comparisons. However, this is not surprising since individuals who perceive themselves as having been harmed might be more likely to report the symptoms found to be significant in this analysis. These results will be reexamined for positive correlations between the complaints and dioxin levels when the serum assay data become available. Additionally, significant group-by-covariate interactions were frequently observed in the adjusted analysis, which often made direct contrast of the two groups with adjustment for significant covariates difficult. The covariates of age, alcohol history, and presence of post-traumatic stress disorder showed strong effects on many of the psychological measurements. There was generally a lack of consistency in the findings of similar variables in the psychological tests.

The gastrointestinal assessment found no significant group difference for historical liver disease, historical and current ulcer, and current hepatomegaly. The Ranch Hand alkaline phosphatase mean was significantly higher than the Comparison mean, but group differences for the other laboratory examination variables (aspartate aminotransferase, alanine aminotransferase, gamma-glutamyl transpeptidase, total bilirubin, direct bilirubin, lactic dehydrogenase, cholesterol, high density lipoprotein [HDL], cholesterol-HDL ratio, triglycerides, creatine kinase, and fasting glucose) were not significant.

In the dermatologic assessment, no cases of chloracne were diagnosed. For participants with no history of acne before the start of the first Southeast Asia (SEA) tour, a greater percentage of Ranch Hands than Comparisons reported the occurrence of acne after the start of the first SEA tour. However, the anatomic pattern of these lesions was not suggestive of chloracne. No other significant group differences were detected in the remainder of the analyses. The exposure index and longitudinal analyses were also essentially negative;

the few positive findings were inconsistent with dose-response effects and the available knowledge of current serum TCDD levels in the Ranch Hand group.

The cardiovascular evaluation showed that the health of the two groups was similar for reported and verified heart disease and central cardiac function. With regard to peripheral vascular function, the Ranch Hands manifested a marginally higher mean diastolic blood pressure than the Comparisons, but the percentage of individuals with a diastolic blood pressure above 90 mm Hg was not significantly different in the two groups. The Ranch Hands had a marginally higher percentage of individuals with carotid bruits, and there were also significant, or marginally significant, differences with respect to femoral pulses, dorsalis pedis pulses, and three aggregates pulse indices (leg, peripheral, and all pulses), as assessed by manual palpation. Significantly more pulse abnormalities in the Ranch Hands were also found at Baseline, when pulses were measured by manual palpation, but not in the 1985 followup, when both manual and Doppler measurements were utilized.

In the hematologic evaluation, red blood cell count, hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, and mean corpuscular hemoglobin concentration were not significantly different in the two groups. The mean white blood cell and platelet counts were significantly greater in the Ranch Hands than in the Comparisons, but the magnitude of the difference was small in each case. The difference in platelet counts was significant despite that in the longitudinal analysis of the changes from Baseline to the 1987 followup examination, platelet counts in the Ranch Hands decreased to a significantly greater degree than in the Comparisons. The percentage of individuals with abnormally high platelet counts was also significantly greater in the Ranch Hand group, but the relative risk was less than 2. In addition, no platelet count was elevated into a pathologic range. Exposure index analyses did not generally support dose-response relationships.

The groups did not differ significantly in reported history of kidney disease/stones or for urinary protein, urinary occult blood, urinary white blood cell count, blood urea nitrogen, or urine specific gravity based on unadjusted analyses. In the adjusted analyses, there was no pattern of results that suggested a detriment to either group.

For the endocrinologic assessment, the Ranch Hand thyroid stimulating hormone (TSH) mean was marginally significantly higher than the Comparison TSH mean, but results of the TSH discrete analyses did not show statistically significant group differences. Mean levels for triiodothyronine percent (T_3 %) uptake, testosterone, and 2-hour postprandial glucose were similar between groups. The percentage of abnormal levels for each of these variables, and the composite diabetes indicator, was higher for the Ranch Hand group than for the Comparison group, but none of these differences was statis tically significant. Self-reported data on current thyroid function and past history of thyroid disease were similar between groups. Also, the percentages of participants with thyroid or testicular abnormalities diagnosed at the physical examination were not statistically different between groups. Overall, the endocrinologic health status of the Ranch Hand group does not appear substantially different from the Comparison group.

For the immunologic assessment of the 1987 followup, Ranch Hands and Comparisons did not differ on the cell surface markers, functional stimulation tests, total lymphocyte counts, or quantitative immunoglobulins. Statistical analyses of the natural killer cell assay variables adjusting for covariate information were conducted within the Black and nonblack strata. These analyses showed that Black Ranch Hands had higher adjusted mean counts and average percent releases than the Black Comparisons for the natural killer assay measures. The meaning of this observation is unknown. Without adjusting for covariate information, significantly more Ranch Hands had a possibly abnormal reading on the composite skin reaction test than the Comparisons. Adjusting for covariate information resulted in performing group contrasts on the composite skin reaction variable within strata of the lifetime cigarette smoking history variable. For the heavier smoking participants, significantly more Ranch Hands had a possibly abnormal reading on the composite skin reaction test than the Comparisons. Within the other strata, there were no significant differences.

The pulmonary health of the two groups was reasonably similar based on the analyses without adjustment for covariates, although the Ranch Hands had significantly more thorax and lung abnormalities and marginally higher prevalence rates for hyperresonance. When significant interactions involving group were ignored, no significant differences were found in the adjusted analyses. Exploration of the interactions did not identify a consistent pattern. The adverse effects of smoking were evident in all analyses.

The process of inferring causality is complex and must be based on careful consideration of many factors. Any interpretations of the data must consider the biological plausibility, clinical significance, specificity and consistency of the findings, and a host of statistical factors, such as strength of the association, lack of independence of the measurements, and multiple testing. Based on direct and indirect evidence, it is concluded that this study is free of overt bias and the measurement systems used to obtain the data were accurate and valid.

In summary, there is not sufficient evidence at this time to implicate a causal relationship between herbicide exposure and adverse health in the Ranch Hand group. No cases of chloracne or porphyria cutanea tarda, the two most commonly accepted effects of dioxin exposure, were detected in this study. There was a single case of soft tissue sarcoma in each group and one case of non-Hodgkin's lymphoma in a Ranch Hand. The differences noted indicate that reanalysis using dioxin body burden levels and continued medical surveillance are warranted.

AIR FORCE HEALTH STUDY

1987 FOLLOWUP MORBIDITY SUMMARY REPORT

BACKGROUND

INTRODUCTION

In January 1962, President John F. Kennedy approved a program of aerial herbicide dissemination, for the purpose of defoliation and crop destruction, in support of tactical military operations in the Republic of Vietnam (RVN). Under this program, code-named Operation Ranch Hand, approximately 19 million gallons of herbicides were dispersed on an estimated 10 to 20 percent of South Vietnam. From 1962 to 1971, approximately 11 million gallons of Herbicide Orange, the primary defoliant of the six herbicides utilized in the program, were disseminated.

Operation Ranch Hand was the subject of intense scrutiny from the start due to the controversial nature of the program and political sensitivity to chemical warfare charges contained in enemy propaganda. The concerns, which were initially based on military, political, and ecological issues, shifted during 1977 to health issues. Numerous claims of exposure to herbicides, particularly Herbicide Orange and its dioxin contaminant, and subsequent adverse health effects among U.S. military service personnel have resulted in class action litigation and substantial controversy. Social concern for the Herbicide Orange issue continues to be manifest by continuing scientific research, media presentations, congressional hearings, and legal action.

The U.S. Air Force Medical Service's concern for the health of Air Force personnel exposed to herbicides was demonstrated in October 1978 when the Air Force Deputy Surgeon General made a commitment to Congress and to the White House to conduct a health study on the Ranch Hand personnel, the men who disseminated the majority of the defoliants in the RVN. The prevailing reasons for the Air Force Health Study (AFHS) commitment included the availability of a definitive occupational exposure to herbicides, a sufficient sample size for survey and clinical research, the ability to ascertain the population at risk, and an opportunity for the Air Force Medical Corps to fulfill its adage "we care" to the Air Force community.

The Air Force School of Aerospace Medicine, Brooks Air Force Base, Texas, was tasked by the Surgeon General to develop the Study Protocol. In 1982, after extensive peer review, the epidemiologic study began, and the Protocol was published.

Since 1978, numerous animal and human studies of dioxin effects have been planned or initiated by governmental agencies, universities, and industrial firms. The key scientific issue in these studies was the extent of exposure, e.g., who was exposed and how much each individual was exposed. Unfortunately, population identification and exposure estimation, which are critical for a valid study of ground troops, have been scientifically elusive.

It is believed that of all the military personnel who served in the RVN, the Ranch Hand population was the most highly exposed to herbicides. In 1987, the Air Force initiated a collaborative study with the Centers for Disease Control (CDC) to measure the serum dioxin levels in the AFHS morbidity population. The results of that study clearly demonstrate that substantially elevated levels of dioxin can still be found in the serum of some Ranch Hands, as opposed to the absence of elevated levels found in ground troups by CDC. Based on the principle of dose-response, the Ranch Hands should manifest more and/or earlier evidence of adverse health. Thus, the results of the AFHS should serve as an indicator of herbicide effects in ground personnel.

STUDY DESIGN

The purpose of the study is to determine whether adverse health effects exist and can be attributed to occupational exposure to Herbicide Orange. The study, consisting of mortality and morbidity components, is based on a matched cohort design in a nonconcurrent prospective setting with followup studies. The interwoven study elements of multiple mortality assessments, a Baseline morbidity study, and five followup morbidity studies over 20 years provide a comprehensive approach to the detection of attributable adverse health effects. Complete details on the design are provided in the Study Protocol.

For the Baseline study, the population ascertainment process identified 1,264 Ranch Hand personnel who served in the RVN between 1962 and 1971. By the time the first followup began in 1985, an additional 9 Ranch Hands had been identified. Two years later for the second followup, four additional Ranch Hands were identified. A Comparison group was formed, consisting of individuals assigned to Air Force units operating C-130 cargo aircraft in Southeast Asia. Using a computerized nearest neighbor selection procedure, a maximum of 10 Comparisons was selected for each Ranch Hand, matching on age, race, and military occupation. After personnel record reviews, each Ranch Hand who was determined to be eligible and fully suitable for study had an average of 8.2 Comparison subjects.

The mortality component addresses mortality from the time of the RVN assignment. A Baseline mortality study was conducted in 1982, and the mortality followup consists of annual mortality updates for 20 years. For the Baseline study and the first four updates, five individuals were randomly selected from the matched Comparison set for each Ranch Hand for a 1:5 design. Subsequent to 1987, the design was expanded to include all of the individuals in the Comparison population.

The Baseline morbidity component, begun in 1982, reconstructed the medical history of each participant by reviewing and coding past medical records. A cross-sectional element, designed to assess the participant's current state of mental and physical health, was based on comprehensive questionnaires and physical examinations given to the participants. For this component of the study, each living Ranch Hand and the first living member of his Comparison set were selected to participate in the examination. Sequential questionnaires, medical record reviews, and physical examinations in 1985, 1987, 1992, 1997, and 2002 comprise the morbidity study followup.

MORBIDITY COMPONENT

The Baseline morbidity assessment, conducted in 1982, disclosed some differences between the Ranch Hands and Comparisons, but those differences were generally not traditional indicators of dioxin-related disease. The sustained commitment of Congress and the Air Force to pursue the Agent Orange question to its scientific conclusion was demonstrated by the conduct of the first two morbidity followups in 1985 and 1987. The first (1985) followup provided the first opportunity to confirm or refute some of the Baseline findings and to explore longitudinal changes. For the 1985 followup, the mental and physical health status of the participants during the 3-year interval since the Baseline study was assessed. The results of the 1985 followup approximated those of the Baseline examination; however, the Ranch Hands continued to manifest slightly more adverse health conditions than the Comparisons.

In 1987, the second followup was initiated. During a 2-1/2 year period, the data were collected, automated, and analyzed. The 1987 followup was conducted by Science Applications International Corporation (SAIC) in conjunction with Scripps Clinic and Research Foundation and National Opinion Research Center, working as a team with the Air Force.

PURPOSE

The 1987 morbidity followup is the subject of this report. The objective of the morbidity followup is to continue the investigation of the possible long-term health effects following exposure to herbicides containing 2,3,7,8-tetrachlorodibenzo-p-dioxin (or TCDD). This summary describes the procedures and results of the second morbidity followup of the AFHS. Although the blood samples for the measurement of serum dioxin levels were collected during the 1987 followup, the results of this testing are not yet available.

SUMMARY

The study population for the 1987 followup of the AFHS consisted of 2,294 participants: 995 Ranch Hands and 1,299 Comparisons. The personal characteristics and habits of the Ranch Hands and Comparisons were contrasted. The variables selected to characterize the two groups included all of the candidate covariates in the adjusted analyses of clinical endpoints.

The two groups were contrasted on the matching variables (age, race, and occupation), drinking habits, smoking habits, sun exposure characteristics, exposure to carcinogens, selected personal and family health variables, risk-taking behavior, and other characteristics (education, blood type, personality type, post-traumatic stress disorder (PTSD), current military status, and 1986 individual income).

No difference between the two groups were found for the matching variables, personal and family health variables, and other characteristics. The Ranch Hands and Comparisons reported similar current and lifetime alcohol use; however, the average current alcohol use was higher for the Comparisons and

the Ranch Hands had a higher average lifetime alcohol history. These differences were not significant. Significantly more Comparisons than Ranch Hands drank wine at the time of the 1987 followup; however, the mean numbers of wine drinks per day were not significantly different. For lifetime wine history, the distribution of wine drinkers (nonwine drinkers, moderate wine drinkers, and heavy wine drinkers) was significantly different for the two groups. The Comparisons had a higher percentage of moderate wine drinkers than the Ranch Hands. However, the mean number of wine drink-years for the two groups was similar.

At the time of the 1987 followup, the Ranch Hands smoked significantly more cigarettes than the Comparisons. The Ranch Hands had a higher average lifetime cigarette smoking history than the Comparisons, but this difference was not significant. The two groups had similar current cigar, current pipe, and past and recent marijuana smoking habits.

The two groups reported similar sun exposure characteristics. However, significantly more Comparisons than Ranch Hands had an average lifetime residential latitude of less than 37 degrees North.

Differences in reported exposure to carinogens were assessed for 21 carcinogens or groups of carcinogens and one composite exposure variable constructed from reported exposure to 15 of 21 carcinogens. As anticipated, significantly more Ranch Hands than Comparisons reported being exposed to herbicides and insecticides. Reported ionizing radiation exposure was significantly higher in the Comparisons. Marginally significant differences were detected in reported exposure to arsenic (Comparisons>Ranch Hands), chromates (Ranch Hands>Comparisons), and naphthylamine (Ranch Hands>Comparisons). More Ranch Hands than Comparisons reported being exposed to at least one of the carcinogens used to construct the composite exposure variable; the difference was marginally significant. No differences were detected for the other 15 carcinogen variables.

The risk-taking behavior of the two groups was characterized by participation in nine potentially dangerous recreational activities. Significantly more Ranch Hands than Comparisons reported that they had ever ridden surfboards. No differences in participation in the other eight activities were identified.

In summary, the 995 Ranch Hands and 1,299 Comparisons who participated in the 1987 AFHS followup were found to have similar personal characteristics and habits.

EXPOSURE INDEX

INTRODUCTION

An increased incidence of adverse health effects at higher levels of exposure represents a classic increasing dose-response relationship. The potential relationship of clinical endpoints with herbicide exposure can be tested using an estimate of exposure, hereinafter called an exposure index, for each member of the Air Force Health Study Ranch Hand cohort.

An index of potential exposure to any of four 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)-containing herbicides from fixed-wing spray missions was constructed for each Ranch Hand from the available historical data. The index serves as an estimate only, since the actual concentration of TCDD in the herbicides varied from lot to lot and individual assessments of actual body burden during or just after exposure in Vietnam were not feasible. The four TCDD-containing herbicides used in the development of the index are Herbicide Orange, Herbicide Purple, Herbicide Pink, and Herbicide Green. The exposure index was designed to correlate as closely as possible with exposure and is not an exact measure of actual individual exposures. Although the index contains errors when used to assess the exposure of a specific individual, it was thought to provide some degree of useful inference for groups of similarly exposed individuals. The exposure index for each subject is defined as the product of the TCDD weighting factor, the gallons of TCDD-containing herbicide sprayed in the Republic of Vietnam (RVN) theater during the tour of the subject, and the inverse of the number of men sharing the subject's duties during the tour of the subject. Each of these factors is described below.

The TCDD weighting factor reflects the estimated relative concentration of TCDD in the herbicides sprayed. The estimated mean concentrations of TCDD in Herbicide Orange, Herbicide Purple, Herbicide Pink, and Herbicide Green are 2 parts per million (ppm), 33 ppm, 66 ppm, and 66 ppm, respectively. Archived samples of Herbicide Purple indicate a mean concentration of approximately 33 ppm, and samples of Herbicide Orange had a mean concentration of about 2 ppm. Since Herbicide Pink and Herbicide Green contained twice as much 2,4,5-T as Herbicide Purple, the estimated mean concentration of TCDD in these two herbicides was approximately 66 ppm. Based on procurement records and dissemination information, a combination of Herbicide Green, Herbicide Pink, and Herbicide Purple was sprayed between January 1962 and 1965. Using available data on the number of gallons procured and sprayed, the estimated mean concentration of TCDD for this time period was 48.0 ppm.

The Herbs Tape and other data sources indicate that only Herbicide Orange was disseminated after 1 July 1965. Normalizing to Herbicide Orange, the weighting factor becomes 24.0 before 1 July 1965 and 1.0 after 1 July 1965.

Using the Herbs Tape, Contemporary Historical Evaluation and Combat Operations Reports, and quarterly operations reports, a table of gallons of TCDD-containing herbicide sprayed for each month of the operation was constructed. Gallons of Herbicides Purple, Pink, and Green were converted to Herbicide Orange equivalent gallons based on the TCDD weighting factor of 24.0. This information is provided in Table E-1 in the main report.

The dates and occupational category of each Ranch Hand's tour(s) in the RVN were obtained by a manual review of military records. The study design specified five occupational categories: (1) officer-pilot, (2) officer-navigator, (3) officer-nonflying, (4) enlisted flyer, and (5) enlisted ground-crew. Based on the review of the records, the Ranch Hand manning for each occupational category by month was compiled.

A numeric exposure index reflecting the effective number of gallons of Herbicide Orange to which each individual was potentially exposed was computed. For analysis purposes, the values were categorized as high, medium, or

low for each occupational category. Only three occupational categories were used. The three officer categories were combined into one since pilots and navigators were exposed in the same manner and the officer-nonflying category, which included a relatively small number of participants, consisted of administrators whose exposure was considered to be essentially zero. The overall group of "nonexposed" Ranch Hands, estimated at approximately 2 percent of the Ranch Hand group, was analyzed in the low exposure category (see Table 1), conceivably leading to dilution of the exposure analyses and group contrasts.

TABLE 1.

Exposure Index Categorization of
995 Compliant Ranch Hands

Occupational Group	Exposure Category	Effective Herbicide Orange Gallons Corresponding to Exposure Category	Number of Ranch Hand Participants in Exposure Category
Officer	Low	<35,000	130
	Medium	35,000-70,000	124
	High	>70,000	125
Enlisted Flyer	Low	<50,000	55
	Medium	50,000-85,000	63
	High	>85,000	53
Enlisted Groundcrew	Low	<20,000	147
	Medium	20,000-27,000	158
	High	>27,000	140
Total			995

The exposure index categorizations developed for the Baseline study and used in this report are provided in Table 1, along with the frequencies of Ranch Hand participants by occupation and exposure level. The cutpoints for the categories of the exposure index were the 33rd and 66th percentiles of the exposure index distributions within each of the three occupational strata (officer, enlisted flyer, and enlisted groundcrew). Ranch Hands with administrative duties were assigned an index of zero.

DISCUSSION

The calculated exposure index is not specific to each individual and, therefore, may underestimate exposure for those individuals whose jobs required routine handling of herbicide. For example, maintenance schedules for the aircraft herbicide spray tank required that an emergency dump valve be periodically greased, requiring entry into the tank. The current exposure

index cannot distinguish between men who received such exposure and men who did not. The extent to which individuals are misclassified by the current exposure index is not known, precluding bias calculations at this time.

Every laboratory and physical examination endpoint in this study was assessed for dose-response effects versus the calculated exposure index. Current TCDD assay results did not correlate with the exposure index, with or without adjustment for time since exposure. These exposure index analyses are presented because some members of the Advisory Committee of the Science Panel of the Agent Orange Working Group advised that they be included in this report.

Because of the acknowledged imprecision of the exposure index, Air Force efforts are under way to measure TCDD levels in serum collected from participants in the 1987 followup. Serum was obtained for 1,999 of the 2,294 participants and is currently being analyzed by the Centers for Disease Control. As of September 1989, results of 1,366 serum specimens (888 Ranch Hands and 468 Comparisons) have been reported. These results are summarized in Table 2.

These results indicate that (1) Comparisons have background levels; (2) Ranch Hands have higher current TCDD levels than Comparisons; and (3) among Ranch Hands, nonflying enlisted personnel have the highest and officers have the lowest TCDD levels.

The relationship between current TCDD body burden and the constructed exposure index will be described in a future report. This report is expected in early 1991.

TABLE 2.
Serum TCDD Results

Stratum	Ranch Hand			Comparison		
	Sample Size	Median*	Range*	Sample Size	Median*	Range*
Officer--Pilot	247	7.3	0.0-42.6	118	4.7	0.0-13.1
Officer--Navigator	63	9.3	1.1-36.0	27	4.9	2.4-7.9
Officer--Nonflying	19	6.7	3.0-24.9	4	4.0	0.0-4.6
Enlisted Flyer	152	17.2	0.0-195.5	76	4.3	0.0-12.8
Enlisted Groundcrew	407	23.6	0.0-617.8	243	4.2	0.0-54.8
All Personnel	888	12.4	0.0-617.8	468	4.4	0.0-54.8

*In parts per trillion.

INTRODUCTION

Careful consideration of bias, interactions, consistency, multiple testing, dose-response patterns and the exposure index, power limitations, strength of association and biologic credibility is essential to the interpretation of these data. Problems inherent in the evaluation of negative results and the summarization of these data should also be considered.

CONSISTENCY

Ideally, an adverse health effect in Ranch Hands attributable to herbicide or dioxin exposure would be revealed by internally and externally consistent findings. A finding would be regarded as internally consistent if it did not contradict prior information, other findings, or medical knowledge. For example, the finding of increased femoral pulse abnormalities is not consistent with the lack of increased posterior tibial pulse abnormalities in Ranch Hands. Further, the lack of interaction with occupation is not consistent with known patterns of dioxin levels in Ranch Hands. A finding would be externally consistent if it had been previously established either in theory or empirically as related to exposure. The observed excess of basal cell carcinoma in Ranch Hands is externally inconsistent since there is no prior evidence that basal cell carcinoma is related to dioxin or herbicide exposure. It is also internally inconsistent because there is no evidence that basal cell carcinoma relative risk is greater among enlisted ground personnel than the relative risk among officers.

MULTIPLE TESTING

The lack of a predefined medical endpoint has necessitated the consideration of literally hundreds of dependent variables. Each dependent variable is analyzed many different ways to accommodate covariate information and different statistical models. In the hypothetical case that Ranch Hand physical health is the same as that of the Comparisons, about 5 percent of the many statistical tests of hypotheses shown in this report should be expected to detect a group difference (produce p-values less than 0.05). The observation of significant results due to multiple testing, even when there is no group difference, is known as the multiple testing artifact and is common in large studies. Unfortunately, there is no statistical procedure available to distinguish between those statistically significant results that arise due to the multiple testing artifact and those which may be due to a bona fide herbicide effect. Instead, the authors have relied on reasoned consideration of strength of association, consistency, dose-response patterns and biologic credibility to weigh and interpret the findings.

DOSE-RESPONSE PATTERNS AND THE EXPOSURE INDEX

Ideally, a dose-response effect would be represented by an increasing trend in disease prevalence from a low rate among Ranch Hands with low expo-

sure to a high rate among Ranch Hands with a high exposure. A dose-response effect may be expected to occur regardless of the presence or absence of a group difference.

Epidemiologic studies of health effects after environmental or occupational exposure to toxic chemicals or substances have generally relied upon indirect measures of exposure, termed exposure indices, to assess dose-response. For example, a study of respiratory cancer mortality in Montana smelter miners exposed to airborne arsenic trioxide and sulfur dioxide used the number of years of employment as an exposure index for an individual miner. With it, a statistically significant dose-response effect was demonstrated. In the aborted Centers for Disease Control (CDC) study of health effects in US Army troops potentially exposed to Agent Orange in Vietnam, study investigators derived several exposure indices in terms of troop locations, known half-lives of dioxin in soil and on plant leaves, and the dates and spray paths of Ranch Hand aircraft. The study was canceled after their exposure indices failed to correlate with current dioxin levels in assay study subjects. In the Air Force Health Study, each Ranch Hand's dioxin exposure was metricized as the product of the gallons of herbicide sprayed during his tour and the dioxin concentration of that herbicide divided by the number of Ranch Hands in his job category during his tour. This exposure index has so far failed to reveal consistent dose-response effects and is not correlated with current dioxin body burden in Ranch Hands.

The AFHS exposure index was based on the best information available during the design phase of this study. The gallons sprayed, dioxin concentrations and personnel figures are considered accurate. The index is based on the logic that exposure should increase with increased spraying or if fewer men in an occupational category become available to do the work. Similarly, it was reasoned that exposure should decrease as spraying decreased or as more men became available to do the work. The validity of this index is limited, however, since the gallons sprayed and personnel figures are not specific to an individual Ranch Hand's assigned base in Vietnam or to his specific daily work schedule. The AFHS exposure index is probably more accurate than the indices attempted by the CDC because the Ranch Hands were much closer to the herbicide than the Army and because recorded troop locations were somewhat inaccurate. Indirect exposure indices based on work history and demographic information have demonstrated significant dose-response effects in studies of long-term occupational exposure with moderate to high relative risks. Such indices have failed to demonstrate significant effects or have failed to correlate with direct measures of exposure, such as the dioxin assay, when exposures are short in duration, are of less than industrial intensity or when the relative risk is small.

Fortunately, the development of the serum dioxin assay and its application to Ranch Hands and Comparisons has obviated our concern about the calculated exposure index.

TRENDS

An assessment of consistent and meaningful trends is an essential element of the interpretation of any large study with multiple endpoints, clinical

areas, and covariates. However, caution must be exercised in the interpretation of trends.

Increased abnormalities or adverse means for the Ranch Hands across medically related variables within a clinical area might indicate an exposure effect. In this case, it is important to note that there is moderate to strong correlation between endpoints. Hence, the strength of the group differences must also be considered in assessing the extent of the suspected exposure effect.

Based on preliminary results, current dioxin levels are strongly associated with occupation. Thus, strong, statistically significant differences between groups in means or percent abnormalities for different occupations (i.e., group-by-occupation interactions) would be indicative of a dose-response effect. In this situation, one would expect to see a steadily increasing relative risk or difference between means as occupational exposure increased (i.e., officers less than enlisted flyers less than enlisted groundcrew). Under these assumptions, significant interaction with occupation could be due to the absence of a true effect, or the power limitations of the statistical test for interactions.

An increasing trend in differences between groups in means or disease rates with levels of a covariate (other than occupation) could also indicate an exposure effect. For example, an increased relative risk for hepatic disease with increased levels of alcohol consumption could be due to an indirect causal relationship between exposure and hepatic disease through alcohol consumption. In assessing potential indirect causal relationships, it is important to consider the strength of the group differences and consistency of both the results with related endpoints and findings over time (i.e., 1982 Baseline, 1985 followup, 1987 followup examinations).

Based on the calculated exposure index, increasing trends in Ranch Hand disease rates with increasing levels of exposure within occupational category would be expected in the presence of an exposure effect. However, preliminary results of serum dioxin assays of the Ranch Hands indicate that the calculated exposure index is not a good measure of actual dioxin exposure. Thus, the results of the exposure index analysis should be interpreted with caution.

POWER LIMITATIONS

The fixed size of the Ranch Hand cohort limits the ability of this study to detect group differences. This limitation is most obvious with regard to specific types of cancer, such as soft tissue sarcoma and non-Hodgkins lymphoma, which are so rare that fewer than one case is expected in each group and, therefore, this study has virtually no statistical power to detect low to moderate group differences regarding them. On the other hand, these sample sizes are sufficient to detect very small mean shifts in the continuously distributed variables. For example, with regard to IgG, this study has approximately 90 percent power to detect a mean shift of 1 percent. The detection of significant mean shifts without a corresponding indication of increased Ranch Hands abnormalities or disease is considered to be of little importance or an artifact of multiple testing. This study has good power to detect relative risks of 2.0 or more with respect to diseases occurring at preva-

lences of at least 5 percent in the Comparison group, such as heart disease and basal cell carcinoma.

In an attempt to overcome the lack of power to detect group differences for specific types of systemic cancer, all kinds of systemic cancer were combined into a single variable. It is still possible, however, that an increased risk could exist for a particular rare type of cancer and that increased risk would be missed in this study.

STRENGTH OF ASSOCIATION

Ideally, an adverse effect, if it exists, would be revealed by a strong association between group and a disease condition, that is, by a statistically significant relative risk greater than 2.0. Statistically significant relative risks less than 2.0 are considered of less importance than larger risks because relative risks less than 2.0 can easily arise due to unperceived bias or confounding; relative risks greater than 5.0 are less subject to this concern. Relative risks greater than 5.0 were generally not found in this study.

BIOLOGIC CREDIBILITY

The assessment of biologic credibility requires consideration of the question: Is it understood in biologic terms how the exposure under study could produce the effect of interest? While lack of biologic credibility or even a contradiction of biologic knowledge can sometimes lead to dismissal of a significant result as spurious, the failure to perceive a mechanism may reflect only ignorance of the state of nature. On the other hand, it has proven all too easy to propose credible biological mechanisms relating most exposures to most cancers. Thus, while pertinent, the response to this question is not especially convincing one way or the other.

SUMMARIZATION OF RESULTS

Many interpreters will attempt to tally statistically significant results across clinical areas and study cycles. A study of this scope having a multitude of endpoints and no prescribed strength of association to declare an effect meaningful, demands and at the same time defies meaningful summary tabulation. Such summaries are misleading because they ignore correlations between the endpoints, correlations between study cycle results, and the nonquantifiable medical importance of each endpoint. In fact, many endpoints are redundant in an effort not to "miss" anything. Additionally, such tabulations combine endpoints that are not medically comparable. For example, sense of smell is of less medical importance than the presence of malignant neoplasm. Nevertheless, given the lack of adequate summary statistics, the tally of significant results will occur. Such summaries can be misleading and must be carefully interpreted.

CONCLUSION

The interpretation of the AFHS requires careful consideration of potential biases, interactions, consistency of results, the multiple testing artifact, dose-response patterns and the exposure index, power limitations, strength of association and biologic credibility. Additionally, any assurances of safety drawn from these data are not scientifically valid and should be avoided. The AFHS is large enough to establish hazard (for disease prevalences on the order of 5%), but is not large enough to establish safety. Simple tabulations of positive results can be misleading.

SUMMARY OF CONCLUSIONS

INTRODUCTION

This section summarizes the conclusions drawn from the statistical analyses that have been conducted on the Air Force Health Study data base. The 1987 followup was the logical extension of the 1982 Baseline and the 1985 followup, building upon the strengths of the previous studies and utilizing the data collected at the Baseline, 1985 followup, and 1987 followup. The high level of participation that characterized the Baseline and 1985 studies was maintained through the 1987 followup.

STUDY PERFORMANCE ASPECTS

Of the 2,919 study subjects who were eligible to attend the 1987 followup, 2,853, or 97.7 percent, were located and asked to participate in the 1987 followup. Participation in the 1987 followup was high. In total, 2,294 study subjects (995 Ranch Hands and 1,299 Comparisons) were fully compliant. This represented compliance rates of 84 percent and 75 percent for Ranch Hands and Comparisons, respectively. Of the living study subjects who were fully compliant at Baseline, 92.2 percent of the Ranch Hands and 93.2 percent of the Comparisons returned to participate in the 1987 examination. Of the 2,853 invited study subjects, 531 (171 Ranch Hands and 360 Comparisons) refused to participate. One Ranch Hand and 27 Comparisons (all new to the study) agreed to complete the Baseline questionnaire, but failed to attend the physical examination and were thus partially compliant.

Study participation was analyzed to assess the potential for compliance bias. The negative findings suggested that there has been no change in the way new and replacement Comparisons self-selected for entry into the 1987 followup from the Baseline and 1985 studies. Based on analysis of telephone interview data, there appeared to be little selection bias due to nonparticipation.

POPULATION CHARACTERISTICS

Overall, the Ranch Hands and Comparisons had similar personal characteristics and lifestyle habits. No significant differences were found in age, race, occupation, education, current military status, and individual income. Although current and lifetime alcohol use were similar for the two groups,

significantly more Comparisons than Ranch Hands reported that they drank wine both at the time of the physical examination and during their lifetimes; however, the current and lifetime wine consumption means were similar for both groups. Ranch Hands smoked significantly more cigarettes per day than the Comparisons at the time of the physical examination, but there was no difference between the groups on lifetime cigarette smoking, current cigar and pipe smoking, and recent and past marijuana smoking habits. In general, risk-taking behavior of the Ranch Hands and Comparisons was comparable.

In addition to the characteristics and habits summarized above, analyses were conducted to detect group differences on all other variables that were candidate covariates in the adjusted analyses of clinical endpoints. In general, the groups were similar for these variables as well.

PATTERNS OF RESULTS

The conclusions reached in this report were carefully considered using the criteria of consistency, specificity, coherence, strength, and plausibility as they apply to the interpretation of group differences. To form an overall assessment, patterns of results that emerged from the clinical evaluations were examined. Few significant group differences were noted for the proportion of abnormalities. In general, the positive associations did not aggregate in the clinical areas of prime concern; some of the statistically significant group differences noted at Baseline or at the 1985 followup examination have disappeared and only a few new associations have emerged. The longitudinal analyses were primarily negative. The unadjusted results have been concordant with the adjusted results, both in terms of the magnitude and statistical significance of the group differences. Associations between the covariates and the dependent variables generally behaved as expected. No consistent pattern of group-by-covariate interactions emerged, and the exposure index analyses were generally not significant and did not support a dose-response relationship. Dose-response relationships were not emphasized in reaching final conclusions because of the acknowledged limitations of the calculated exposure index used in this report. Dioxin body burden levels will be analyzed in a subsequent report and will provide a more valid indicator of the level of exposure.

CLINICAL ASPECTS

This section provides the conclusions from the analyses of the 12 clinical areas. The results for the dichotomous and continuous variables are summarized in Appendix R of the main report.

General Health

General health in the Ranch Hand and Comparison groups was assessed by five measures (self-perception of health, appearance of illness or distress, relative age, percent body fat, and the erythrocyte sedimentation rate). There were no significant group differences, either unadjusted or adjusted for covariates (age, race, occupation, and, in the case of self-perception of health and sedimentation rate, personality type), nor any significant

group-by-covariate interactions for self-perception of health, appearance of illness or distress, relative age, or percent body fat. There was little difference in the geometric mean values of erythrocyte sedimentation rate in the two groups, but the Ranch Hand group had a significantly higher percentage of individuals with an abnormal sedimentation rate (>20 mm/hr) than the Comparisons. However, only three participants (two Ranch Hands and one Comparison) were found to have rates in excess of 100 mm/hr. One participant (a Comparison) proved to have lung cancer and died in early 1989. For neither of the two Ranch Hands was a diagnosis established during the course of the 1987 followup. Exposure index analyses did not detect any consistent dose-response relationships. Longitudinal analyses revealed a similar decline in both groups over time in the percentage of individuals reporting their health as fair or poor. For sedimentation rate, there was a significant difference between groups in the change from Baseline to the 1987 followup examination, with a relatively greater number of Ranch Hands than Comparisons shifting from normal at Baseline to abnormal at the followup examination. The clinical meaning of this observation is unknown.

Malignancy

The unadjusted analysis of all verified neoplasms indicated that the proportion of Ranch Hands with neoplasms was significantly greater than that of the Comparisons. After including suspected neoplasms with verified neoplasms, the Ranch Hand proportion was marginally greater than the Comparison proportion. The majority of malignant neoplasms observed in the Ranch Hands were basal cell carcinomas, a nonlife-threatening form of cancer. When the analysis was performed only on skin neoplasms for nonblack participants, significantly more Ranch Hands had skin neoplasms than did the Comparisons for both the verified and the verified and suspected diagnoses. A significantly greater proportion of Ranch Hands had verified malignant skin neoplasms than did the Comparisons. Given the presence of a neoplasm, a marginally significant higher proportion of Ranch Hands had skin neoplasms than did the Comparisons.

No significant group differences were found in the analyses of systemic neoplasms by number, behavior (malignant, benign, uncertain, or unspecified), or by location and site. Thus, the increase in overall malignancy was due to elevated relative risks for skin cancer and basal cell carcinoma. Also, given the presence of any systemic neoplasm, Ranch Hands and Comparisons did not differ significantly for malignant systemic neoplasms. The number of soft tissue sarcomas and non-Hodgkin's lymphomas were comparable in the two groups.

For unadjusted analyses of verified basal cell carcinoma, a borderline significant group difference was found. The unadjusted analysis of the verified and suspected basal cell carcinomas was not significant. After adjustment for covariates was performed, the group contrast was statistically significant for verified basal cell carcinoma and borderline significant for the verified and suspected diagnoses. Ranch Hands and Comparisons differed significantly on the frequency of participants with zero, one, or multiple verified basal cell carcinomas. Also, the Ranch Hands had a significantly higher percentage of participants with multiple verified basal cell carcinomas than did the Comparisons.

Sun exposure-related malignant skin neoplasms also exhibited group differences. Approximately 90 percent of the participants with sun exposure-related malignant neoplasms had basal cell carcinomas. For the unadjusted analysis, the group contrast was significant for the verified diagnoses and borderline significant for the combination of verified and suspected sun exposure-related malignant skin neoplasms. For the adjusted analyses of these neoplasms, the Ranch Hands and Comparisons differed significantly for both the verified and combined diagnoses. Verified neoplasms of the upper extremities for the sun exposure-related malignant skin neoplasms also exhibited a significant unadjusted group difference. Examining the sun exposure-related malignant skin neoplasms by occupation produced a borderline significant group difference between the Ranch Hand and Comparison officers for verified malignancies of the ear, face, head, and neck.

The fixed size of the Ranch Hand cohort limits the ability of the study to detect group differences, particularly for the rare occurrences of soft tissue sarcoma and non-Hodgkin's lymphoma. The study has virtually no statistical power to detect low to moderate group differences for these malignancies. The study has good power to detect relative risks of 2.0 or more with respect to disease occurring at prevalences of at least 5 percent in the Comparison group, such as basal cell carcinoma.

Neurological Assessment

The neurological health of the Ranch Hand group was not substantially different from the Comparison group. Of the six questionnaire variables relating to neurological disease, the only significant finding was that Ranch Hands had a higher incidence of hereditary and degenerative neurological disease, such as Parkinson's disease and benign essential tremor. The statistical results of the group contrasts for 30 physical examination variables relating to cranial nerve function, peripheral nerve status, and central nervous system coordination processes were generally not significant. Unadjusted analyses disclosed marginally more balance/Romberg sign and coordination abnormalities for Ranch Hands than for Comparisons. Conversely, Ranch Hands had significantly fewer biceps reflex abnormalities than Comparisons. The adjusted analyses revealed a significant group-by-insecticide exposure interaction for the cranial nerve index (excluding neck range of motion). Stratified results showed a relative risk significantly greater than 1 for participants who had never been exposed to insecticides, and a relative risk marginally less than 1 for participants who had been exposed to insecticides. The adjusted analysis for coordination detected differences in the relative risks with occupation and insecticide exposure. Stratified analyses found a significant group difference for enlisted groundcrew who had never been exposed to insecticides. There were no significant differences for the other strata. Further investigation found a significant group difference for enlisted groundcrew after excluding the insecticide interaction, and a significant adjusted group difference overall after excluding both interactions. Ranch Hands had significantly more coordination abnormalities than Comparisons for each analysis. The trend of increasing abnormality in the enlisted groundcrew for coordination will be more fully evaluated in the analyses of serum 2,3,7,8-tetrachlorodibenzo-p-dioxin levels. The exposure index analyses for each occupational cohort did not reveal significant differences supportive

of a herbicide effect. The longitudinal analyses for the cranial nerve index and the central nervous system index were not significant.

Psychological Assessment

The psychological assessment was based on verified psychological disorders; reported sleep disorders; and two clinical psychological tests, the Symptom Check List-90-Revised (SCL-90-R) and the Millon Clinical Multiaxial Inventory (MCMI). The verified data on lifetime psychological disorders showed no differences for psychoses, drug dependence, and anxiety. However, marginally more Ranch Hands than Comparisons had a verified history of alcohol dependence and other neuroses based on unadjusted analyses. The Ranch Hands reported experiencing great or disabling fatigue during the day and talking in their sleep more frequently than the Comparisons. No group differences were detected in the other 13 sleep disorder variables in the unadjusted analyses. Although no significant differences between the Ranch Hands and Comparisons were found in the unadjusted analyses of the 12 SCL-90-R variables, the Ranch Hands had marginally more abnormalities than the Comparisons for depression, somatization, and an index of the general severity of symptoms. The results of the unadjusted analyses of the MCMI scores revealed that the Ranch Hands had significantly higher mean antisocial and paranoid scores than the Comparisons. Marginally significant differences were identified on the narcissistic and psychotic delusion scores, where the mean score of the Ranch Hands exceeded that of the Comparisons. After adjustment for covariates, a significant difference remained on the narcissistic score. The Comparisons had a significantly higher mean dependent score than the Ranch Hands. Significant group-by-covariate interactions were frequently noted in the adjusted analyses, which made direct contrast of the two groups difficult. The exposure index analyses did not reveal evidence of consistent dose-response relationships.

Gastrointestinal Assessment

Overall, the gastrointestinal assessment did not find the health of the Ranch Hand group to be significantly different from the Comparison group. Group differences based on verified historical data from the questionnaire were not significant for eight categories of liver disease. No significant group difference was found for past or present occurrence of peptic ulcers. The prevalence of hepatomegaly diagnosed at the physical examination was also not significantly different between groups. The only significant finding from the laboratory examination variables was that the Ranch Hands had a higher mean alkaline phosphatase than the Comparisons. This was also noted at the 1985 followup examination. Group differences for the other laboratory variables (aspartate aminotransferase [AST], alanine aminotransferase [ALT], gamma-glutamyl transpeptidase [GGT], total bilirubin, direct bilirubin, lactic dehydrogenase, cholesterol, high density lipoprotein [HDL], cholesterol-HDL ratio, triglycerides, creatine kinase, and fasting glucose) were not significant. Stratified analyses to explore group-by-covariate interactions did not disclose any consistent pattern of significant group differences within a subgroup. The exposure index data often exhibited positive dose-response relationships, but results of the statistical analyses were generally not significant. The longitudinal analyses of AST, ALT, and GGT showed that the

group differences did not change significantly between the Baseline examination and the 1987 followup examination.

Dermatologic Evaluation

Except for more Ranch Hands reporting at least one occurrence of acne during their lifetime than Comparisons, no significant group differences were detected in the dermatologic evaluation. Subsequent analysis of the occurrence of acne indicated that, for participants with no history of acne before the start of the first Southeast Asia (SEA) tour, a higher percentage of Ranch Hands than Comparisons reported the occurrence of acne after the start of the first SEA tour. However, the anatomic distribution of these lesions did not suggest chloracne as a cause. No cases of chloracne were diagnosed in the physical examination. Analyses were conducted on historical occurrence and duration of acne, six dermatologic disorders, a composite variable of other disorders, and a dermatology index of four disorders. All of these analyses found no significant group differences. Exposure index analyses did not reveal consistent patterns that supported an increasing dose-response relationship. The longitudinal analysis, based on the dermatology index, showed no significant differences between groups over time.

Cardiovascular Evaluation

The cardiovascular evaluation was based on reported and verified heart disease (essential hypertension, overall heart disease, and myocardial infarction) and measurement of central cardiac function and peripheral vascular function. Based on reported and verified hypertension and heart disease, the health of the two groups was similar. For reported/verified myocardial infarction, there was a statistically significant difference in the relative risk with family history of heart disease. The relative risk was less than 1 in those with no family history of heart disease and greater than 1 in those with a family history of heart disease, although neither of these within-stratum relative risks was statistically significant.

The assessement of the central cardiac function also found the groups to be similar, although significantly fewer Ranch Hands than Comparisons had bradycardia and more had arrhythmias (borderline significant). There were differences in the relative risk with levels of covariates for systolic blood pressure and nonspecific T-waves, but none of the relative risks was statistically significant in any particular stratum of individuals.

For the peripheral vascular function, significant or borderline differences were detected for five of the eight measurements. The Ranch Hands had a higher or marginally higher mean or percent abnormal for diastolic blood pressure (continuous), carotid bruits, femoral pulses, and dorsalis pedis pulses than did the Comparisons. (No difference between the two groups was detected in the discrete analysis of diastolic blood pressure.) The percentage of radial pulse abnormalities was marginally higher in the Comparisons than in the Ranch Hands. On the three pulse indices (leg, peripheral, and all pulses), the Ranch Hands had marginally or significantly higher percentages of abnormalities than the Comparisons. Arterial occlusive disease is often unilateral rather than bilateral and can affect large vessels proximally or

smaller vessels distally in segmental fashion. Distal circulation may be maintained by good collateral vessels even in the presence of proximal, partial pulse deficits. The Doppler should be more reliable than palpation in such cases, but neither method is perfect. The peripheral pulses were measured by manual palpation in the 1987 followup and at Baseline, when differences were also detected. In the 1985 followup, pulses were assessed manually and by the Doppler technique, and the two groups were found to be similar. The exposure index analyses did not reveal consistent patterns suggestive of a dose-response relationship, except possibly for the presence of arrhythmias in the enlisted flyer cohort, where there were six abnormalities in the high exposure-level category, compared to none in the medium and low exposure-level categories. Longitudinal analysis of electrocardiograph findings and combined mortality-morbidity analyses did not indicate excess cardiovascular risk in the Ranch Hands.

Hematologic Evaluation

The hematologic status of the Ranch Hand and Comparison groups was assessed by the examination of eight variables: red blood cell count (RBC), white blood cell count (WBC), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and platelet count. There were no statistically significant differences between the two groups for RBC, hemoglobin, hematocrit, MCV, MCH, and MCHC, in analyses either unadjusted or adjusted for the covariates of age, race, occupation, current cigarette smoking, and lifetime cigarette smoking history. For WBC, the mean level was significantly greater in the Ranch Hands than in the Comparisons, but the magnitude of the difference was small. The difference was not statistically significant after adjustment for covariates, nor were significant differences detected in the percentage of individuals with abnormal values. Mean platelet counts were also significantly greater in the Ranch Hands than in the Comparisons, as was the percentage of individuals with abnormally high values. While these differences remained significant after adjustment for covariates, no platelet count was elevated into a pathologic range.

Exposure index analyses detected significant exposure level effects in the discrete analysis of hematocrit in the officer cohort, in the continuous analysis of MCV in the enlisted groundcrew, in the continuous analysis of MCHC in the enlisted flyers, and in the discrete analysis of platelet count in the enlisted flyers. Several exposure index-by-covariate interactions were also significant. Only the hematocrit and MCV findings were consistent with a dose-response relationship, however. Longitudinal analyses detected a significantly greater decrease in the mean platelet count from Baseline to the 1987 followup examination in the Ranch Hands than in the Comparisons, despite the higher overall mean count. The clinical importance of these observations is unclear.

Renal Assessment

Without adjustments for covariates, none of the variables of reported history of kidney disease/stones, urinary protein, urinary occult blood, urinary white blood cell count, blood urea nitrogen, and urine specific gravi-

ty showed a significant difference between the two groups. In general, these findings were supported by the adjusted analyses. Examination of the group-by-covariate interactions did not yield a consistent pattern to suggest renal detriment to either group. Lack of a group difference in the reported history of kidney disease/stones (consistent with the 1985 followup results) was in contrast with the Baseline findings, where Ranch Hands reported significantly more disease. A nonsignificant difference in the percentage of participants with urinary protein was also inconsistent with the Baseline examination, when the Comparisons had a marginally significant higher prevalence rate. Like the 1982 and 1985 studies, the exposure index analyses showed very little evidence of a dose-response relationship. In the longitudinal analysis of blood urea nitrogen, no difference in the change over time was detected.

Endocrine Assessment

Findings from the endocrinologic assessment did not disclose any statistically significant differences between the Ranch Hand and Comparison groups. The percentage of participants who indicated problems with current thyroid disease was similar between groups, as were the percentages of thyroid and testicular abnormalities determined by palpation at the physical examination. Of the six laboratory examination variables that were examined (triiodothyronine percent [T_3 %] uptake, thyroid stimulating hormone [TSH], follicle stimulating hormone, testosterone, 2-hour postprandial glucose, and a composite diabetes indicator), the Ranch Hand TSH mean was marginally significantly higher than the Comparison TSH mean, a finding that was statistically significant at the 1985 followup examination. Ranch Hand and Comparison mean levels for the other laboratory variables, including testosterone, were similar. For all laboratory variables, the percentage of Ranch Hands with abnormal values was higher than the percentage of Comparisons with abnormal values, but none of these differences was statistically significant. Exposure index results generally did not support the presence of a herbicide effect. The enlisted groundcrew and officer cohorts exhibited increasing dose-response patterns for diabetes, but the associations were not significant. Conversely, the overall result for diabetes was significant for enlisted flyers, but was due to the presence of relatively more diabetics in the medium exposure category than in either the low or high categories. The longitudinal analyses for the T_3 % uptake, TSH, and testosterone did not show significant differences between groups in the changes over time.

Immunologic Evaluation

For the immunologic assessment of the 1987 followup examination, composite skin reaction test results were analyzed from the physical examination data, and various laboratory examination measurements from cell surface marker studies, three groups of functional stimulation tests, and quantitative immunoglobulins were also analyzed. Ranch Hands had a higher frequency of individuals with possibly abnormal reactions on skin testing than the Comparisons. The analysis of the composite skin tests results, adjusting for covariate information, contained a significant group-by-lifetime cigarette smoking history interactions. Followup analyses showed that, among those individuals with the heaviest smoking histories, Black Ranch Hands had a higher frequency of possibly abnormal readings when contrasted with Comparisons. Within the

other strata, there were no significant differences. The unadjusted analyses of the laboratory examination data indicated no significant group difference between Ranch Hands and Comparisons. For the adjusted analyses of the natural killer assay measurements with and without Interleukin 2, significant interactions between group and race were present. Exploration of these interactions revealed that the Black Ranch Hands had higher adjusted means than the Black Comparisons for the natural killer assay measures. The clinical significance of these findings is not apparent and does not point to any known clinical endpoints. In general, the immunologic assessment revealed no medically important differences between the Ranch Hands and Comparisons.

Pulmonary Disease

The pulmonary assessment was based on five self-reported respiratory illnesses, seven clinical observations, and eight laboratory measurements. No evidence of a herbicide effect was detected in the assessment of the reported respiratory illnesses. The health of the two groups was reasonably comparable based on the clinical and laboratory variables, although the Ranch Hands had a significantly higher percentage of thorax and lung abnormalities on examination than the Comparisons, based on the unadjusted analysis, and a marginally higher percentage after adjustment for covariates. No significant group differences were detected in the adjusted analyses without significant interactions involving group. Exploration of the group-by-covariate interactions did not reveal a consistent pattern indicating a herbicide effect. The adverse effects of smoking on pulmonary status were evident in all analyses.

CONCLUSIONS

In the 1987 followup, relatively few differences in health status were found between the Ranch Hands and Comparisons. No cases of chloracne or porphyria cutanea tarda, the most commonly accepted effects of dioxin exposure, were detected in this study. There was a single case of soft tissue sarcoma in each group and one case of non-Hodgkin's lymphoma in a Ranch Hand. The results do not indicate that the health of the Ranch Hands is related to herbicide exposure in Vietnam. Although few differences were noted, reanalysis of the data using the dioxin body burden levels and continued medical surveillance are warranted.

In summary, there is not sufficient scientific evidence at this time to support a causal relationship between herbicide exposure and adverse health in the Ranch Hand group.

FUTURE DIRECTIONS

The scope and complexity of the Air Force Health Study (AFHS) has required gradual refinement and correction to meet the challenges of changing technology and scientific direction, and to ensure continued participation of all participants.

The selection of procedures to be included in each of the followup physical examinations has been driven by the findings and experiences of the earlier phases of the study. Similarly, changes for the 1992 followup examination will be based on the findings covered in this report. The opportunity to measure dioxin accurately in the blood of study participants will significantly enhance the ability to identify relationships between dioxin and medical findings. It is anticipated that additional serum studies to further explore and characterize dioxin half-life will be performed as part of the 1992 examination.

Additional modifications to the examination format being considered for 1992 include enhanced assessments of psychological and neurological status. The evaluation of immune function will continue to be emphasized in the next examination and the current set of skin test reagents will be used; however, modifications may be made in the dosages of the antigens to reflect World Health Organization recommendations. Modifications to the battery of cell surface marker and functional studies may be made to reflect state-of-the-art laboratory practice. Similarly, advances in laboratory technology, such as the use of fluorometric enzyme assays for thyroid function, will be incorporated as well.

Statistical methodology in the longitudinal analyses will be modified so that data from all four physical examinations can be included in these important analyses. In addition, it is anticipated that a more complete characterization of sun exposure will be available through the use of data that describe the average hours of daily sunshine at each geographic location.

The next 12 to 16 months will see several significant milestones in the AFHS: (1) completion of the reanalysis of verified birth defect and reproductive outcome data; (2) reanalysis of the 1987 examination data in conjunction with the serum dioxin results; and (3) a mortality analysis of deaths through December 1989. These reports should provide information that will be useful in the resolution of the scientific and political questions surrounding the military use of Agent Orange in Vietnam.

REPRODUTIVE OUTCOME
CHILDREN

EXECUTIVE SUMMARY

The Air Force is conducting a 20-year prospective study of the health of veterans of Operation Ranch Hand, the unit responsible for aerial spraying of herbicides in Vietnam from 1962 to 1971. The health of those veterans is compared to that of a group of other Air Force veterans who served in Southeast Asia (SEA) and who had no occupational exposure to herbicides. The study, called the Air Force Health Study (AFHS), is in its tenth year and is designed to determine whether exposure to the herbicides or their contaminant, 2,3,7,8-tetrachlorodibenzo-p-dioxin (dioxin), has adversely affected the health, survival or reproductive outcomes of Ranch Hands.

This report summarizes the findings of an investigation of reproductive outcomes of the 791 Ranch Hands and 942 Comparisons for whom a dioxin level had been determined by August, 1991. These men have fathered 5,489 pregnancies including 4,514 live births. These men are a subset of all Ranch Hands (n=1,098) and Comparisons (n=1,549) who have fathered 8,263 pregnancies and 6,792 live births. All data in this report have been verified by review of birth certificates, newborn clinic records, health records and death certificates. The health status of each child was verified through the age of 18.

The analysis of birth defects in the baseline AFHS report, released in 1984, found that the Ranch Hand rate of reported pre-SEA defects was less than the Comparison rate and the Ranch Hand rate of reported post-SEA defects was greater than the Comparison rate. The baseline finding motivated the verification of conception outcomes and birth defects which are the subject of this report. Reanalysis using verified data also found similar results; however, additional analyses found no indication that these group differences were related to paternal dioxin levels. Furthermore, analyses within each of 13 categories of birth defects found no evidence that this finding was attributable to any specific category of anomalies.

This study is the first to combine a direct measurement of dioxin level with documented and verified reproductive outcomes in a population of sufficient size to provide a reasonable opportunity to detect possible associations between paternal dioxin levels and a range of common reproductive outcomes. This study has good statistical power to detect relative risks of 2 for common birth defects such as musculoskeletal deformities but low statistical power for relative risks of 2 for rare conditions such as chromosomal abnormality or infant death.

Three types of analyses that compared reproductive outcomes to dioxin levels are presented here. Dioxin levels were measured in 1987, and in one analyses those measurements were used to estimate initial doses received in SEA. Current dioxin levels with adjustment for time since departure from SEA, and current categorized dioxin levels were used in the other two analyses.

Analyses of miscarriage, total adverse outcome, total conceptions, birth weight, birth defects, birth defect severity, specific birth defects, infant death and neonatal death were carried out on all conceptions and children and with restriction to full siblings (step-children were excluded) to minimize genetic variation. Additionally, all reproductive outcomes except sperm count, percent abnormal sperm and multiple birth defects were analyzed with and without consideration of the pre-SEA reproductive experiences of these men.

Verification of all live births and conceptions revealed that Ranch Hands and Comparisons misreported birth defects similarly. About 2% of all pre-SEA and post-SEA children had parent-reported birth defects that could not be verified. Both groups under-reported 7% of pre-SEA defects and 14% of post-SEA defects.

Semen

The association between the father's dioxin level and sperm count and the percentage of abnormal sperm was assessed based on semen specimens collected during the baseline examination in 1982. No significant association was found between dioxin and sperm count or the percentage of abnormal sperm.

Miscarriage, Total Adverse Outcome and Total Conceptions

Analyses of miscarriage adjusted for the outcomes of pre-SEA pregnancies were generally negative. Although miscarriages increased with dioxin in conceptions fathered by Ranch Hands with late tours, they decreased in those with early tours. Since it seems implausible that dioxin would act differently in the two groups, it is concluded that dioxin does not affect miscarriage rates. Furthermore, the highest number of post-SEA conceptions was found in Ranch Hands having the highest dioxin levels, which argues against a relationship between dioxin exposure and miscarriage.

Similar to the results obtain for miscarriage, the rate of adverse outcomes increased with dioxin in Ranch Hands with early tours and decreased in Ranch Hands with late tours. Like those for miscarriage, these findings don't make biologic sense and appear unrelated to dioxin.

Birth Weight

Analyses of birth weight with adjustment for birth weights of pre-SEA children were mostly negative. The few significant findings were not suggestive of a dioxin effect. Among pre-SEA children, the rate of abnormally low birth weight in children of Ranch Hands with the lowest dioxin levels (61.2 per 1000) was less than that in children of Comparisons (73.5 per 1000) and in

post-SEA children, the rate in Ranch Hand children (93.3 per 1000) was greater than that in children of Comparisons (41.9 per 1000). This change is due as much to the decrease in the Comparison rate as to the increase in the Ranch Hand rate.

Analyses of birth weight without statistical adjustment for birth weights of pre-SEA children were generally negative or were complicated by interactions with covariates that lack biological explanation. After restriction to full siblings, birth weight decreases with dioxin in some strata and increases in others, suggesting that these findings are chance occurrences. We find no evidence in these data to suggest that birth weight is adversely associated with the father's dioxin level.

Birth Defects

The significance of the association between paternal dioxin level and birth defects was assessed within each of 13 categories of anomalies (total congenital, nervous system, eye, ear face and neck, circulatory system and heart, respiratory system, digestive system, genital, urinary, musculoskeletal, skin, chromosomal and other unspecified). Analyses were first conducted on all children and then with restriction to full sibling children. Each analysis was carried out first without and then with adjustment for covariates.

Few significant associations were found. Those that were found did not appear consistently across analyses and most were not suggestive of a plausible dioxin effect. Some analyses of total congenital anomalies and musculoskeletal deformities found significant relative risks, but no consistent patterns emerged. For example, an analysis of total congenital anomalies found that children of Ranch Hand officers with low dioxin levels had a lower anomaly rate than children of Comparisons. Children of Ranch Hand enlisted flyers and enlisted ground personnel with low dioxin levels had higher rates than children of Comparisons, but the rates in children of fathers with the highest dioxin levels were not elevated. These findings are consistent with the conclusion that the apparent associations are chance occurrences and that there is no underlying association between paternal dioxin and birth defects.

Birth Defect Severity

No consistent pattern of association between birth defect severity and dioxin levels was found. For instance, in some analyses, the highest rates were found in children born to Ranch Hands with intermediate dioxin levels, while the lowest rates were found in children born to Ranch Hands with the highest dioxin levels. We conclude that there is no evidence in these data to suggest that dioxin is adversely associated with birth defect severity.

Specific Birth Defects and Developmental Anomalies

Twelve specific birth defects (anencephaly, spina bifida, hydrocephalus, cleft palate, cleft lip/palate, esophageal atresia, anorectal atresia, polydactyly, limb reduction defects, hypospadias, congenital hip dislocation, Down's syndrome) and 4 developmental anomalies (disturbance of emotion, hyperkinetic syndrome of childhood, specific delays in development, mental retardation) were investigated. Of these, there were only enough occurrences of specific delays in development and hyperkinetic syndrome of childhood to permit statistical analysis.

Analyses of hyperkinetic syndrome with pre-SEA adjustment were entirely negative. Two unadjusted analyses of specific delays in development found significant associations but these were not supportive of a hypothesis of adverse effects of dioxin. One of these findings was caused by decreasing rates with extrapolated initial dioxin. The other was due to high post-SEA rates in children of Ranch Hands with intermediate dioxin levels and lower rates in children of Ranch Hands with high dioxin levels. Analyses of hyperkinetic syndrome without pre-SEA adjustment found one significant association. This finding was caused by a decreasing rate with dioxin in children of Ranch Hands, a finding opposite to the expected dose-response and most likely due to chance. Analyses of specific delays in development without pre-SEA adjustment found one significant association, caused by the rate being higher in children of Ranch Hands with low dioxin levels than in children of Comparisons. The rate in children of Ranch Hands with high dioxin levels was not significantly different from the rate in children of Comparisons.

These findings are weak, inconsistent and often opposite to the expected dose response. They are not supportive of a hypothesis of an adverse association between dioxin and delays in development or hyperkinetic syndrome.

Multiple Birth Defects

Of 1772 post-SEA children included in these analyses, 57 had multiple defects that could not be attributed to recognized syndromes. The few significant associations with dioxin were caused by increased rates of multiple birth defects in children of Ranch Hands with low dioxin levels relative to children of Comparisons. The rates in children of Ranch Hands with the highest dioxin levels were not significantly elevated. These findings are weak and inconsistent with the expected dose-response. We conclude that there is no evidence in these data that dioxin is adversely associated with multiple birth defects.

Infant and Neonatal Mortality

Analyses of infant death were either negative or could not be carried out due to insufficient data. Analyses of neonatal death found two significant associations, both caused by the rate of neonatal death being lower in children of Ranch Hands with the highest dioxin levels than in children of Comparisons (opposite to the expected dose response). We conclude that there is no association between dioxin and infant or neonatal mortality.

Summary

Extensive analyses of verified birth defects and other reproductive outcomes were conducted with the father's serum dioxin level as the measure of exposure.

The lack of significant association between dioxin and total conceptions and between dioxin and any considered semen characteristic provide no support for the claim that Ranch Hand dioxin exposure is adversely related to the ability to father children. Similarly, the lack of association between dioxin and miscarriage, total adverse outcome, birth weight, any of 13 categories of birth defects and neonatal death provides no support for the idea that dioxin is adversely related to reproductive outcomes in this population.

The few positive associations found between dioxin and reproductive outcomes were generally weak, inconsistent or biologically implausible. These data provide no support for the hypothesis that paternal dioxin exposure is adversely associated with reproductive outcomes. Whether dioxin exposure of the mother before or during pregnancy results in abnormalities in the developing fetus or child could not be addressed in this study and remains an open question.

EXECUTIVE SUMMARY

1987 FOLLOWUP MORBIDITY REPORT

The Air Force Health Study is an epidemiologic investigation to determine whether adverse health effects exist and can be attributed to occupational exposure to Herbicide Orange. The study consists of mortality and morbidity components, based on a matched cohort design in a nonconcurrent prospective setting with followup studies. The Baseline study was conducted in 1982, and the first two followup morbidity studies were performed in 1985 and 1987. The purpose of this report is to present the results of the 1987 followup.

In the Baseline morbidity study, each living Ranch Hand was matched to the first living and compliant member of a randomly selected Comparison set based on age, race, and military occupation, producing an approximate 1:1 contrast. The Comparisons had served in numerous flying organizations that transported cargo to, from, and within Vietnam but were not involved in the aerial spraying of Herbicide Orange. All previous participants and refusals, newly located study members, and replacements (matched on reported health status) were invited. Eighty-four percent (995/1,188) of the eligible Ranch Hands and 77 percent (939/1,224) of the eligible Original Comparisons participated in the 1987 followup examination and questionnaire process. Participation among those who were fully compliant at Baseline was very high. Ninety-two percent of the Ranch Hands and 93 percent of the Comparisons who were fully compliant at Baseline also participated in the 1987 followup. In total, 2,294 study subjects, 995 Ranch Hands and 1,299 Comparisons, participated in the 1987 followup.

The followup study was conducted under contract to the Air Force by Science Applications International Corporation, in conjunction with the Scripps Clinic and Research Foundation and the National Opinion Research Center. Most of the data were collected through face-to-face interviews and physical examinations conducted at the Scripps Clinic in La Jolla, California. Other data sources included medical and military records and the 1982 and 1985 data bases. As a contract requirement, all data collection personnel were unaware of each participant's exposure status, and all phases of the study were monitored by stringent quality control. The statistical analyses were based on analysis of variance and covariance, chi-square tests, Fisher's exact tests, general linear models, logistic regression, proportional odds models, t-tests, and log-linear models, all of which were specified in an analytical plan written prior to data analysis.

The questionnaire and physical examination data were analyzed by major organ system. The primary focus was on the assessment of differences between the Ranch Hand and Comparison groups based on data from the 1987 followup. Additionally, dose-response relationships within the Ranch Hand group were examined, and longitudinal assessments of differences in the changes of the two groups between the examinations were conducted for selected variables.

In the analyses in this report, Ranch Hand exposure to dioxin was quantified by use of a calculated index based on the quantity of herbicides containing dioxin sprayed each month and the number of Ranch Hands assigned to each occupational category in those months. The statistical relationships between the evaluated conditions and the calculated index were assessed for

significance and patterns suggestive of dose-response. However, early results of serum dioxin studies in Ranch Hand personnel conducted at the Centers for Disease Control indicate the calculated index is not a good measure of actual dioxin exposure. Therefore, the results of analyses using the calculated exposure index should be interpreted with caution. A full report relating the serum assay results to the medical data contained in this report is expected in 1991.

The fixed size of the Ranch Hand cohort limits the ability of the study to detect group differences, particularly for the rare occurrences of soft tissue sarcoma and non-Hodgkin's lymphoma. The study has virtually no statistical power to detect low to moderate group differences for these malignancies. The study has good power to detect relative risks of 2.0 or more with respect to disease occurring at prevalences of at least 5 percent in the Comparison group, such as basal cell carcinoma.

Self-perception of health, appearance of illness or distress, relative age, and percent body fat were similar in the two groups. There has been a decline in the percentage of individuals reporting their health as fair or poor in both groups since the Baseline examination. A significantly greater percentage of Ranch Hands than Comparisons, however, had abnormal erythrocyte sedimentation rates. Only three participants (two Ranch Hands and one Comparison) had rates in excess of 100 mm/hr. The Comparison had lung cancer and died in early 1989. In neither of the Ranch Hands was a diagnosis established during the course of the 1987 followup. A significant difference was also detected at the 1985 followup examination, and it will be important to monitor the sedimentation rates in subsequent examinations.

For all verified neoplasms combined, Ranch Hands had a significantly greater frequency than the Comparisons. Ranch Hands also had a marginally significant greater frequency than the Comparisons when suspected neoplasms were included in the analysis. Because cancers fall into systemic or skin categories, group contrasts were performed within each category. Analyses restricted to systemic neoplasms revealed no significant differences between the Ranch Hands and Comparison groups. Focusing only on skin neoplasms, Ranch Hands had significantly or marginally significant higher frequencies for the following categories: all verified skin neoplasms, all verified and suspected skin neoplasms, all verified malignant skin neoplasms, and sun exposure-related malignant skin neoplasms. Significant group differences for the sun exposure-related malignant skin neoplasms are not surprising because approximately 90 percent of the participants with those neoplasms had verified basal cell carcinomas, and Ranch Hands had significant or marginally significant higher frequencies of verified basal cell carcinoma than the Comparisons.

The neurological assessment did not disclose significant findings detrimental to the health of the Ranch Hands, although several differences were noted. Of the six reported and verified neurological diseases and disorders, the only significant finding was that Ranch Hands had a higher incidence of hereditary and degenerative neurological diseases. Unadjusted analyses for the 30 physical examination variables showed marginally more balance/Romberg sign and coordination abnormalities in the Ranch Hand group than in the Comparison group. In the adjusted analyses, a significant difference in the relative risk for the cranial nerve index (without range of

motion) occurred with insecticide exposure. Stratified results showed that among those who had never been exposed to insecticides, significantly more Ranch Hands than Comparisons were abnormal on this index. Of those who had been exposed to insecticides, the percentage of abnormalities on this index was marginally higher in the Comparisons. The adjusted analysis for coordination detected two significant group-by-covariate interactions (group-by-occupation and group-by-insecticide exposure). Stratified analyses found a significant group difference for enlisted groundcrew after excluding the group-by-insecticide exposure interaction, and a significant adjusted group difference overall after excluding both group-by-covariate interactions. Ranch Hands had significantly more coordination abnormalities than Comparisons for each analysis. The trend of increasing abnormality in the enlisted groundcrew for coordination will be more fully evaluated in the analyses of serum 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) levels.

The psychological assessment was based on the analysis of 52 variables, which included reported illnesses verified by medical record review, reported sleep disorders, and scores from two clinical psychological tests. The results showed that significant or marginally significant differences between the Ranch Hands and the Comparisons were found for some verified psychological disorders, reported sleep disorders, and the self-administered Symptom Checklist-90-Revised and Millon Clinical Multiaxial Inventory psychological examinations. For these differences, the Ranch Hands generally manifested higher percentages of abnormalities or higher mean scores than the Comparisons. However, this is not surprising since individuals who perceive themselves as having been harmed might be more likely to report the symptoms found to be significant in this analysis. These results will be reexamined for positive correlations between the complaints and dioxin levels when the serum assay data become available. Additionally, significant group-by-covariate interactions were frequently observed in the adjusted analysis, which often made direct contrast of the two groups with adjustment for significant covariates difficult. The covariates of age, alcohol history, and presence of post-traumatic stress disorder showed strong effects on many of the psychological measurements. There was generally a lack of consistency in the findings of similar variables in the psychological tests.

The gastrointestinal assessment found no significant group difference for historical liver disease, historical and current ulcer, and current hepatomegaly. The Ranch Hand alkaline phosphatase mean was significantly higher than the Comparison mean, but group differences for the other laboratory examination variables (aspartate aminotransferase, alanine aminotransferase, gamma-glutamyl transpeptidase, total bilirubin, direct bilirubin, lactic dehydrogenase, cholesterol, high density lipoprotein [HDL], cholesterol-HDL ratio, triglycerides, creatine kinase, and fasting glucose) were not significant.

In the dermatologic assessment, no cases of chloracne were diagnosed. For participants with no history of acne before the start of the first Southeast Asia (SEA) tour, a greater percentage of Ranch Hands than Comparisons reported the occurrence of acne after the start of the first SEA tour. However, the anatomic pattern of these lesions was not suggestive of chloracne. No other significant group differences were detected in the remainder of the analyses. The exposure index and longitudinal analyses were also essentially negative; the few positive findings were inconsistent with

dose-response effects and the available knowledge of current serum TCDD levels in the Ranch Hand group.

The cardiovascular evaluation showed that the health of the two groups was similar for reported and verified heart disease and central cardiac function. With regard to peripheral vascular function, the Ranch Hands manifested a marginally higher mean diastolic blood pressure than the Comparisons, but the percentage of individuals with a diastolic blood pressure above 90 mm Hg was not significantly different in the two groups. The Ranch Hands had a marginally higher percentage of individuals with carotid bruits, and there were also significant, or marginally significant, differences with respect to femoral pulses, dorsalis pedis pulses, and three aggregates pulse indices (leg, peripheral, and all pulses), as assessed by manual palpation. Significantly more pulse abnormalities in the Ranch Hands were also found at Baseline, when pulses were measured by manual palpation, but not in the 1985 followup, when both manual and Doppler measurements were utilized.

In the hematologic evaluation, red blood cell count, hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, and mean corpuscular hemoglobin concentration were not significantly different in the two groups. The mean white blood cell and platelet counts were significantly greater in the Ranch Hands than in the Comparisons, but the magnitude of the difference was small in each case. The difference in platelet counts was significant despite that in the longitudinal analysis of the changes from Baseline to the 1987 followup examination, platelet counts in the Ranch Hands decreased to a significantly greater degree than in the Comparisons. The percentage of individuals with abnormally high platelet counts was also significantly greater in the Ranch Hand group, but the relative risk was less than 2. In addition, no platelet count was elevated into a pathologic range. Exposure index analyses did not generally support dose-response relationships.

The groups did not differ significantly in reported history of kidney disease/stones or for urinary protein, urinary occult blood, urinary white blood cell count, blood urea nitrogen, or urine specific gravity based on unadjusted analyses. In the adjusted analyses, there was no pattern of results that suggested a detriment to either group.

For the endocrinologic assessment, the Ranch Hand thyroid stimulating hormone (TSH) mean was marginally significantly higher than the Comparison TSH mean, but results of the TSH discrete analyses did not show statistically significant group differences. Mean levels for triiodothyronine percent (T_3 %) uptake, testosterone, and 2-hour postprandial glucose were similar between groups. The percentage of abnormal levels for each of these variables, and the composite diabetes indicator, was higher for the Ranch Hand group than for the Comparison group, but none of these differences was statistically significant. Self-reported data on current thyroid function and past history of thyroid disease were similar between groups. Also, the percentages of participants with thyroid or testicular abnormalities diagnosed at the physical examination were not statistically different between groups. Overall, the endocrinologic health status of the Ranch Hand group does not appear substantially different from the Comparison group.

For the immunologic assessment of the 1987 followup, Ranch Hands and Comparisons did not differ on the cell surface markers, functional stimulation

tests, total lymphocyte counts, or quantitative immunoglobulins. Statistical analyses of the natural killer cell assay variables adjusting for covariate information were conducted within the Black and nonblack strata. These analyses showed that Black Ranch Hands had higher adjusted mean counts and average percent releases than the Black Comparisons for the natural killer assay measures. The meaning of this observation is unknown. Without adjusting for covariate information, significantly more Ranch Hands had a possibly abnormal reading on the composite skin reaction test than the Comparisons. Adjusting for covariate information resulted in performing group contrasts on the composite skin reaction variable within strata of the lifetime cigarette smoking history variable. For the heavier smoking participants, significantly more Ranch Hands had a possibly abnormal reading on the composite skin reaction test than the Comparisons. Within the other strata, there were no significant differences.

The pulmonary health of the two groups was reasonably similar based on the analyses without adjustment for covariates, although the Ranch Hands had significantly more thorax and lung abnormalities and marginally higher prevalence rates for hyperresonance. When significant interactions involving group were ignored, no significant differences were found in the adjusted analyses. Exploration of the interactions did not identify a consistent pattern. The adverse effects of smoking were evident in all analyses.

The process of inferring causality is complex and must be based on careful consideration of many factors. Any interpretations of the data must consider the biological plausibility, clinical significance, specificity and consistency of the findings, and a host of statistical factors, such as strength of the association, lack of independence of the measurements, and multiple testing. Based on direct and indirect evidence, it is concluded that this study is free of overt bias and the measurement systems used to obtain the data were accurate and valid.

In summary, there is not sufficient evidence at this time to implicate a causal relationship between herbicide exposure and adverse health in the Ranch Hand group. No cases of chloracne or porphyria cutanea tarda, the two most commonly accepted effects of dioxin exposure, were detected in this study. There was a single case of soft tissue sarcoma in each group and one case of non-Hodgkin's lymphoma in a Ranch Hand. The differences noted indicate that reanalysis using dioxin body burden levels and continued medical surveillance are warranted.

EXECUTIVE SUMMARY

BASELINE MORBIDITY STUDY

The Ranch Hand II epidemiologic study uses a matched cohort design in a nonconcurrent prospective setting, and incorporates mortality, morbidity, and follow-up studies. The purpose of this report is to present the baseline morbidity study.

The morbidity study design matched each living Ranch Hander (by age, job, and race) to the first living and compliant member of a randomly selected comparison mortality set of 5 individuals, producing a 1:1 contrast. The comparison group was formed from numerous flying organizations which transported cargo to, from, and within Vietnam, but were not involved in aerial spray operations of Herbicide Orange. Of the potential study participants, 99.5% were located. Early in the physical examination phase of the study, it was discovered that 18% of the entire comparison group was ineligible to participate because of inappropriate selection. Thereafter, study eligibility was certified only after a hand-review of personnel records. Next-in-line compliant comparisons entered the study as replacements after fully completing the questionnaire and physical examination. Statistical analyses of these replacement individuals later showed that they differed from the original comparisons in a variety of subtle and often opposite ways. As a conservative measure to avoid possible bias by the inclusion of the replacements in the analyses, a management decision was made to base the statistical tests in this report primarily upon contrasts of the Ranch Hand group to the original comparison group.

The preponderance of data was obtained from the in-home interviews and the physical examination, each conducted under contract to the Air Force by Louis Harris and Associates, Inc., New York NY, and the Kelsey-Seybold Clinic, P.A., Houston TX, respectively. All contacts with the participants were carried out with utmost professionalism and sensitivity. Other morbidity data sources included reviews of medical records, military personnel documents, and birth certificates; in-home questionnaires and telephone questionnaires of the study participant's wives, former wives and, occasionally, their next-of-kin. All aspects of the study were voluntary. As a contract requirement, data collection personnel were blind as to the exposure status of the participants. Ninety-seven percent of the Ranch Handers and 93% of the comparisons participated in the in-home interview. For the physical examination, 87% of the Ranch Handers and 76% of the comparison group participated, a total of 2,272 individuals. This differential attendance at the examination may have introduced a potential participation bias that, in a military population predominantly engaged in flying duties, is multifactorial and complex. All study phases were monitored by stringent quality control standards. Statistical analyses of the data consisted primarily of log-linear models, logistic regression techniques, generalized linear models, matched covariate analyses, and Kolmogorov-Smirnov, chi-square, and t tests.

The physical examination and the in-home questionnaire data were analyzed by major organ system. In terms of general health, more Ranch Handers perceived themselves to be in fair or poor health than did their comparisons. No

group differences were detected for hematocrit or percent body fat determinations. Unadjusted group differences in sedimentation rate were not observed; however, significantly more young comparisons had abnormalities in sedimentation rate than did their Ranch Hand counterparts. There were no statistically significant differences in the occurrence of malignant or benign systemic tumors between the groups. One case of soft tissue sarcoma was found in a comparison member. Significantly more nonmelanotic skin cancer was noted in the Ranch Hand group, but these analyses have not yet considered (adjusted for) sunlight exposure, the prime etiology of these cancers. Such nonmelanotic skin cancer (predominantly basal cell carcinoma) is the most common neoplasm in the White population of the United States. Up to the statistical limits of the study there were no consistent data that showed that the Ranch Handers were developing uncommon cancers, or cancer in unusual sites, or at an unusual age. Measures of fertility and reproductive outcome showed mixed results. It is emphasized that the fertility and reproductive results are preliminary at this time as they are based largely upon subjective self reports that await full medical record and birth certificate verification. Four measures of fertility: number of childless marriages, couples with the desired number of children, the infertility index and the fertility index, showed no difference between the Ranch Hand and comparison groups. A semen specimen obtained from those willing and able to provide one showed no group differences with respect to total sperm count or percent abnormal sperm. There were no significant findings in conception outcomes for miscarriages, stillbirths, induced abortions, or live births. For live birth outcomes no differences were observed for prematurity, learning disability, or infant deaths. There was no significant disparity between groups for the classifications of severe or moderate birth defects. By parental history, however, Ranch Hand offspring showed significantly more minor birth defects (birth marks, etc). Reported neonatal deaths and physical handicaps were also significantly excessive in the Ranch Hand group when contrasted to the total comparison group. All fertility and reproductive findings in the Ranch Hand group showed inconsistent relationships to the herbicide exposure index. Medical records and birth certificates are currently being chronicled for complete verification of all historical findings. A comprehensive neurological examination showed no consistent abnormalities in the cranial nerves, peripheral nerves or central nervous system function of the Ranch Handers. As expected, there was a profound influence of diabetes and alcohol in both groups upon numerous neurological tests. Detailed psychologic data were obtained on all participants at both the in-home interview and the physical examination. It is emphasized that the majority of psychological data was derived from self reported responses during interview and has not been fully assessed for the effect of differential reporting. A variety of subjective deficits (fatigue, anger, fear, anxiety, etc) were significantly more common in the high school educated Ranch Handers. Educational level significantly and consistently influenced most subjective test results. In sharp contrast, more objective performance testing by the Halstead-Reitan battery and IQ testing did not reveal any significant intergroup differences. The roles of overreporting and the Post Vietnam Stress Syndrome in these analyses have not as yet been assessed. Liver function tests and clinical history data showed mixed results. Ranch Handers had some elevated liver enzyme tests and lower cholesterol levels. More Ranch Handers were found to have hepatomegaly and verified histories of prior hepatic disease than their counterpart comparisons. Exposure to alcohol, degreasing chemicals, and industrial chemicals in general, influenced

the liver test results. Ranch Handers reported significantly more symptoms resembling porphyria cutanea tarda than the comparisons, but these data have not been verified by medical record reviews nor were they substantiated by laboratory testing or by physical examination. Exposure index analyses were essentially negative. In the dermatologic evaluation, no cases of chloracne were diagnosed clinically or by biopsy. A thorough questionnaire analysis of acne showed that the incidence, severity, duration, and anatomic location did not differ between groups, and suggested that the historical occurrence of chloracne was highly unlikely in the Ranch Handers. Evaluation of the cardiovascular system showed equal proportions of abnormalities in blood pressures, electrocardiograms, past electrocardiograms, and heart sounds in both groups. Ranch Handers are not having premature heart attacks or generalized heart disease. However, the Ranch Handers showed significant deficits in 2 specific peripheral leg pulses and all leg pulses as a group. These puzzling findings were highly correlated with age and smoking patterns, and verified past heart disease. The assessment of the immune system by laboratory testing was compromised by excessive test variability. An independent review committee determined which test data were suitable for statistical analysis. As an unexpected finding, the test data were significantly influenced by the age and smoking history of the participant; no group differences were detected after adjustment for these factors. A hematologic test battery revealed three red cell abnormalities in the Ranch Hand group, but these were difficult to place into a clinical or epidemiologic context. Evaluation of renal, pulmonary, and endocrine functions generally disclosed small and inconsistent proportions of abnormalities between groups, and were deemed clinically unimportant. An unrefined assessment of all summed and weighted organ system abnormalities by group did not show an aggregation of multisystem disease or malfunction.

Any interpretation of these study data, in whole or in part, must carefully consider the methodical steps required for a proper inference of causality. It is specifically pointed out that many group differences were largely based upon subjective data, and that a subtle effect of differential reporting is suggested but has not been fully evaluated. For objective data, group differences were generally within normal ranges and were not correlated to the herbicide exposure index, nor fell within the expected latency periods following Vietnam service. The proposed clinical end points of dioxin exposure, chloracne, soft tissue sarcoma, and porphyria cutanea tarda, were not found in the Ranch Hand group (study power limitations recognized). Overall, substantial credence is given to the objective study findings, particularly after observing the consistent duplication of the classical effects of risk factors such as age, smoking, alcohol, etc., in almost all clinical areas. Additional work with these baseline data is still required in the areas of data base refinement, statistical testing and bias analysis, exposure index refinement, establishment of the follow-up examination requirements, and collaboration with other dioxin research studies.

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. The 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 their families at this time.

PREFACE

In October 1978, the United States Air Force (USAF) Surgeon General made the commitment to the Congress and to the White House to conduct an epidemiologic study of the possible adverse health effects arising from the herbicide exposure of Air Force personnel who conducted aerial dissemination missions in Vietnam (Operation Ranch Hand). The purpose of this epidemiologic investigation is to determine whether long-term adverse health effects exist, and whether they can be attributed to occupational exposure to herbicides and their contaminants. The study protocol for this effort incorporates a matched cohort design placed in a nonconcurrent prospective setting. The study approach includes mortality, morbidity, and follow-up elements linked tightly in time in order to produce the most data in the shortest time. The study addresses the question: Have there been, are there currently, or will there be any adverse health effects among former Ranch Hand personnel caused by repeated occupational exposure to 2,4,5-Trichlorophenoxyacetic acid (2,4,5,-T) containing herbicides and the contaminant, 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD)? At the request of the Principal Investigators (see Appendix I) the study protocol was extensively and independently reviewed. The review agencies included: The University of Texas School of Public Health, Houston TX; the USAF Scientific Advisory Board; the Armed Forces Epidemiological Board; and the National Research Council of the National Academy of Sciences. In 1980, the Science Panel of the Agent Orange Working Group was created as an additional peer review agency. This group, redesignated as the Advisory Committee on Special Studies Relating to the Possible Long-Term Health Effects of Phenoxy Herbicides and Contaminants, has consented to the oversight responsibility of the Ranch Hand study and continues to monitor the conduct of this epidemiologic investigation (see Appendix II).

The Air Force Health Study (Ranch Hand II) protocol emphasizes the suboptimal statistical power of the mortality study. The mortality study was motivated by the desire to use a full spectrum epidemiologic approach to the herbicide question. Additionally, the investigators were scientifically obliged to pursue the mortality study because of previous and emerging studies (some with small sample sizes) which suggested the possibility of a soft tissue sarcoma end point (Honchar, 1981; Hardell, 1979; Erikson, 1979). Within the inherent sample size limitation of the Ranch Hand population, detection of such a rare condition will be missed unless there is marked case clustering and correspondingly high relative risks.

Also, because of sample size limitations as well as the myriad of proposed clinical end points, a case-control design was not entertained. In the morbidity phase of the study, the investigators have attempted to enhance statistical power and analytic sensitivity where possible by using (1) precise matching procedures with a replacement strategy to maintain statistical power while averting a loss-to-study bias, (2) exacting quality control procedures, (3) mortality-morbidity linkages, (4) a lengthy follow-up study, (5) state-of-the-art statistical methodology, (6) continuously distributed physical examination variables, and (7) data collection focused on verifiable end points.

The mortality analyses have not revealed any adverse death experience in the herbicide/dioxin exposed cohort. The results of the analyses were consistent: at this time, there is no indication that Ranch Hand personnel have experienced any increased mortality or any unusual patterns of death in time or by cause. They are not dying in increased numbers, at earlier ages, or by unexpected causes.

The fact that only a relatively small number of Ranch Hand deaths were available for analysis is reassuring in itself. However, the fact that adverse effects have not yet been detected does not imply that an effect will not become manifest at a future time or after covariate-adjusted analyses. For this reason, further analyses are intended and mortality in the study population will be ascertained annually for the next 20 years.

The morbidity portion of the study was conducted in two phases; an in-home, face-to-face interview, and a comprehensive physical and psychological examination. Both phases were conducted by civilian organizations under contract to the Air Force, using materials and procedures prescribed by the contract. One thousand, one hundred seventy four (97%) of the Ranch Hand group and 1,156 (93%) of the initially selected comparison group participated in the questionnaire. An additional 376 comparison subjects were interviewed as replacement subjects, bringing the total number of comparison participants to 1,532. Two thousand, seven hundred eight current and former wives of the study participants were interviewed. One thousand forty five (87%) of the Ranch Hand group participated in the physical examination, and 936 (76%) of the initially selected comparison subjects participated. Two hundred eighty-eight replacement subjects also participated in the examination process, giving a total of 2,269 participants, resulting in 1,024 matched pairs for analysis.

The first chapter of this report is devoted to a discussion of the background of the study and the next seven chapters present a summary of the methodology used in gathering, analyzing, and interpreting the data. The results and discussion of these analyses, organized by organ system and/or disease end point, are contained in the remaining chapters.

This report assumes that readers are familiar with statistical and epidemiologic techniques. It also assumes that the reader has a familiarity with the herbicide/dioxin issue and a detailed knowledge of the protocol of the Air Force study, the baseline questionnaire, and the baseline mortality results. In the interest of brevity, the reader is referred to the protocol published as US Air Force School of Aerospace Medicine Technical Report 82-44, the baseline questionnaires published as US Air Force School Aerospace Medicine Technical Report 82-42, and the Baseline Mortality Study Results, 30 June 1983. These reports are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161.

ACKNOWLEDGMENTS

The Principal Investigators gratefully acknowledge the outstanding support given to this project by:

The Ranch Hand Association and its elected officers, for sustained encouragement of the study, assistance in population ascertainment, and camaraderie and patriotism which contributed to unparalleled participation rates.

Our peer review groups, the University of Texas, School of Public Health, the Air Force Scientific Advisory Board, the Armed Forces Epidemiological Board, the National Research Council, and our Advisory Committee for their scientific contributions which have facilitated the conduct of this study and enhanced public credibility.

The Air Force Research Management Structure at the Aerospace Medical Division, Brooks AFB TX; 3303d Contracting Squadron, Air Training Command, Randolph AFB TX; Office of the Command Surgeon, Air Force Systems Command, Andrews AFB MD; and the Air Force Surgeon General's Office, Bolling AFB DC, for program advocacy and provision of resources.

The over 100 professionals, consultants, technicians, military and civilian, whose dedication and hard work over the past five years have made this report possible.