

TABLE 12.

**Overall Summary Results of Unadjusted and Adjusted
Group Contrast Analyses of Renal Variables**

Variable	Type of Analysis	Unadjusted	Adjusted
<u>Questionnaire</u>			
History of Kidney Disease/Stones	D	NS	NS
<u>Laboratory</u>			
Urinary Protein	D	NS	** (NS)
Urinary Occult Blood	D	NS	** (NS)
Urinary White Blood Cell Count	D	NS	NS
Blood Urea Nitrogen	C	NS	** (NS)
Urine Specific Gravity	C	NS	NS

D: Discrete analysis performed.

C: Continuous analysis performed.

NS: Not significant ($p > 0.10$).

** (NS): Group-by-covariate interaction ($0.01 \leq p \leq 0.05$); not significant when interaction is deleted; refer to Table N-2 in the main report for a detailed description of this interaction.

The current renal function was evaluated by five laboratory variables: urine protein, urinary occult blood, urinary white blood cell count, blood urea nitrogen, and urine specific gravity.

There was no significant difference detected between the two groups based on the unadjusted analysis of urinary protein. In the adjusted analysis, there was a significant interaction between group and occupation. Stratifying by occupation revealed that the Comparison enlisted flyers had a higher percentage of abnormalities than the Ranch Hand enlisted flyers ($p = 0.036$). After deleting the group-by-occupation interaction, no difference between the two groups was observed. This result differed from the twofold increase of proteinuria observed in Comparisons at Baseline.

No difference was identified between the Ranch Hands and the Comparisons based on the analysis of urinary occult blood without adjustments for covariates. However, after stratifying by race due to a significant group-by-race interaction, the estimated prevalence rate for the Black Ranch Hands was noted as being statistically higher than the corresponding rate for the Black Comparisons ($p=0.013$). The estimated prevalence rates were not detected as being different based on an adjusted model without the group-by-race interaction.

Based on the analyses of urinary white blood cell count, no differences were detected between the two groups in either the unadjusted or adjusted analyses.

The mean blood urea nitrogen levels of the Ranch Hands and Comparisons did not vary significantly when compared without adjustments. The adjusted analysis detected a significant group-by-race interaction. Stratifying by race revealed that the mean of the Black Comparisons was statistically higher than the mean of the Black Ranch Hands ($p=0.022$). The adjusted means were also not significantly different when estimated without the group-by-race interaction in the model.

There was no evidence that the mean urine specific gravity was different between the Ranch Hands and Comparisons in either the unadjusted or adjusted analysis.

The exposure index analyses showed very little evidence of a dose-response relationship at the 1987 followup examination. No pattern in the relationship of abnormality rates or mean levels was seen within occupational cohort.

The longitudinal analysis was based solely upon the contrast of blood urea nitrogen levels between the 1982 and 1987 examinations. The unadjusted mean levels increased slightly from 1982 to 1987, but the change between the Ranch Hands and Comparisons over time was not significantly different.

In conclusion, none of the six variables of the renal assessment showed a significant difference based on the unadjusted analyses. For three of the variables, the adjusted results supported the findings of the unadjusted analyses; there were significant group-by-covariate interactions for the other variables. Further examination by strata revealed that in one case the Ranch Hand prevalence rate was higher than the Comparison rate and that the opposite relationship existed for another case. In the third instance, the Comparison mean was higher than the mean of the Ranch Hands; however, both means were within the normal range. The adjusted analyses without the group-by-covariate interactions supported the findings of the unadjusted analyses; the renal status of the Ranch Hands and Comparisons was generally similar.

ENDOCRINE ASSESSMENT

INTRODUCTION

The human endocrine system is not considered to be a major target of chlorophenol or 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) exposure. This is not so in animals, however. A wide range of endocrine abnormalities in many species has been induced experimentally by TCDD and includes hypoglycemia, hypothyroxinemia, reduced progesterone levels, and increased testosterone levels, the latter presumably reflecting decreased liver catabolism due to parenchymal liver damage or an inhibition of the cytochrome P-450 system.

DISCUSSION

The historical and laboratory data collected in the endocrine assessment provide a valid reflection of thyroid, gonadal, and pancreatic functions by indices that are well established in clinical practice.

As would be expected, comparison of the current historical data with those of the 1985 followup revealed an increase over time in the incidence of thyroid disease, with similar trends in both the Ranch Hand (n=55 in 1985 vs. n=93 in 1987) and Comparison (n=78 in 1985 vs. n=113 in 1987) groups.

An increase in the presence of thyroid nodules as a result of advancing age is well documented in autopsy and ultrasound studies. However, a decrease was detected in thyroid abnormalities in the 1987 followup (n=592 abnormal) versus the 1985 followup (n=773 abnormal). Prior to the 1987 examination, an attempt was made to reduce interobserver variability among the examining physicians by employing more uniform techniques of palpation and by defining more objective endpoints for palpable abnormalities. Comparison of the data revealed virtually identical trends in both the Ranch Hands (342 abnormal in 1985 vs. 258 abnormal in 1987; 34% vs. 26% incidence) and Comparisons (431 abnormal in 1985 vs. 334 abnormal in 1987; 33% vs. 26% incidence).

Though not reported in the endocrine assessment, several additional physical findings beyond simple palpation are recognized as relevant to the clinical evaluation of thyroid and gonadal function. Body habitus, ocular and integumentary signs, and deep tendon reflexes are among the variables that were included in the general health, neurological, and dermatologic examinations, and are reported in those sections respectively.

Of the two laboratory variables used, the triiodothyronine (T_3) uptake, though far less sensitive than the serum thyroid stimulating hormone (TSH), assumes importance as the only index common to all three physical examination cycles. In the current study, as in the Baseline and 1985 followup, no significant differences were detected between the Ranch Hand and Comparison groups. Further, the few covariate associations defined fail to document any consistent clinical trends over time.

In lacking a lower limit of normal, the radioimmunoassay (RIA) technique of serum TSH determination is not sensitive to hyperthyroid states. Nonetheless, the current 1987 followup data can be validly compared with those from

the 1985 followup, which exhibited significant differences between Ranch Hands and Comparisons for both the unadjusted and adjusted analyses. For the 1987 followup, the Ranch Hand group was found to have TSH levels that were marginally higher than the TSH levels of the Comparisons in both the unadjusted and adjusted analyses (unadjusted: 1.01 2IU/ml vs. 0.97 2IU/ml; adjusted: 0.96 2IU/ml vs. 0.93 2IU/ml).

With respect to gonadal function, no significant group differences were found, and two established clinical correlations were confirmed. With advancing age, a gradual decline in serum testosterone levels would be expected and was evident only in those participants born in or before 1922.

The correlation between testosterone levels and obesity is less well defined. While extremes of obesity (i.e., in excess of 100% of ideal body weight) are usually associated with gonadal suppression, no consistent relationship has been defined between serum testosterone and percent body fat. Further, the apparent differences in serum testosterone levels may in fact reflect changes in sex hormone binding globulin rather than the biologically active-free fraction. Finally, the finding of slightly lower testosterone levels in Type B individuals is of doubtful clinical significance but consistent with the increased frequency of endomorphic body habitus in this personality type. The earlier examinations in this series found that the Ranch Hands had higher levels of testosterone than did the Comparisons, a difference that is no longer evident.

An expected incidence of overt diabetes mellitus and of glucose intolerance was documented in the current study with no significant group differences defined. In ambulatory medicine, the 2-hour postprandial blood sugar has replaced the traditional glucose tolerance test as a screen for diabetes. Consistent with the insulin resistance that occurs in Type II diabetes, strong covariate associations were defined relating glucose intolerance to age and percent body fat. Independent of weight, a 10-15 percent incidence of glucose intolerance will occur by age 70. For each decade over age 40, there is an increase in the 2-hour postprandial blood sugar of 10-15 mg percent, and an average increase of 5.0 mg percent per decade in the fasting blood sugar.

The results of the endocrine assessment confirmed numerous associations that would be expected in clinical practice, and no significant or clinically relevant group differences were found.

SUMMARY

Table 13 summarizes the results of Ranch Hand and Comparison group contrasts for the 10 variables examined in 1987 to assess the endocrine system.

Two variables were constructed from the review-of-systems questionnaire and the health interval questionnaire to determine the thyroid status for each participant. No significant group difference was noted for both the self-administered response to current thyroid function and the interviewer-administered response to history of thyroid disease.

TABLE 13

Overall Summary Results of Unadjusted and Adjusted
Group Contrast Analyses of Endocrine Variables

Variable	Unadjusted		Adjusted		Direction of Results
	Discrete	Continuous	Discrete	Continuous	
<u>Questionnaire</u>					
Current Thyroid Function (Self- Administered)	NS	--	--	--	
History of Thyroid Disease (Interviewer Administered)	NS	--	--	--	
<u>Physical Examination</u>					
Thyroid Gland	NS	--	--	--	
Testes	NS	--	--	--	
<u>Laboratory</u>					
T ₄ % Uptake	NS	NS	NS	NS	
TSH	NS	NS*	NS	NS*	RH>C
FSH	NS	NS	NS	NS	
Testosterone	NS	NS	****	NS	
2-Hour Postprandial Glucose	NS	NS	NS	NS	
Composite Diabetes Indicator	NS	--	** (NS)	--	

NS: Not significant ($p > 0.10$).

--Analysis not performed or not applicable.

NS*: Borderline significant ($0.05 < p \leq 0.10$).

RH>C: Higher prevalence rate or mean in Ranch Hands than in Comparisons.

****: Group-by-covariate interaction ($p \leq 0.01$).

** (NS): Group-by-covariate interaction ($0.01 < p \leq 0.05$); not significant when interaction is deleted; refer to Table 0-2 in the main report for a detailed description of this interaction.

The thyroid gland and the testes were evaluated at the physical examination. The percentage of abnormalities did not differ significantly between groups for either organ.

Six laboratory examination variables were analyzed to assess current endocrine function: T_3 uptake, TSH, follicle stimulating hormone (FSH), testosterone, 2-hour postprandial glucose, and a composite diabetes indicator. Each variable was analyzed in continuous and discrete forms, except for the composite diabetes indicator, which was only analyzed discretely.

No significant unadjusted group differences were found for any of these variables. However, the Ranch Hand TSH mean was marginally significantly higher than the Comparison mean ($p=0.099$). A statistically significant TSH difference was noted in the 1985 followup. The only change in findings after adjustment for significant covariates was due to the presence of four group-by-covariate interactions for testosterone discretized. Initial stratification by occupation revealed no significant group differences for the officer and enlisted flyer cohorts. Further stratification by personality type and age for the enlisted groundcrew cohort detected no significant strata, but results from this analysis were limited due to sparse data in many cells (in several strata the abnormally low testosterone values were either all Ranch Hands or all Comparisons). Although no significant group differences were found for the laboratory test variables, the direction of the unadjusted results showed that Ranch Hands consistently had more abnormalities than Comparisons. These trends are discussed in the interpretive considerations section.

Results from the exposure index analyses generally did not support a herbicide effect. For T_3 uptake, TSH, FSH, and 2-hour postprandial glucose, there were no statistically significant findings. Unadjusted testosterone means differed significantly for the officer cohort, exhibiting a pattern consistent with a decreasing dose-response relationship; after covariate adjustment, this difference became nonsignificant. Adjusted results were significant for the enlisted flyer cohort, but did not indicate a dose-response effect since the highest levels were found in the medium exposed group. Testosterone results for the enlisted groundcrew were not significant. The enlisted groundcrew and officer cohorts showed increasing dose-response patterns for diabetes, but the association was not significant. In contrast, a significant result ($p=0.010$) was found for the enlisted flyer cohort but was due to most diabetics falling in the medium exposure category.

Longitudinal analyses for T_3 uptake, TSH, and testosterone showed no significant group differences from the Baseline to the 1987 followup examination.

In conclusion, statistical analysis of the 10 endocrinologic variables did not reveal any significant group differences. The Ranch Hand TSH mean was marginally significantly higher than the Comparison mean; at the 1985 examination, a significant difference was found. Means for the other variables were very similar between groups. For all laboratory examination variables, the percentage of abnormalities was higher for the Ranch Hand group than for the Comparison group, but not statistically significant. The significant differences in testosterone and 2-hour postprandial glucose found in the 1985 examination were no longer evident.

IMMUNOLOGIC EVALUATION

INTRODUCTION

Overt damage to organs of the immune system and depressed immunologic function have been noted in a variety of animals exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). As the fields of immunology and immunotoxicology have grown within the past 10 years, a significant spectrum of subtle immunotoxic effects has also been described in animals, but for many possible reasons, comparable adverse effects have not been consistently recorded in exposed human individuals or cohorts.

Thus, an intensive search is under way to ascertain the effects of TCDD on the human immune system, particularly with respect to the development of cancer. Most ongoing dioxin morbidity studies in the United States have incorporated comprehensive laboratory assessments of the immune system.

DISCUSSION

Immunologic competence was assessed by analysis of data from cell surface marker studies, immunoglobulin quantitation, functional stimulation assays, and skin tests for delayed hypersensitivity response on a randomized subset of the study population. The tuberculin skin test is the prototype test for DCH. This test has been used throughout the 20th century as the traditional method of diagnosing infection with Mycobacterium tuberculosis in individual patients, contacts of diseased individuals, occupational groups, and epidemiologic studies of populations.

The absence of a response to a series of skin test antigens is usually indicative of an impaired immune defense mechanism (anergy). Anergy can occur in elderly individuals in the setting of certain viral, bacterial, and fungal infections; or with advanced protein deficiency, underlying malignancy, or treatment with corticosteroids and other immunosuppressive agents. Skin tests for DCH are occasionally used to test for anergy as a prognostic indicator in individuals in compromised states such as the acquired immunodeficiency syndrome or those at risk of infection following surgery.

Skin tests for DCH are subject to numerous variables including the dose and method of administration of the antigen and the techniques employed in reading and interpreting the response. Following quality control concerns over the 1985 skin test data, stringent protocols were established to ensure consistent methods and interpretation. In the current study, a premium was placed on uniform and consistent methods and interpretation. There was a 92 percent concordance between readers and duplicate interpretations by the same reader. More than 99.6 percent of the sample population had interpretable skin tests. The 94.9 percent incidence of intact DCH is consistent with clinical experience in the general population. Analysis of the data suggested interactive effects of cigarette and alcohol use. Clarification of the observed group difference in the composite skin test diagnosis must await the analysis of the quantitative serum dioxin results.

Cell surface marker studies for CD2 (total T cell), CD4 (helper T cell), CD8 (suppressor T cell), CD25 (activated T cells), CD20 (total B cell), CD14 (monocytes), and HLA-DR positive cell populations were analyzed. The CD4/CD8 ratio was calculated and also analyzed. Both the unadjusted and adjusted analyses of the various cell surface markers measured did not indicate significant group differences between Ranch Hands and Comparisons. Significant covariate associations with age were found for CD2, CD4, CD8, CD20, and HLA-DR cells. These variables consistently decreased with increasing age, which is consistent with established clinical findings. Statistically significant race and alcohol associations were found for CD20 and CD14. Overall, cell surface marker counts increased with cigarette smoking. The clinical significance of these findings is unknown.

Functional stimulation assay data analyzed include the unstimulated and stimulated responses for both the PHA and MLC assays. No significant unadjusted or adjusted group differences between Ranch Hands and Comparisons were found for either the PHA or MLC assays. Both PHA and MLC responses appeared to decrease with age. Race appeared to affect PHA response, but biologic significance was difficult to evaluate given the lack of established clinical endpoints associated with these differences and the lack of a consensus as to what the normal range is for these assays. Implications of mild to moderate increases and decreases are not known. The ability to respond to a challenge with increased cell counts and functional reactions is desirable but a hyperactive response may not be desirable since it might indicate a constantly challenged immune system.

Other functional stimulation assay data evaluated included the net responses for the natural killer cell assays (with and without the addition of Interleukin 2 as a response stimulator). Unadjusted analyses for both natural killer cell assays revealed no significant Ranch Hand and Comparison differences; however, there was a significant group-by-race interaction for both assays. When analyzing the data within each racial grouping, there was a statistically significant difference between Black Ranch Hands and Black Comparisons.

The adjusted group contrast analysis for the four natural killer cell variables and the MLC net response variable each contained group-by-race interactions. The clinical significance of these findings is not apparent.

The exposure index analyses failed to reveal any consistent trends in the many variables analyzed. For the adjusted analyses, many exposure index-by-covariate interactions were found. These interactions primarily involved the covariates of cigarette smoking, age, and alcohol use. Final interpretation of these data must await the results of the serum TCDD assays and the development of interpretive criteria for these immunologic assays.

As seen in the 1985 followup, there were no significant group differences for either the unadjusted or adjusted analyses of any of the laboratory immunologic variables examined. Consistently decreasing values for the cell surface markers and functional stimulation assays were associated with increasing age, while increases in lifetime smoking were usually associated with increases in the values of those variables. Longitudinal analysis of the CD4/CD8 ratio results did not reveal a significant group difference over time.

The immunologic assessment of laboratory data revealed no statistically significant differences between the Ranch Hand and Comparison populations. Covariate associations with age and lifetime smoking were noted in the adjusted analyses of these immunologic tests. The finding of a group difference in the proportion of participants possibly abnormal on the composite skin test diagnosis is of interest and will be reevaluated in the context of quantitative serum dioxin levels. Overall, there appears to be no indication of impaired immunologic competence in the Ranch Hand group versus the Comparison group over time.

SUMMARY

For the 1987 followup immunologic assessment, a number of unadjusted and adjusted analyses were performed using physical examination (composite skin test diagnosis) and laboratory examination data (cell surface marker studies, TLC, quantitative immunoglobulin measurements, and functional stimulation tests). The results of the Ranch Hand and Comparison group contrasts performed using the physical examination and laboratory examination data are summarized in Table 14.

For the composite skin test diagnosis, the unadjusted group contrast of the relative frequency of participants with possibly abnormal composite readings was significantly greater ($p=0.019$) for the Ranch Hands than the Comparisons. The adjusted model for the composite skin test results contained a significant group-by-lifetime cigarette smoking history interaction. Because of this interaction, the skin test results were analyzed for group differences through stratification of lifetime cigarette smoking history. Ranch Hands who smoked for over 10 pack-years had a significantly greater frequency of individuals with possibly abnormal skin test results than Comparisons with the same lifetime cigarette smoking history ($p=0.005$). Without the cited interaction, a significant adjusted group difference ($p=0.011$) remained.

For the cell surface marker studies of the 1987 followup, there were no significant group differences for either the unadjusted or the adjusted analyses. Except for CD25, the same cell surface marker variables were analyzed in both the 1985 and the 1987 followup studies. The 1985 followup unadjusted analyses for group differences were not significant. The 1985 followup adjusted analyses were not significant for CD4, CD8, and the CD4/CD8 ratio; the remaining 1985 followup cell surface marker variables had significant group-by-covariate interactions in the adjusted models.

Unadjusted and adjusted group contrasts were not significant for TLC.

For each of the quantitative immunoglobulins (IgG, IgA, and IgM), the unadjusted and adjusted group contrasts were not significant.

For the functional stimulation tests of the 1987 followup study, unadjusted and adjusted analyses were performed on a number of measures pertaining to responses after mitogen stimulation with phytohemagglutinin (PHA), mixed lymphocyte culture (MLC) responses to stimulation from donor lymphocytes, and natural killer cell assay (NKCA) and natural killer cell assay with Interleukin 2 (NKCI).

TABLE 14.

**Overall Summary Results of Unadjusted and
Adjusted Analyses of Immunologic Variables**

Variable	Type of Analysis	Unadjusted	Adjusted	Direction of Results
Physical Examination				
Composite Skin Test Diagnosis	D	0.019	** (0.011)	RH>C
Laboratory Examination: Quantitative Studies				
CD2 Cells	C	NS	NS	
CD4 Cells	C	NS	NS	
CD8 Cells	C	NS	NS	
CD20 Cells	C	NS	NS	
CD14 Cells	C	NS	NS	
CD25 Cells	D	NS	--	
	C	NS	NS	
HLA-DR Cells	C	NS	NS	
CD4/CD8 Ratio	C	NS	NS	
TLC	C	NS	NS	
IgG	C	NS	NS	
IgA	C	NS	NS	
IgM	C	NS	NS	
Laboratory Examination: Functional Stimulation Tests				
Unstimulated PHA Response	C	NS	NS	
PHA Net Response: Day 1			** (NS)	
Concentration 1	C	NS	NS	
Concentration 2	C	NS	NS	
Concentration 3	C	NS	NS	
PHA Net Response: Day 2				
Concentration 1	C	NS	NS	
Concentration 2	C	NS	NS	
Concentration 3	C	NS	NS	
Overall PHA Net Response	C	NS	NS	
Maximum PHA Net Response	C	NS	NS	
Unstimulated MLC Response	C	NS	NS	
MLC Net Response	C	NS	** (NS)	
NKCA 50/1 Net Response	C	NS	** (NS)	
NKCA 50/1 Percent Release	C	NS	** (NS)	
NKCI 50/1 Net Response	C	NS	****	
NKCI 50/1 Percent Release	C	NS	****	

TABLE 14. (continued)

Overall Summary Results of Unadjusted and
Adjusted Analyses of Immunologic Variables

-
- D: Discrete analysis performed.
** (0.011): Group-by-covariate interaction ($0.01 < p < 0.05$); significant
($p=0.011$) when interaction is deleted.
RH>C: More abnormalities in Ranch Hands than in Comparisons.
C: Continuous analysis performed.
NS: Not significant ($p > 0.10$).
--Analysis not done.
** (NS): Group-by-covariate interaction ($0.01 < p < 0.05$); not significant when
interaction is deleted; refer to Table P-3 in the main report for a
detailed description of this interaction.
****: Group-by-covariate interaction ($p \leq 0.01$); refer to Table P-3 in the main
report for a detailed description of this interaction.
-

For the PHA responses, the group contrasts were performed for each of the following: unstimulated PHA responses for 2 harvest days concurrently; net responses to PHA at each of three concentrations on two different days; all PHA net responses concurrently for the six concentration and day combinations; and the maximum of the six PHA net responses.

For the 1987 followup, as in 1985, the unadjusted and adjusted group contrasts of the unstimulated PHA responses were not significant.

For the PHA net response for day 1, the unadjusted group contrast at each of the three concentration levels was not significant. The adjusted group contrasts of the PHA net response for day 1 at concentration levels 2 and 3 were also not significant. However, the adjusted analysis of the PHA net response for day 1 at concentration level 1 had a significant group-by-current alcohol use interaction. For participants having over four drinks per day, Comparisons had a significantly greater net response to PHA for day 1 at concentration level 1 than Ranch Hands ($p=0.024$). For the PHA net response for day 2 at each of three concentration levels, the unadjusted and adjusted group contrasts were not significant. For the 1985 followup data, both the unadjusted and the adjusted group contrasts of the PHA net response did not exhibit significant group differences.

The unadjusted and adjusted simultaneous contrast of the six PHA net responses was not significant. The unadjusted and adjusted analyses of the maximum PHA net responses were not significant for the Ranch Hand versus Comparison group contrasts.

For the unstimulated MLC response, both the unadjusted and the adjusted group contrasts were not significant. For the MLC net response, the unadjusted group contrast was not significant and the adjusted analysis had a significant group-by-race interaction. Because of this interaction, group contrasts were performed within race strata. Among Blacks, the Ranch Hands

had a marginally significantly lower average MLC net response than the Comparisons ($p=0.059$). An interaction with smoking history was seen in 1985.

For the NKCA and NKCI, 50/1 net responses and 50/1 percent releases were analyzed. In the Ranch Hand and Comparison group contrasts, the unadjusted analyses were not significant. For each of the adjusted analyses of the NKCA and NKCI variables, there was a significant group-by-race interaction. Because of these interactions, the NKCA 50/1 net responses and the 50/1 percent releases were analyzed within race strata. Black Ranch Hands had a borderline significantly greater average net response than Black Comparisons ($p=0.065$), and Black Ranch Hands had a significantly higher average percent release than their Comparisons ($p=0.031$). Deleting these interactions yielded nonsignificant group contrasts. For the NKCI assay, the group contrasts were also examined by race because of the significant group-by-race interaction. Black Ranch Hands had a significantly greater mean net response for NKCI than did the Black Comparisons ($p=0.007$). Black Ranch Hands had a significantly greater average percent release of NKCI than Black Comparisons ($p=0.008$), and nonblack Ranch Hands had a marginally significant lower average than nonblack Comparisons ($p=0.069$).

The unadjusted exposure index analysis of the composite skin test diagnosis was not significant for the enlisted flyers and for the enlisted groundcrew, and it was borderline significant ($p=0.090$) for the officers. For the adjusted exposure index analysis, officers had a significant exposure index-by-lifetime cigarette smoking history interaction and a significant exposure index-by-current alcohol use interaction. For enlisted flyers, there was a significant exposure index-by-lifetime alcohol history interaction. For enlisted groundcrew, there was a significant exposure index-by-lifetime alcohol history interaction and a significant exposure index-by-current alcohol use interaction.

For the exposure index analysis of the cell surface marker measures, the unadjusted analysis generally showed no significant difference for each occupation. For the adjusted exposure index analyses of an individual cell surface marker variable, an exposure index-by-covariate interaction was generally found for at least one occupation. For the most part, the interactions involved the covariates of age, lifetime cigarette smoking history, current alcohol use, or lifetime alcohol history.

The unadjusted and adjusted exposure index analyses of TLC were not significant for officers and enlisted flyers. For the enlisted groundcrew, the unadjusted exposure index analysis was not significant, and the adjusted analysis contained a significant exposure index-by-lifetime cigarette smoking history interaction.

In general, the unadjusted exposure index analyses of the immunoglobulins were not significant for each occupation. For officers, the adjusted exposure index analysis of IgG was significant ($p=0.032$). For enlisted groundcrew, there was a significant exposure index-by-lifetime cigarette smoking history interaction for IgG. For officers and enlisted groundcrew, the adjusted exposure index analyses of IgA had significant exposure index-by-current cigarette smoking and exposure index-by-lifetime alcohol history interactions, respectively. The adjusted exposure index analyses of IgM were not significant.

For the exposure index analysis of the unstimulated PHA responses, the unadjusted and adjusted analyses for officers and for enlisted flyers were not significant. For enlisted groundcrew, the unadjusted exposure index analysis was not significant and the adjusted analysis contained significant interactions between the exposure index and both alcohol use covariates. For the PHA net responses for day 1 at each of three different concentration levels, the unadjusted and adjusted exposure index analyses were generally not significant for the three occupations. The exceptions occurred for enlisted flyers at concentration level 2 on the adjusted analysis ($p=0.053$), and for enlisted flyers at concentration level 3 on the unadjusted and the adjusted analyses ($p=0.067$ and $p=0.056$, respectively). For the PHA net responses for day 2 at each of three concentration levels, the unadjusted analyses were not significant for the three occupations. For the adjusted exposure index analyses of the PHA net responses for day 2, a significant exposure index-by-age interaction was found for the enlisted groundcrew at concentration level 1 and a significant exposure index-by-current cigarette smoking interaction was found for the officers at concentration level 3. For the simultaneous analysis of the six PHA net responses, neither the unadjusted nor the adjusted analysis was significant for each occupation. Similarly, neither the unadjusted nor the adjusted exposure index analysis of the maximum PHA net response was significant for each occupation.

The unadjusted exposure index analyses of the unstimulated MLC responses were not significant for each occupation. For the adjusted exposure index analysis of the unstimulated MLC responses, the enlisted flyers had a significant exposure index-by-age interaction, and the officers and the enlisted groundcrew displayed no significant difference for exposure index. For the MLC net responses, both the unadjusted and the adjusted exposure index analyses were not significant for each occupation.

The unadjusted exposure index analyses of the NKCA and NKCI net responses and percent releases were not significant for each occupation. For the exposure index adjusted analysis of the NKCA net response, the enlisted flyers had a significant exposure index-by-lifetime cigarette smoking history interaction. For the exposure index adjusted analyses of the NKCA and the NKCI percent release, the enlisted groundcrew had significant exposure index-by-age interactions. Overall, the exploration of covariate interactions in the exposure index analyses detected scattered increases and decreases in cell count and functional assays that are impossible to interpret in the absence of a consensus as to what is abnormal for these measures of immunity.

The longitudinal analysis of the CD4/CD8 ratio results for the 1985 and 1987 followup examinations did not exhibit a significant group difference over time.

The immunologic assessment of laboratory data revealed no statistically significant differences between the Ranch Hands and Comparisons. The finding of a group difference in the proportion of participants possibly abnormal on the composite skin test diagnosis is of interest and will be reevaluated in the context of the quantitative serum dioxin levels. Overall, there appears to be no indication of clinically relevant impaired immunologic competence in the Ranch Hand group versus the Comparison group over time.

PULMONARY DISEASE

INTRODUCTION

Pulmonary dysfunction and overt pulmonary disease are not recognized clinical entities resulting from exposure to chlorophenols or 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD).

Acute exposure to chlorophenols, phenoxy herbicides, and TCDD have caused the traditional acute symptoms of cough, nasal/lung irritation, shortness of breath, and, occasionally, bronchitis. These symptoms have been noted almost exclusively in industrial workers and not in individuals experiencing casual contact. Long-term sequelae arising from the acute symptom stage in ill individuals have not been generally known because of minimal followup and surveillance of the pulmonary symptoms.

Further, due to the profound effect of smoking on pulmonary function, great emphasis must be placed on the collection of highly accurate, detailed, and validated smoking data as an adjustment variable.

DISCUSSION

While the presence of pulmonary disease is often evident based on a careful history and physical examination, definitive diagnosis usually requires the collection of data from a number of other sources. The standard radiographic examination of the chest and pulmonary function studies are routinely ordered and were included as variables in the Air Force Health Study examination. In addition, because the lung is often involved secondarily in numerous infectious, inflammatory, and neoplastic disorders, the assessment of pulmonary disease should include the type of comprehensive multisystem review conducted in this examination cycle and reported in other sections.

Historical information on the occurrence of pulmonary disease must be interpreted with caution in the absence of medical record verification. Many of the cardinal symptoms of lung disease, including dyspnea, chest pain, and exercise intolerance, are common to cardiovascular disease as well (particularly ischemic heart disease) and are frequently misinterpreted as to cause. Wheezing, assumed by the patient to be indicative of asthma, may in fact be reflective of hemodynamic compromise in congestive heart failure. A positive purified protein derivative skin test, indicative of subclinical tuberculous infection, may be erroneously interpreted and reported as prior active infection. "Pneumonia" and "pneumonitis" are often confused by patients in relating the medical history.

The physical examination variables studied can provide valuable clues to the presence of pulmonary disease. In lacking specificity, however, these data are often of limited utility in confirming a specific diagnosis. Wheezes and hyperresonance, for example, will occur in obstructive airway disease in

asthma or in emphysema secondary to cigarette use. Dullness to percussion, a finding common to many disorders, will occur in consolidation from atelectasis, infections, pleural thickening, or pleural effusion.

In view of the limitations of the history and physical examination noted above, added emphasis is placed on screening laboratory data in the diagnosis of respiratory disease. The chest x ray, when normal, is highly reliable in excluding pulmonary parenchymal disease, though several exceptions must be recognized. Solitary lesions less than 6 millimeters, miliary granulomatous infection, and early interstitial disease, among others, may be present but not detectable radiographically. On the other hand, the chest x ray may reveal an early occult malignancy in an asymptomatic patient and thus afford an opportunity for cure.

Spirometry has been used as a clinical tool to measure static lung volumes and to detect respiratory disease for over a century. Dynamic indices, relating changes in lung volume to time, were first developed over 50 years ago and, with computerization, have been refined to a high degree of accuracy and reproducibility. To be valid, spirometry requires that particular attention be paid to technician training and, with proper coaching, to eliciting the full cooperation of the patient. In any longitudinal study emphasis must be placed on the use of identical techniques to ensure comparability of data.

In broad terms, the spirometric indices evaluated in this chapter are designed to measure lung volume (vital capacity) and respiratory air flow (FEF). Static lung volume is principally determined by height and is independent of weight, while dynamic volume measurements depend in part on physical strength. Accordingly, all indices require correction for age and gender. Further, as confirmed in the present study, normal values for whites cannot be applied to other ethnic groups.

In clinical practice, respiratory disease can be divided into two broad categories. "Restrictive" disease is characterized by reduced vital capacity as seen in interstitial fibrosis or reduced lung volume after surgical resection. In "obstructive" airways disease, usually emphysema associated with cigarette use, there is abnormal prolongation of the flow-dependent indices [forced expiratory volume in one second (FEV1), forced expiratory volume in two seconds (FEV2), forced expiratory volume in three seconds (FEV3), and maximum forced expiratory flow (FEFmax)].

With few exceptions, the dependent variable-covariate associations found in the statistical analyses, confirm observations that are well established in clinical practice. With advancing age, an increased incidence of respiratory disease would be expected and was confirmed by history, on physical examination, and in the laboratory. The age-related decline in vital capacity is considered "physiologic" over time and will occur independent of acquired pulmonary disease.

The cause of the increased incidence of bronchitis and pneumonia in non-blacks is uncertain and cannot be explained on the basis of any previously established genetic or ethnic susceptibility. Differential access and use of medical care may play a role. In contrast, Blacks were found to be at detriment by all spirometric indices.

In the exposure index analyses, the ratio of FEV1 to forced vital capacity (FVC) revealed similar trends in the enlisted flyer and enlisted groundcrew cohorts. Although the data may reflect some herbicide-related health detriment, two confounding variables must be taken into consideration. As an index reflective of obstructive airways disease, the FEV1 will diminish with increased cigarette smoking over time. Secondly, as an effort-dependent index, the FVC is subject to performance bias and requires a fully compliant participant in order to be valid. Even in those studies considered technically adequate, the self-perception of prior herbicide exposure could introduce subtle bias sufficient to affect the results. It will be important to reexamine the FEV1/FVC ratio data when the body burden of herbicide can be defined more objectively by serum levels.

As expected, current and lifetime cigarette use were associated with significant abnormalities in all variables examined. Enlisted participants, with greater lifetime and current cigarette exposure, were at detriment relative to officers.

Finally, though limited to a single variable, the longitudinal analysis revealed no significant difference in the Ranch Hands versus the Comparisons. These observations will be greatly strengthened by longitudinal analysis of the spirometric variables in future examination cycles.

Data collected in the pulmonary assessment provide a valid reflection of lung function in the population under study. There was a similar incidence of respiratory disease and similar respiratory function in the Ranch Hand and Comparison groups.

SUMMARY

The 1987 pulmonary assessment was based on five questionnaire variables, seven variables from the physical examination, and eight laboratory variables. The results of the Ranch Hand and Comparison contrasts are summarized in Table 15.

The five questionnaire variables were based on self-reported data for the occurrence of the following conditions: asthma, bronchitis, pleurisy, pneumonia, and tuberculosis. There were no differences identified between the Ranch Hands and the Comparisons based on the unadjusted analyses. The results of the adjusted analyses supported this finding for asthma, bronchitis, and pleurisy. Due to the low number of participants reporting tuberculosis, no adjusted analysis was conducted. In the adjusted analysis of pneumonia, there was a significant interaction between group and lifetime cigarette smoking history ($p=0.004$). Stratifying by the covariate showed that a significantly higher percentage of Comparisons in the heavy cigarette smoking category reported pneumonia than heavy smoking Ranch Hands ($p=0.005$).

TABLE 15.

**Overall Summary Results of Unadjusted and Adjusted
Group Contrast Analyses of Pulmonary Variables**

Variable	Type of Analysis	Unadjusted	Adjusted	Direction of Results
<u>Questionnaire</u>				
Asthma	D	NS	NS	
Bronchitis	D	NS	NS	
Pleurisy	D	NS	NS	
Pneumonia	D	NS	****	
Tuberculosis	D	NS	--	
<u>Physical Examination</u>				
Thorax and Lung Abnormalities	D	0.020	NS*	RH>C
Asymmetric Expansion	D	NS	--	
Hyperresonance	D	NS*	** (NS)	RH>C
Dullness	D	NS	--	
Wheezes	D	NS	NS	
Rales	D	NS	NS	
X-Ray Interpretation	D	NS	** (NS)	
<u>Laboratory</u>				
FVC	C	NS	NS	
FEV ₁	C	NS	** (NS)	
FEV ₂	C	NS	** (NS)	
FEV ₃	C	NS	NS	
FEFmax	C	NS	NS	
Ratio of Observed FEV ₁ to Observed FVC	C	NS	NS	
Loss of Vital Capacity	D	NS	NS	
Obstructive Abnormality	D	NS	NS	

D: Discrete analysis performed.

NS: Not significant ($p > 0.10$).

****: Group-by-covariate interaction ($p \leq 0.01$); refer to Table Q-2 in the main report for a detailed description of this interaction.

--: Analysis not done.

NS*: Borderline significant ($0.05 < p \leq 0.10$).

RH>C: Higher prevalence rate in Ranch Hands than in Comparisons.

** (NS): Group-by-covariate interaction ($0.01 < p \leq 0.05$); not significant when interaction is deleted; refer to Table Q-2 in the main report for a detailed description of this interaction.

C: Continuous analysis performed.

The physical examination variables of the pulmonary assessment were: thorax and lung abnormalities, asymmetric expansion, hyperresonance, dullness, wheezes, rales, and x-ray interpretation. The Ranch Hands had significantly more thorax and lung abnormalities than the Comparisons based on the unadjusted analysis ($p=0.020$). After adjusting for age, occupation, current cigarette smoking, and lifetime cigarette smoking history, the difference was borderline significant ($p=0.072$).

There was only one participant--a Comparison--with asymmetric expansion. No significant difference was detected in the unadjusted analysis, and due to the low number of participants with this condition, no adjusted analysis was conducted.

The unadjusted analysis of hyperresonance showed a borderline difference between the two groups with a higher prevalence rate among the Ranch Hands ($p=0.100$). In the adjusted analysis, there was a significant group-by-occupation interaction ($p=0.017$). Stratifying by occupation revealed that the Ranch Hand enlisted flyers had a significantly higher rate of hyperresonance than the Comparison enlisted flyers ($p=0.006$). Without the group-by-occupation interaction in the model, no difference between the two groups was detected.

There was a total of three participants diagnosed with dullness of the lungs: two Ranch Hands and one Comparison. No difference was found in the unadjusted analysis. Due to the low prevalence rate of dullness, no adjusted analysis was performed.

Neither the unadjusted nor adjusted analyses for wheezes and rales detected a difference between the Ranch Hands and the Comparisons.

No significant difference between the two groups was identified based on the unadjusted analysis of x-ray abnormalities. In the adjusted analysis, there was a significant group-by-race interaction ($p=0.023$). Exploring the interaction by stratifying on race showed a borderline significant difference between the Black Ranch Hands and the Black Comparisons, with the Ranch Hands having more x-ray abnormalities ($p=0.068$). Without the interaction in the model, no significant difference was found.

The eight laboratory variables of the pulmonary assessment were: FVC, FEV1, FEV2, FEV3, FEFmax, ratio of observed FEV1 to observed FVC, loss of vital capacity, and obstructive abnormality. For six of the eight variables, no significant difference was detected between the Ranch Hands and the Comparisons in both the unadjusted and adjusted analyses. These six variables were: FVC, FEV3, FEFmax, ratio of observed FEV1 to observed FVC, loss of vital capacity, and obstructive abnormality.

No significant difference was identified in the unadjusted analysis of FEV1. This result was supported by the adjusted analysis without the significant group-by-age interaction ($p=0.037$). When the interaction was explored, the Ranch Hands born between 1923 and 1941 were found to have a significantly lower adjusted mean than the Comparisons in the same age category ($p=0.022$). However, the Ranch Hands who were born in or before 1922 had a marginally higher adjusted mean than the Comparisons in that category ($p=0.081$).

The results of the analyses of FEV2 were similar to the results of FEV1. No difference between the two groups was detected based on the unadjusted analysis. In the adjusted analysis, there was a significant interaction between group and age ($p=0.042$). Of the participants born between 1923 and 1941, the Ranch Hands had a significantly lower adjusted mean FEV2 than the Comparisons ($p=0.017$). Among the participants who were born in or before 1922, a borderline significant group difference was found with the adjusted mean of the Comparisons being lower than the adjusted mean of the Ranch Hands ($p=0.070$).

Although the results were primarily not significant or borderline significant, the relative risk was greater than 1 or the mean of the Ranch Hands was less favorable than the mean of the Comparisons in 17 of the 20 unadjusted analyses. In general, this pattern was repeated in the adjusted analyses, where the models were adjusted for the effects of cigarette smoking; again, however, the results were primarily not significant. Trends such as these are discussed in the section on interpretive considerations.

Longitudinal analyses showed no changes over time between groups for the ratio of observed FEV1 to observed FVC. The exposure index analyses detected significant results suggestive of a dose-response relationship infrequently, and no pattern in the results emerged. Exposure index-by-covariate interactions observed were primarily with the two smoking covariates.

In conclusion, 14 variables demonstrated nonsignificant results in both unadjusted and adjusted Ranch Hand versus Comparison group contrasts. Two variables exhibited a significant or borderline significant result affecting the Ranch Hands in either the unadjusted or adjusted analyses. The Ranch Hands had more thorax and lung abnormalities than the Comparisons based on the unadjusted analysis; after adjustment for age and current cigarette smoking, the difference was borderline significant. A borderline significant difference in hyperresonance was found in the unadjusted analysis, and a group-by-occupation interaction was present in the adjusted analysis. Four additional variables were nonsignificant in unadjusted analyses with a group-by-covariate interaction present in the adjusted analyses. Of the five interactions, two variables showed a significant detriment to the Ranch Hands, one a significant detriment to the Comparisons, and two variables demonstrated mixed results; that is, significant or borderline significant results were present for both Ranch Hands and Comparisons, depending on which covariate stratum was examined. Without the group-by-covariate interactions in the final model, no significant effects due to group were seen. Although the pulmonary health of the two groups was reasonably comparable, assessment of the pulmonary function should be included in the future examinations.

INTERPRETIVE CONSIDERATIONS

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 exposure 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 prevalences 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 assessment 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 gravity 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.