

## Table of Contents

---

<b>7</b>	<b>STATISTICAL METHODS</b>	<b>7-1</b>
7.1	INTRODUCTION	7-1
7.2	MODELS AND ASSUMPTIONS	7-1
7.2.1	Model 1: Group and Occupation as Estimates of Exposure	7-2
7.2.2	Models 2 through 4: Serum Dioxin as an Estimate of Exposure	7-3
7.2.2.1	Prior Knowledge Regarding Dioxin	7-3
7.2.2.2	Fundamental Limitations of the Serum Dioxin Data	7-4
7.2.2.3	Model 2: Health versus Initial Dioxin in Ranch Hands	7-4
7.2.2.4	Model 3: Health versus Dioxin in Ranch Hands and Comparisons	7-6
7.2.2.5	Model 4: Health versus 1987 Dioxin in Ranch Hands	7-7
7.3	FACTORS DETERMINING THE STATISTICAL ANALYSIS METHOD	7-9
7.4	ANALYSIS METHODOLOGIES	7-11
7.4.1	Methods for Analyzing Continuous and Discrete Variables	7-11
7.4.2	Modeling Strategy	7-16
7.4.3	Longitudinal Analysis	7-16
7.5	INTERPRETIVE CONSIDERATIONS	7-17
7.5.1	Adjustments for Covariates	7-17
7.5.2	Multiple Testing	7-18
7.5.3	Trends	7-18
7.5.4	Interpretation of the Coefficient of Determination	7-18
7.5.5	Clinical Interpretation of Discrete versus Continuous Data	7-18
7.5.6	Power	7-19
7.6	EXPLANATION OF TABLES	7-21
7.6.1	Exposure Analysis	7-21
7.6.1.1	Continuous Variables	7-21
7.6.1.2	Discrete Variables	7-24
7.6.1.2.1	Discrete Variable with Two Categories	7-24
7.6.1.2.2	Discrete Variable with More Than Two Categories	7-26
7.6.2	Longitudinal Analysis	7-28
7.6.2.1	Continuous Variables	7-28
7.6.2.2	Discrete Variables with Two Categories	7-30
7.6.2.2.1	Discrete Variable with More Than Two Categories	7-32
	<b>REFERENCES</b>	<b>7-33</b>

## List of Tables

---

Table 7-1. Model 1: Assessing Health versus Group Status in Ranch Hands and Comparisons: Assumptions, Advantages, and Disadvantages .....	7-3
Table 7-2. Model 2: Assessing Health versus Initial Dioxin in Ranch Hands: Assumptions, Advantages, and Disadvantages .....	7-5
Table 7-3. Model 3: Assessing Health versus Categorized Dioxin in Ranch Hands and Comparisons .....	7-7
Table 7-4. Model 4: Assessing Health versus 1987 Dioxin in Ranch Hands: Assumptions, Advantages, and Disadvantages .....	7-8
Table 7-5. Summary of Statistical Analysis Situations by Dependent Variable Form, Exposure Estimate, Analysis Cohort, and Analysis Type .....	7-9
Table 7-6. Summary of Statistical Procedures .....	7-12
Table 7-7. Approximate Power To Detect a Group Effect at a 5-Percent Level of Significance (Discrete Dependent Variable) .....	7-19
Table 7-8. Approximate Power To Detect a Group Effect at a 5-Percent Level of Significance (Continuous Dependent Variable) .....	7-20
Table 7-9. Location of Table Results from Different Exposure Analysis Models .....	7-21
Table 7-10. Location of Table Results from Different Longitudinal Analysis Models .....	7-28

## **7 STATISTICAL METHODS**

---

### **7.1 INTRODUCTION**

This chapter summarizes the statistical methods used in the analysis of Air Force Health Study (AFHS) 1997 follow-up examination data to investigate relations between the health status of the 2,121 participants attending this examination and their corresponding group (Ranch Hand or Comparison) or serum dioxin estimates and measurements. Group contrast models were similar to analyses performed for the 1982 baseline and 1985, 1987, and 1992 follow-up examinations (1, 2, 3, 4). Models relating health to dioxin estimates and measurements were based on analyses performed for the Serum Dioxin Analysis Report for the 1987 Follow-up and 1992 follow-up examinations (4, 5).

The statistical methods used in this report encompassed four different forms of hypotheses or models applied to 266 study endpoints. Each of these models specified the study cohort or subset of participants included in the respective analyses together with the dioxin exposure or proxy estimates used in the analysis. The first model (Model 1) specified contrasts between Ranch Hands and Comparisons using group as a proxy for exposure, and it did not incorporate serum dioxin measurements. The remaining three models (Models 2, 3, and 4) all incorporated serum dioxin measurements. A summary description of each of the four models is provided in section 7.2, “Models and Assumptions.”

Each model and exposure estimate combination was implemented for study variables and type of analysis (unadjusted, adjusted, or longitudinal). The specific statistical procedures (e.g., analysis of variance or logistic regression) that were used are presented in section 7.3, “Factors Determining Statistical Analysis Method.” The relation between the factors and statistical procedures is presented in section 7.4, “Analysis Methodologies.” That presentation is followed by a discussion of “Interpretive Considerations” (section 7.5), and a description of the contents of tables used to report statistical analysis results throughout the report is given in the “Explanation of Tables” (section 7.6).

### **7.2 MODELS AND ASSUMPTIONS**

The statistical analysis was based primarily on four models, each using a different estimate of exposure. The first model used group and military occupation (officer, enlisted flyer, and enlisted groundcrew) to assess health effects and dose-response relations related to herbicide exposure. Serum dioxin measurements were not used in this model. The other three models accounted for dioxin effects either through estimated initial dioxin levels for Ranch Hands or using current or recent serum dioxin levels for Ranch Hands and Comparisons to assess health endpoints (e.g., cholesterol, diabetes) and dose-response relations related to exposure. These analyses were accomplished with and without adjustment for covariates.

Throughout this report, dioxin levels are used as measures of both exposure to dioxin itself and exposure to dioxin-contaminated herbicides, including Herbicide Orange. Direct contrasts of Ranch Hand and Comparison veterans (Model 1) address the hypothesis of health effects attributable to any herbicide exposure experienced by Ranch Hand veterans during Operation Ranch Hand. Models involving dioxin measurements address the hypothesis that health effects change with the amount of exposure. Dioxin measurements are used as a measure of exposure to dioxin-contaminated herbicides because it is expected that as exposure to such herbicides increased, dioxin levels should increase. Therefore, the dioxin measurement serves as direct biomarker of exposure to dioxin-contaminated herbicides. No other

direct measure or estimate of herbicide exposure is available with which to address hypothetical dose-response relations with health. Some indirect measures, such as self-report of skin contact among enlisted groundcrew, or simply being a Ranch Hand enlisted groundcrew member, are valuable alternatives because dioxin measures suggest that enlisted groundcrew experienced the heaviest exposures. Reported skin exposure is not addressed in this report, but enlisted groundcrew status is addressed in Model 1. The use of dioxin as a measure of exposure to dioxin-contaminated herbicides is consistent with the goal of the study, which is to determine whether health effects exist and can be attributed to occupational exposure to Herbicide Orange (6).

### 7.2.1 Model 1: Group and Occupation as Estimates of Exposure

This section describes the model that used the exposure group (Ranch Hand, Comparison) to assess the relation between health status and dioxin exposure. Statistical analyses based on this model were termed “Model 1” in the assessment of the clinical areas. Analyses of this type are straightforward, easy to interpret, and well established in epidemiological studies. In this model, exposure was defined as “yes” for Ranch Hands and “no” for Comparisons without regard to the magnitude of the exposure. As an attempt to quantify exposure, three contrasts of Ranch Hands and Comparisons were performed along with the overall Ranch Hand versus Comparison contrast. These three contrasts compared Ranch Hands and Comparisons within each occupational category (officers, enlisted flyers, and enlisted groundcrew). As described in the analyses performed for the Serum Dioxin Analysis Report for the 1987 Followup (5), the average levels of exposure to dioxin were highest for enlisted groundcrew, followed by enlisted flyers, then officers.

Table 7-1 provides a description of Model 1 and gives the assumptions, advantages, and disadvantages for a continuously distributed health endpoint,  $y$ . The model presented in Table 7-1 is unadjusted for any covariates—adjusted models are a straightforward extension.

**Table 7-1. Model 1: Assessing Health versus Group Status in Ranch Hands and Comparisons: Assumptions, Advantages, and Disadvantages**

<b>Model 1: <math>y = \mu + G_i + e</math> (All Ranch Hands and Comparisons)</b>	
<b><math>y = \mu + G_i + O_j + (GO)_{ij} + e</math> (Ranch Hands and Comparisons by occupation)</b>	
where	
y	= health variable in group i and occupation j
$G_i$	= effect due to group status (i = 1,2 – Comparisons, Ranch Hands)
$O_j$	= effect due to occupation (j = 1,2,3 – Officers, Enlisted Flyers, Enlisted Groundcrew)
$(GO)_{ij}$	= interaction between group status and occupation (i = 1,2; j = 1,2,3); used to examine Ranch Hand and Comparison differences for each occupation
e	= zero mean error.
Assumptions:	Comparisons were unexposed and Ranch Hands were exposed.
	For the purposes of investigating dose-response effects, enlisted groundcrew were more heavily exposed than enlisted flyers, and enlisted flyers were more heavily exposed than officers.
	The error variance does not change with group status or occupation.
Advantages:	Easily interpretable.
Disadvantages:	Results are biased toward the null hypothesis of no dioxin effect if unexposed Ranch Hands are misclassified (i.e., remain in the analysis as exposed Ranch Hands). It is not possible to fully distinguish unexposed Ranch Hands from exposed Ranch Hands.

### 7.2.2 Models 2 through 4: Serum Dioxin as an Estimate of Exposure

Current dioxin levels in 1987 were determined by the Centers for Disease Control and Prevention from serum samples taken from approximately 2,000 Ranch Hands and Comparisons. Additional serum samples were taken from selected Ranch Hands and Comparisons at the 1992 and 1997 follow-up examinations to provide insight on dioxin levels and the elimination of dioxin from the body, and to provide measurements for new subjects and those who were not previously measured. A discussion of the details of dioxin measurement is found in Chapter 2, Dioxin Assay.

Investigation of the mechanics of dioxin elimination is currently under study by the Air Force. Based on samples collected in 1982, 1987, 1992, and 1997, issues such as half-life estimation and first-order pharmacokinetic assumptions are being investigated.

#### *7.2.2.1 Prior Knowledge Regarding Dioxin*

This section presents analytic strategies based on assumptions and models conceived after the Ranch Hand half-life study published in 1996 (7). Available data have suggested that the dioxin elimination process is first-order, based on measurements subsequent to the ingestion of dioxin by an individual (8). Data on 213 Ranch Hand veterans with dioxin measured in blood collected in 1982, 1987, and 1992 produced a half-life estimate of 8.7 years (7); this estimate was used in all calculations involving half-life.

The term “elimination” denotes the overall removal of dioxin from the body. Some of the analyses assumed that the amount of dioxin in the body (C) decreases exponentially with time according to the model  $C = I \cdot \exp(-rt)$ , where I is the initial level,  $r = \log(2)/h$  is the elimination rate, h is the half-life, and t is the number of years from the end of service in Southeast Asia (SEA) to the time of the blood measurement for dioxin. If a participant had measurements at more than one point in time, the measurement closest to the time of duty in SEA was used. This exponential elimination law is termed “first-order elimination.”

The first-order elimination assumption is equivalent to assuming a one-compartment model for dioxin distribution within the body. While a multicompartment model incorporating body composition and dioxin binding to tissue receptors would provide a detailed description of dioxin concentrations in different compartments, published multicompartment models for dioxin distribution within the body predict first-order elimination of dioxin, overwhelmingly because of fecal elimination (9).

The lipid-weight concentration of dioxin, expressed in parts per trillion (ppt) (10, 11), is a derived quantity calculated from the formula  $\text{ppt} = \text{ppq} \cdot 102.6/W$ , where ppt is the lipid-weight concentration, ppq (parts per quadrillion) is the actual whole weight of dioxin in the sample in femtograms, 102.6 corrects for the average density of serum, and W is the total lipid weight of the sample (9).

The relation between the serum lipid-weight concentration of dioxin and lipid-weight concentrations in adipose tissue is a subject of continuing research. The correlation between the serum lipid-weight concentration and adipose tissue lipid-weight concentration of dioxin has been observed by Patterson, et al., to be 0.98 in 50 persons from Missouri (12). Using the same data, Patterson, et al., calculated the partitioning ratio of dioxin between adipose tissue and serum on a lipid-weight basis as 1.09 (95% confidence interval: [0.97,1.21]). On the basis of these data, a one-to-one partitioning ratio of dioxin between lipids in adipose tissue and the lipids in serum could not be excluded. Measurements of dioxin in adipose tissue generally have been accepted as representing the body burden concentration of dioxin. The high correlation between serum dioxin levels and adipose tissue dioxin levels in the Patterson, et al., study suggests that serum dioxin is also a valid measurement of dioxin body burden.

#### 7.2.2.2 *Fundamental Limitations of the Serum Dioxin Data*

There are two evident limitations to the available data:

- While Ranch Hand data did not appear to violate a first-order elimination assumption, no serially repeated dioxin assay results, taken over many years and with which to evaluate directly the adequacy of the first-order elimination model in humans, were available.
- It was not known whether Ranch Hands with body burdens of dioxin at or below 10 ppt were exposed and their body burdens had decreased to these levels since their time of duty in SEA, or whether they were not exposed at all during their time of duty in SEA.

#### 7.2.2.3 *Model 2: Health versus Initial Dioxin in Ranch Hands*

The relation between estimated initial dioxin levels and health was assessed in Ranch Hands using the model described in Table 7-2. Statistical analyses based on this model were termed “Model 2” in the assessment of the clinical areas. In this model, an initial dioxin level was estimated for a Ranch Hand from a current or recent lipid-adjusted dioxin measure, the length of time between the time of duty in SEA and the date of the blood measurement of dioxin, and an estimated half-life of 8.7 years. From studies conducted by the Air Force, body fat at the time of the blood measurement of dioxin appeared to be related to the dioxin half-life for a participant (7). Hence, this body fat measure was included in this

model as a covariate. Model 2 differs from Model 1 in that the estimate of exposure in Model 1 (group: Ranch Hand, Comparison) was not dependent upon extrapolation to an earlier date.

Table 7-2 also includes assumptions, advantages, and disadvantages of the model for a continuously distributed health variable, y. The model presented in Table 7-2 is unadjusted for any additional risk factors, but extension to an adjusted model is straightforward.

**Table 7-2. Model 2: Assessing Health versus Initial Dioxin in Ranch Hands: Assumptions, Advantages, and Disadvantages**

<b>Model 2: <math>y = b_0 + b_1 \log_2(I) + b_2 BF + e</math></b>	
where	
y	= health variable
I	= extrapolated initial dose, assuming first-order elimination, $I = 4 + (C - 4) \cdot \exp(\log(2) \cdot t/h)$ , where 4 ppt is considered the median background level of lipid-adjusted dioxin; t = length of time between the time of duty in SEA and the date of the blood measurement of dioxin in 1987, 1992, or 1997; C = lipid-adjusted dioxin, determined in 1987, 1992, or 1997; and h = dioxin half-life in Ranch Hands assuming first-order elimination (8.7 years assumed for analysis)
BF	= body fat at the time of the blood measurement of dioxin, calculated from the formula shown below
e	= zero mean error.
Body fat was calculated from a metric body mass index (13); the formula is	
$\text{Body Fat (in percent)} = \frac{\text{Weight (kg)}}{[\text{Height (m)}]^2} \cdot 1.264 - 13.305.$	
Assumptions:	Ranch Hands received a single dioxin dose in Vietnam and background exposure thereafter. Ranch Hands experienced first-order dioxin elimination. The error variance does not change with health status or initial dioxin dose.
Advantages:	Easily interpretable. Most efficient if first-order elimination and half-life are valid and y is linearly related to $\log_2(I)$ . The logarithm (base 2) of initial dioxin presents the dioxin data as a more symmetric distribution than the distribution of initial dioxin in its original units. In addition, the relative risk based on the logarithm (base 2) of initial dioxin is more meaningful than on the original scale (i.e., a doubling of initial dioxin rather than a 1 ppt increase in dioxin).
Disadvantages:	Results are biased if first-order elimination or constant half-life assumptions are not valid.

In Table 7-2, the phrase “single dioxin dose” is a simplification of the process by which Ranch Hands accumulated dioxin during their time of duty in SEA. This process, which undoubtedly varied from individual to individual, is unknown; however, the time of duty in SEA for an individual Ranch Hand generally was short (1 to 3 years) relative to the time elapsed since his duty in SEA. Hence, additional knowledge regarding the accumulation of dioxin during an individual Ranch Hand’s time of duty in SEA, were it to become available, would not likely change conclusions drawn from any of the statistical analyses.

Analyses were performed on Ranch Hands who had lipid-adjusted dioxin levels greater than 10 ppt at either the 1987, 1992, or 1997 physical examinations. The value 10 ppt corresponds to the approximate 98th percentile of the Comparison lipid-adjusted dioxin distribution. Based on this Comparison dioxin distribution, it was believed that participants with greater than 10 ppt lipid-adjusted dioxin were definitely exposed. It was not known whether Ranch Hands with dioxin burdens at or below 10 ppt were exposed and their body burdens had decreased to these levels since their time of duty in SEA, or whether they were not exposed at all during their time of duty in SEA. Lipid-adjusted dioxin levels less than 10 ppt are subsequently called “background” levels.

#### 7.2.2.4 *Model 3: Health versus Dioxin in Ranch Hands and Comparisons*

An assessment of the health consequences of dioxin above background levels was carried out with a model that was applied to both Ranch Hand and Comparison data. This model assessed health versus dioxin body burden categorized into four levels, given below:

- Comparisons—Comparisons with up to 10 ppt lipid-adjusted dioxin
- Background—Ranch Hands with up to 10 ppt lipid-adjusted dioxin
- Low—Ranch Hands with more than 10 ppt lipid-adjusted dioxin but at most 94 ppt estimated initial dioxin
- High—Ranch Hands with more than 10 ppt lipid-adjusted dioxin and more than 94 ppt estimated initial dioxin.

Statistical analyses based on this model were termed “Model 3” in the assessment of the clinical areas. The low and high Ranch Hand categories, of approximately equal size, were determined by the median estimated initial dioxin level (94 ppt) of the Ranch Hands with more than 10 ppt lipid-adjusted dioxin (i.e., the sample used in Model 2). In this model, an initial dioxin level was estimated for a Ranch Hand from a current or recent lipid-weight dioxin measure, the length of time between the time of duty in SEA and the date of the blood measurement of dioxin, and an estimated half-life of 8.7 years. From studies conducted by the Air Force, body fat at the time of the blood measurement of dioxin appeared to be related to the dioxin half-life for a participant. This body fat measure was included in this model as a covariate. Using this body fat measure in Model 3 for all Comparisons and Ranch Hands with dioxin measurements allowed body fat to act as a potential risk factor as well as an adjusting variable to explain half-life differences.

For a continuously distributed health variable,  $y$ , for example, the mean values of  $y$  within the background, low, high, and low plus high categories were contrasted with the mean values of  $y$  within the Comparison category. The mean value of  $y$  for the low plus high category was calculated as a linear combination of the low dioxin category and the high dioxin category, with weights based on the sample size in each of these categories. Relative frequencies were contrasted for discrete health variables. Table 7-3 shows this model and the assumptions, advantages, and disadvantages for the unadjusted analysis of a continuous variable; extension to an adjusted model is straightforward.

**Table 7-3. Model 3: Assessing Health versus Categorized Dioxin in Ranch Hands and Comparisons**

<b>Model 3: <math>y = b_0 + b_1I_1 + b_2I_2 + b_3I_3 + b_4I_4 + b_5BF + e</math></b>	
where	
y	= health variable
I <sub>1</sub>	= indicator variable for categorized dioxin; I <sub>1</sub> = 1 if participant is a Comparison with a background level of dioxin, I <sub>1</sub> = 0 if participant is not a Comparison
I <sub>2</sub>	= indicator variable for categorized dioxin; I <sub>2</sub> = 1 if participant is in background dioxin category, I <sub>2</sub> = 0 if participant is not in background dioxin category
I <sub>3</sub>	= indicator variable for categorized dioxin; I <sub>3</sub> = 1 if participant is in low dioxin category, I <sub>3</sub> = 0 if participant is not in low dioxin category
I <sub>4</sub>	= indicator variable for categorized dioxin; I <sub>4</sub> = 1 if participant is in high dioxin category, I <sub>4</sub> = 0 if participant is not in high dioxin category
BF	= body fat at the time of blood measurement of dioxin, calculated from the formula shown below
e	= zero mean error.
Body fat was calculated from a metric body mass index (13); the formula is	
$\text{Body Fat (in percent)} = \frac{\text{Weight (kg)}}{[\text{Height (m)}]^2} \cdot 1.264 - 13.305.$	
Assumptions:	Dioxin body burden has been eliminated with time. The error variance does not change with categorized dioxin body burden.
Advantages:	Initial dioxin is probably a better measure for determining low and high exposure than current or recent lipid-adjusted dioxin measurements. Less dependent on the accuracy of the estimation algorithm for determining initial dioxin than Model 2.
Disadvantages:	Makes no use of prior belief that some Ranch Hands received an unusually large dioxin dose in Vietnam; all Ranch Hands with high dioxin levels are treated similarly. “Background” Ranch Hand category is probably a mixture of exposed and unexposed Ranch Hands. Analysis may be biased toward the null hypothesis of no dioxin effect. “Low” and “high” Ranch Hand categories are based on initial dioxin model, which is based on valid half-life and first-order dioxin elimination. Bias is possible if model is incorrect. Also, a conditional null hypothesis is tested using these categories (“Is there a dioxin effect, given a specified level of exposure?”).

7.2.2.5 *Model 4: Health versus 1987 Dioxin in Ranch Hands*

The relation between 1987 dioxin and health was assessed using the model described in Table 7-4. This measure of dioxin is termed “1987 dioxin” because most Ranch Hands were assayed for dioxin initially at the 1987 follow-up examination. This table also describes the assumptions, advantages, and disadvantages for the unadjusted analysis of a continuously distributed health variable, y.

Ranch Hands with a dioxin measurement may have had their blood collected at the pilot study in April 1987, at the 1987 physical examination, at the 1992 physical examination, or at the 1997 physical examination. If an individual had measurements at more than one of these points in time, the measurement closest to the time of duty in SEA was used. If only a 1992 serum dioxin measurement was available, the level was extrapolated to the date of the 1987 physical examination. The model

$$C_{1987} = 4 + (C_{1992} - 4) \cdot \exp(rt)$$

was used for extrapolation of lipid-adjusted dioxin to 1987 levels ( $C_{1987}$ ), where  $C_{1992}$  is the lipid-adjusted dioxin level in 1992, 4 ppt is considered the median background level for lipid-adjusted dioxin,  $r = \log(2)/h$  is the elimination rate,  $h$  is the half-life (8.7 years), and  $t$  is the length of time between the physical examination in 1987 and the physical examination in 1992. This model was used only if the lipid-adjusted dioxin level in 1992 was greater than 10 ppt; otherwise, the 1992 measurement was used. A similar strategy was used for participants who had only a 1997 serum dioxin measurement. The estimate of exposure in Model 4 (1987 dioxin) was based on extrapolation to 1987 for only 39 out of the 863 Ranch Hands. Most measurements were based on 1987 dioxin measurements and extrapolation was not needed. Consequently, body fat at the time of the blood measurement of dioxin was not used in Model 4, which was different from the strategy used for Models 2 and 3.

**Table 7-4. Model 4: Assessing Health versus 1987 Dioxin in Ranch Hands: Assumptions, Advantages, and Disadvantages**

<b>Model 4: <math>y = b_0 + b_1 \log_2(\text{ppt} + 1) + e</math></b>	
where	
y	= health variable
ppt	= lipid-adjusted dioxin = $\text{ppq} \cdot 102.6 / W$ , where ppq = whole weight of dioxin in the sample in femtograms (102.6 corrects for the average density of serum) and W = total lipid weight of the sample
e	= zero mean error.
Assumptions:	Ranch Hands received a single dioxin dose in Vietnam and background exposure thereafter. The error variance does not change with health status or 1987 dioxin.
Advantages:	Using 1987 dioxin has less inherent variation than initial dioxin, which is extrapolated by a first-order elimination model across a 20- to 30-year time period. The logarithm (base 2) of (1987 dioxin + 1) presents the dioxin data as a more symmetric distribution than the distribution of 1987 dioxin in its original units. In addition, the relative risk based on the logarithm (base 2) of (1987 dioxin + 1) is more meaningful than on the original scale (i.e., a doubling of 1987 dioxin + 1, rather than a 1 ppt increase in dioxin).
Disadvantages:	1987 dioxin may not be a good surrogate for exposure if elimination rate differs among individuals. Individuals with measurements in 1992 only or 1997 only are extrapolated to 1987, and variation is increased with estimation using a first-order elimination model.

The relation between current health and dioxin was assessed using a model, termed “Model 4,” with lipid-adjusted 1987 dioxin as the estimate of exposure. Model 4 used the logarithm (base 2) of lipid-adjusted 1987 dioxin and is described in Table 7-4.

### 7.3 FACTORS DETERMINING THE STATISTICAL ANALYSIS METHOD

For a specified questionnaire-based or clinical measurement determined from the physical or laboratory examination, the selection of an analytical method depended on each of the following:

- Dependent Variable Form: Continuous or discrete
- Exposure Estimate and Analysis Cohort:
  - Model 1: Group—All Ranch Hands and Comparisons
  - Model 2: Initial dioxin—Ranch Hands having a dioxin body burden of greater than 10 ppt of lipid-adjusted dioxin, based on 1987 dioxin levels as defined in Section 7.2.2.5
  - Model 3: Categorized dioxin—Comparisons with a dioxin body burden of 10 ppt lipid-weight dioxin or less, based on 1987 dioxin levels, and all Ranch Hands with a dioxin measurement
  - Model 4: 1987 dioxin—All Ranch Hands with a dioxin measurement
- Analysis Type: Unadjusted, adjusted, or longitudinal.

Table 7-5 specifies 22 separate analysis situations based on dependent variable form, exposure estimate, analysis cohort, and analysis type. For each of the 22 situations, the statistical method is specified. For example, linear regression models were used for adjusted analyses of initial dioxin for continuous dependent variables.

**Table 7-5. Summary of Statistical Analysis Situations by Dependent Variable Form, Exposure Estimate, Analysis Cohort, and Analysis Type**

Exposure Estimate	Analysis Cohort	Analysis Type	Statistical Methods	Independent Variables
<b>Continuous</b>				
Model 1: Group (Ranch Hands vs. Comparisons)	All RH & C	Unadjusted	Analysis of Variance	Group
		Adjusted	Analysis of Covariance	Group; Covariates
		Longitudinal <sup>a</sup>	Analysis of Covariance	Group; Age at the 1997 Follow-up Examination; 1982 Measurement
Model 2: Log <sub>2</sub> (Initial)	RH >10 ppt lipid-adjusted 1987 dioxin	Unadjusted	Linear Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin
		Adjusted	Linear Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin; Covariates

**Table 7-5. Summary of Statistical Analysis Situations by Dependent Variable Form, Exposure Estimate, Analysis Cohort, and Analysis Type (Continued)**

Exposure Estimate	Analysis Cohort	Analysis Type	Statistical Methods	Independent Variables
		Longitudinal <sup>a</sup>	Linear Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin; Age at the 1997 Follow-up Examination; 1982 Measurement
Model 3: Categorized Dioxin	All RH with a dioxin measurement, C ≤10 ppt lipid-adjusted 1987 dioxin	Unadjusted	Analysis of Covariance	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin
		Adjusted	Analysis of Covariance	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin; Covariates
		Longitudinal <sup>a</sup>	Analysis of Covariance	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin; Age at the 1997 Follow-up Examination; 1982 Measurement
Model 4: Log <sub>2</sub> (1987 Dioxin + 1)	All RH with a dioxin measurement	Unadjusted	Linear Regression	Log <sub>2</sub> (1987 Dioxin + 1)
		Adjusted	Linear Regression	Log <sub>2</sub> (1987 Dioxin + 1); Covariates
<b><u>Discrete</u></b>				
Model 1: Group (Ranch Hands vs. Comparisons)	All RH & C	Unadjusted	Chi-Square Contingency Table, Logistic Regression	Group
		Adjusted	Logistic Regression	Group; Covariates
		Longitudinal <sup>b</sup>	Logistic Regression	Group; Age at the 1997 Follow-up Examination
Model 2: Log <sub>2</sub> (Initial)	RH >10 ppt lipid-adjusted 1987 dioxin	Unadjusted	Logistic Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin
		Adjusted	Logistic Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin; Covariates
		Longitudinal <sup>b</sup>	Logistic Regression	Log <sub>2</sub> (Initial); Body Fat at the Time of the Blood Measurement of Dioxin; Age at the 1997 Follow-up Examination
Model 3: Categorized Dioxin	All RH with a dioxin measurement, C ≤10 ppt lipid-adjusted 1987 dioxin	Unadjusted	Chi-Square Contingency Table; Logistic Regression	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin
		Adjusted	Logistic Regression	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin; Covariates

**Table 7-5. Summary of Statistical Analysis Situations by Dependent Variable Form, Exposure Estimate, Analysis Cohort, and Analysis Type (Continued)**

Exposure Estimate	Analysis Cohort	Analysis Type	Statistical Methods	Independent Variables
		Longitudinal <sup>b</sup>	Logistic Regression	DXCAT; Body Fat at the Time of the Blood Measurement of Dioxin; Age at the 1997 Follow-up Examination
Model 4: Log <sub>2</sub> (1987 Dioxin + 1)	All RH with a dioxin measurement	Unadjusted	Logistic Regression	Log <sub>2</sub> (1987 Dioxin + 1)
		Adjusted	Logistic Regression	Log <sub>2</sub> (1987 Dioxin + 1); Covariates

<sup>a</sup> Dependent variable usually paired difference score of (1997 to 1982) dependent variable values. For some clinical areas, paired difference scores were (1997 to 1985) differences.

<sup>b</sup> Analysis performed subject to the constraint that participant was normal at the 1982 baseline (or 1985) examination.

Note: Log<sub>2</sub> (Initial) = Logarithm (base 2) of estimated initial dioxin level.

Log<sub>2</sub> (1987 Dioxin + 1) = Logarithm (base 2) of (1987 dioxin level + 1).

DXCAT = Categorized dioxin (incorporating group membership—three categories for Ranch Hands, one category for Comparisons).

RH = Ranch Hand.

C = Comparison.

## 7.4 ANALYSIS METHODOLOGIES

### 7.4.1 Methods for Analyzing Continuous and Discrete Variables

For analyses of continuous dependent variables, the general linear models approach was used for applying such techniques as simple and multiple linear regression, analysis of variance, analysis of covariance, repeated measures analysis, and survival time analysis. This approach permitted model fitting of the dependent variable as a function of group or dioxin and specified covariates. Continuous dependent variables were examined to ensure that assumptions underlying appropriate statistical methods were met. Transformations (e.g., square root, logarithmic) were used to enhance normality for specific continuous health variables. A further discussion of general linear models, as well as other methods used for the statistical analyses in this report, is found in Table 7-6.

For these continuous analyses, the SAS<sup>®1</sup> general linear models analysis (PROC GLM) (14) was used. After a model was fitted, tests of significance for a group or dioxin effect were developed. Associations with a p-value less than or equal to 0.05 were described as significant, and associations with a p-value greater than 0.05 but less than or equal to 0.10 were described as marginally significant.

The SAS<sup>®</sup> procedures LIFEREG and LIFETEST (14) were used for the time to diabetes onset variable in the endocrine clinical assessment. Statistical methods used to analyze measures of this type implemented

<sup>1</sup> SAS and all other SAS Institute, Inc., product and service names are registered trademarks or trademarks of SAS Institute, Inc., in the USA and other countries.

a technique known as “survival time” analysis. A further discussion of survival time analysis is found in Table 7-6.

For dichotomous discrete dependent variables, logistic regression was performed using SAS<sup>®</sup> PROC GENMOD (15). For dependent variables with more than two categories, polytomous logistic regression was performed using SAS<sup>®</sup> PROC CATMOD (14). Parameter estimation and model selection for polytomous logistic regression and ordinary logistic regression are similar. Both forms of regression use the maximum likelihood principle to obtain parameter estimates. For a model with k parameters for two equations, 2k parameters are estimated, k for each logit function. If ordinary logistic regression is applied twice (for example, once for abnormal low versus normal and then for abnormal high versus normal), 2k parameters are estimated; however, ordinary logistic regression maximizes two likelihood equations, each with k parameters, while polytomous logistic regression estimates all 2k parameters simultaneously with one likelihood equation. Polytomous logistic regression also can be used for dependent variables that have more than three levels and require more than two contrasts with a normal category. A further discussion of logistic regression and polytomous logistic regression is found in Table 7-6.

A chi-square statistic, adjusted for the continuity of the chi-square distribution, was used when a test of the relative frequency of abnormal measurements between Ranch Hands and Comparisons was performed, and the relative frequency of either the Ranch Hand or the Comparison group was zero. This test statistic yields p-values approximately equal to Fisher’s exact test (16) for a two-sided alternative hypothesis.

---

**Table 7-6. Summary of Statistical Procedures**

<p><b>Chi-Square Contingency Table Test</b></p> <p>The chi-square test of independence (17) is calculated for a contingency table by the following formula:</p> $\chi^2 = \sum \frac{(f_O - f_E)^2}{f_E}$ <p>where the sum is taken over all cells of the contingency table and  <math>f_O</math> = observed frequency in a cell  <math>f_E</math> = expected frequency under the hypothesis of independence.</p> <p>Large values indicate deviations from the null hypothesis and are tested for significance by comparing the calculated <math>\chi^2</math> to the tables of the chi-square distribution.</p> <p>For 2x2 tables, the chi-square statistic above can be adjusted for the continuity of the <math>\chi^2</math> distribution. This test statistic yields p-values approximately equal to Fisher’s exact test (16) for a two-sided alternative and is as follows:</p> $\chi^2 = \sum \frac{\max(0, ( f_O - f_E  - \frac{1}{2}))^2}{f_E}$
--

**Table 7-6. Summary of Statistical Procedures (Continued)**

**Correlation Coefficient (Pearson's Product-Moment)**

The population correlation coefficient  $\rho$  (18) measures the strength of the linear relation between two random variables X and Y. A commonly used sample-based estimate of this correlation coefficient is

$$\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2]}}$$

where the sum is taken over all (x, y) pairs in the sample. A student's t-test based on this estimator is used to test for a significant correlation between the two random variables of interest. For the sample size of approximately 2,121 in this study, a sample correlation coefficient of 0.04254 is sufficient to attain a statistically significant correlation at a 5-percent level for a two-sided hypothesis test, assuming normality of X and Y.

**Survival Time Analysis**

The survival time model (19) permits a dependent variable with censored observations to be modeled in a general linear models framework. For example, if the time to diabetes onset is defined as an event, the time for participants for which this event has not occurred is right-censored. The survival time model is

$$y = X\beta + \sigma\epsilon$$

where

- y = vector of responses (e.g., time to diabetes onset), usually the logarithm of the survival times
- X = matrix of covariates, or risk factors (e.g., group status and age)
- $\beta$  = vector of unknown regression parameters
- $\sigma$  = unknown scale parameter
- $\epsilon$  = vector of errors assumed to have a known distribution.

For a model with a dependent variable containing right-censored data, the log likelihood function is a combination of a probability density function for noncensored values and a survival distribution function for right-censored values. The model parameters can be estimated by maximum likelihood in SAS<sup>®</sup> PROC LIFEREG, using a Newton-Raphson algorithm, where the distribution of the random error term can be specified. The distributional assumptions regarding the error term can be tested by examining plots of the Kaplan-Meier survival functions using SAS<sup>®</sup> PROC LIFETEST.

PROC LIFEREG provides estimates, standard errors, and p-values associated with a chi-square test on each parameter (i.e., risk factor) in the model. These are used to test the significance of the group or dioxin term in the unadjusted and adjusted models. In this procedure, percentile estimates also can be produced for each group or each dioxin category in the unadjusted model. The percentile estimates are used to determine parameter estimates from the Weibull distribution. The Weibull distribution parameter estimates are then used in an iterative nonlinear estimation procedure (SAS<sup>®</sup> PROC NLIN [14]) to produce estimated means from a censored Weibull distribution. The loss function that is minimized in the estimation procedure is

$$Loss = -\log[x \cdot (\frac{\beta}{\theta^\beta} \cdot y^{\beta-1} \cdot e^{-(\frac{y}{\theta})^\beta}) + (1-x) \cdot (1 - e^{-(\frac{y}{\theta})^\beta})]$$

- where x = 1 if diabetic
- x = 0 if not diabetic
- and y = time to onset of diabetes.

**Table 7-6. Summary of Statistical Procedures (Continued)**

**General Linear Models Analysis**

The form of the general linear model (18) for two independent variables is

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

where

- Y = dependent variable (continuous)
- $\alpha$  = level of Y at  $X_1 = 0$  and  $X_2 = 0$  (i.e., the intercept)
- $X_1, X_2$  = measured value of the first and second independent variables, respectively, which may be continuous or discrete (e.g., group status and age)
- $\beta_1, \beta_2$  = coefficient indicating linear association between Y and  $X_1$ , Y and  $X_2$ , respectively; each coefficient reflects the effect on the model of the corresponding independent variable adjusted for the effect of the other independent variable
- $\epsilon$  = error term.

This model assumes that the error terms are independent and normally distributed with a mean of 0 and a constant variance. Extension to more than two independent variables is immediate. Simple linear regression, multiple linear regression, analysis of variance, analysis of covariance, and repeated measures analysis of variance are all examples of general linear models analysis.

**Logistic Regression Analysis**

The logistic regression model (20) enables a dichotomous dependent variable to be modeled in a regression framework with continuous and discrete independent variables. For two risk factors, such as dioxin and age, the logistic regression model is

$$\text{logit } P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

where

- P = probability of disease for an individual with risk factors  $X_1$  and  $X_2$
- logit P =  $\ln (P/(1-P))$  (i.e., the log odds for disease)
- $X_1$  = first risk factor (e.g., dioxin)
- $X_2$  = second risk factor (e.g., age).

The parameters are interpreted as follows:

- $\alpha$  = log odds for the disease when  $X_1 = 0$  and  $X_2 = 0$
- $\beta_1$  = coefficient indicating the dioxin effect adjusted for age
- $\beta_2$  = coefficient indicating the age effect adjusted for dioxin
- $\epsilon$  = error term.

For a dichotomous measure, the term  $\exp(\beta_1)$  equals the adjusted odds ratio of abnormal versus normal for Ranch Hands ( $X_1 = 1$ ) compared to Comparisons ( $X_1 = 0$ ). If the probability of being abnormal is small compared to being normal for both the Ranch Hand and Comparison groups, the odds ratio is approximately equal to the relative risk of being abnormal between the two groups. If  $X_1$  is a continuous covariate,  $\exp(\beta_1)$  represents the adjusted odds ratio of outcome 1 versus outcome 0 for a unit increase in  $X_1$ . If the risk factor is expressed in logarithmic (base 2) form,  $\exp(\beta_1)$  reflects the adjusted odds ratio for a twofold increase in the risk factor. Throughout this report and previous reports, the adjusted odds ratio was referred to as an adjusted relative risk. Correspondingly, in the absence of covariates (i.e., unadjusted analysis), the unadjusted odds ratio was referred to as an estimated relative risk.

This technique also was used for longitudinal analyses of dichotomous dependent variables to examine changes in health status between 1982 (or 1985) and 1997 in relation to the dioxin measures.

**Table 7-6. Summary of Statistical Procedures (Continued)**

**Polytomous Logistic Regression Analysis**

Polytomous logistic regression (20, 21) allows a categorical dependent variable with more than two outcomes to be modeled in a regression environment with continuous and discrete independent variables. For polytomous logistic regression, the model equation depends on the scale of the dependent variable. This discussion focuses on nominal scaled dependent variables.

Suppose Y is a nominal scaled dependent variable with three outcomes labeled 0, 1, or 2 (normal, low, or high). Polytomous logistic regression models two logit functions, one for Y = 1 versus Y = 0 and the other for Y = 2 versus Y = 0. The zero outcome for Y is called the reference category. To model Y with two covariates such as group status and age, the polytomous regression model would be

$$\text{logit } P_1 = \alpha_1 + \beta_{1(1)}X_1 + \beta_{1(2)}X_2 + \varepsilon_1$$

$$\text{logit } P_2 = \alpha_2 + \beta_{2(1)}X_1 + \beta_{2(2)}X_2 + \varepsilon_2$$

where

- $P_i$  = probability that Y = i (outcome i) with covariates  $X_1$  and  $X_2$ ,  $i = 0, 1, 2$
- $\text{logit } P_i$  =  $\ln(P_i/P_0)$  (i.e., the log odds of outcome i versus outcome 0,  $i = 1, 2$ )
- $X_1$  = first effect (e.g., group status)
- $X_2$  = second effect (e.g., age).

The parameters are interpreted as follows:

- $\alpha_i$  = log odds of outcome i versus outcome 0 when  $X_1 = 0$  and  $X_2 = 0$ ,  $i = 1, 2$
- $\beta_{i(1)}$  = coefficient indicating the group status effect on the logit  $P_i$ , adjusted for age;  $i = 1, 2$
- $\beta_{i(2)}$  = coefficient indicating the age effect on the logit  $P_i$ , adjusted for group status;  $i = 1, 2$
- $\varepsilon_i$  = error term for logit  $P_i$ ,  $i = 1, 2$ .

This model assumes independent multinomial sampling.

Because the interpretation of each logistic modeling function is similar, consider the logit  $P_1$  and suppose  $X_1$  is a binary covariate ( $X_1 = 1$  for Ranch Hands or  $X_1 = 0$  for Comparisons). The term  $\exp(\beta_{1(1)})$  equals the adjusted odds ratio of low versus normal for Ranch Hands ( $X_1 = 1$ ) compared to Comparisons ( $X_1 = 0$ ). If the probability of being low is small compared to being normal for both the Ranch Hand and Comparison groups, the odds ratio of low versus normal is approximately equal to the relative risk of being low between the two groups. If  $X_1$  is a continuous covariate,  $\exp(\beta_{1(1)})$  represents the adjusted odds ratio of outcome 1 versus outcome 0 for a unit increase in  $X_1$ .

The abnormal and normal categorizations for many of the discrete analyses were defined by categorizing laboratory and physical examination measures according to laboratory and clinic reference values. Cutpoints for the dependent variables erythrocyte sedimentation rate, cholesterol, and total testosterone were age-dependent. Consequently, normal and abnormal levels were constructed according to a participant's laboratory value and age at the physical examination.

### 7.4.2 Modeling Strategy

In general, based on one of the adjusted analysis models described in Table 7-5, a model for dependent variables was based on the exposure effect (group or dioxin) and medically relevant covariates, as identified in Chapters 9 through 18 for each clinical category. As described previously, body fat at the time of the blood measurement of dioxin was included in Models 2 and 3.

The general modeling strategy did not remove any covariates from the model; however, the modeling strategy for the adjusted analysis of dependent variables in certain clinical areas was modified as necessary because of the large number of covariates or sparse number of participants with abnormal measurements. Stepwise elimination of covariates was conducted to allow for proper estimation of model parameters. When this strategy of removing covariates was necessary, the covariates removed from (or retained in) a model for a given health endpoint and model were specified in footnotes to the tables.

### 7.4.3 Longitudinal Analysis

Selected longitudinal analyses were performed to investigate changes in health status between 1982 and 1997 for Models 1, 2, and 3 as a function of dioxin exposure. Model 4 was not examined in longitudinal analyses because lipid-adjusted dioxin, the estimate of exposure in this model, changes over time and was not available for all participants in 1982 or 1997. All three models were adjusted for age at the time of the 1997 follow-up physical examination. Age was a well-known risk factor for nearly all clinical areas, and although Ranch Hands and Comparisons were matched on age, the estimates of dioxin exposure in Models 2 and 3 were not.

In the longitudinal analysis of discrete variables, only those participants whose health was classified as normal in 1982 were included in the analysis of the participants' health at the 1997 follow-up examination. Participants classified as "abnormal" in 1982 were excluded because the focus of the analysis was to investigate the temporal effects of dioxin exposure between 1982 and 1997. Participants classified as "abnormal" in 1982 were already abnormal before this period; consequently, only participants classified as "normal" at the 1982 examination were considered to be at risk when the effects of dioxin over time were explored. The rate of abnormalities under this restriction approximated the cumulative incidence rate between 1982 and 1997 (22).

The dependent variable in this type of analysis was the health of participants at the 1997 examination whose health was normal in 1982. The independent variables were the appropriate exposure estimate and age at the time of the 1997 follow-up physical examination. The analyses of Models 2 and 3 also were adjusted for body fat at the time of the blood measurement of dioxin. Tabular displays of the longitudinal analysis results of discrete dependent variables include summary statistics for 1982 and 1997, as well as 1985, 1987, and 1992 summaries, if available. The results of the statistical analysis restricted to those participants who were normal in 1982 also were provided.

In the longitudinal analysis of continuous variables, a general linear model approach, as explained in Table 7-6, was used. The dependent variable was the difference between the 1997 measurement and the 1982 measurement. This difference, measuring the change in the endpoint over this period of time, was modeled as a function of the estimate of exposure (group or dioxin), the participant's age at the time of the 1997 follow-up physical examination, and the 1982 measurement of the continuous dependent variable. The analyses of Models 2 and 3 also were adjusted for body fat at the time of the blood measurement of dioxin. The reasons for using the health endpoint measurement in 1982 for longitudinal analysis of continuous variables were as follows:

- A linear relation between measurements of the dependent variable in 1982 and 1997 because of a difference in measuring devices was accounted for by using the 1982 measurement as an independent variable.
- The difference between two measurements taken over a period of time was generally correlated with the first measurement (23).
- The relation between the difference of the 1997 and 1982 measurements and the estimate of exposure may be confounded with the 1982 measurement, especially if the endpoint and the estimate of exposure were related.

Tabular displays of the results of longitudinal analysis of continuous dependent variables include summary statistics for 1982 and 1997, as well as 1985, 1987, and 1992 summaries, if available. Results of the statistical analysis relating the difference in the 1997 and 1982 measurements to the estimate of exposure also were provided. For some variables, 1985 clinical measurements were substituted for 1982 measurements because the variable was not analyzed at the 1982 examination or was inherently different from the 1997 variable due to differing clinical methods.

## **7.5 INTERPRETIVE CONSIDERATIONS**

Several specific issues to consider when interpreting the results found in this report are discussed in this section. The issues discussed here include adjustments for covariates, multiple testing, trends in the results of endpoints within a clinical area, the proportion of variation explained by the model ( $R^2$ ), interpretation of discrete and continuous analyses of a health endpoint, and statistical power to detect the effects of dioxin.

### 7.5.1 Adjustments for Covariates

In contrasts between all Ranch Hands and all Comparisons (Model 1), the matching variables age, race, and occupation were effectively eliminated as confounders. The initial and 1987 dioxin analyses within Ranch Hands (Models 2 and 4) and the categorized dioxin analysis within Ranch Hands and Comparisons (Model 3) did not benefit from the matched design. For example, military occupation was a strong confounder because it is highly correlated with dioxin levels in Ranch Hands and is related to some health variables through socioeconomic differences between officers and enlisted personnel. Education was highly associated with military occupation and certain psychometric results. Consequently, with the exception of a few analyses where the prevalence or history of abnormal results was sparse, all health endpoints were analyzed with and without adjustment for clinically relevant covariates.

### 7.5.2 Multiple Testing

Numerous dependent variables were considered because of the lack of a predefined medical endpoint. Each dependent variable was analyzed in many different ways to accommodate covariate information and different statistical models. Under the hypothesis of no relation between physical health and dioxin, approximately 5 percent of the many statistical tests (group or dioxin effects) in this report detected an association between group or dioxin and health (p-values  $\leq 0.05$ ). Observing significant results because of multiple testing, even when there is no relation between dioxin and health, is known as the multiple-comparisons problem (24) and is common in all large studies with multiple endpoints. It is generally difficult to distinguish between those statistically significant results that arise because of the multiple testing artifact and those that may be due to an actual dioxin effect. In order to weigh and interpret the findings, the strength of the association, consistency, dose-response patterns, and biologic plausibility were considered.

### 7.5.3 Trends

Assessing consistent and meaningful trends is essential when interpreting any comprehensive study with multiple endpoints, clinical areas, and covariates; however, caution must be used. Increased numbers of abnormalities or mean values with increased dioxin levels across medically related variables within a clinical area might indicate a group or dioxin effect. There may, however, be a moderate-to-strong correlation between these endpoints, where a change in one variable leads directly to a change in the other. Hence, the strength of the trends also was considered when assessing the suspected association.

### 7.5.4 Interpretation of the Coefficient of Determination

The coefficient of determination ( $R^2$ ) measures the proportionate reduction of the total variation in a continuously distributed health variable,  $y$ , associated with the set of independent variables in a linear regression. A large value of  $R^2$  does not necessarily imply that the fitted model is a useful one. Large values of  $R^2$  would occur, for example, if  $y$  is regressed on an independent variable with only a few observed values. On the other hand, small values of  $R^2$  are generally seen in observational studies because little or no control has been applied in the assignment of the values of the “treatment” (dioxin) or the conditions under which the “treatment” has been applied. In this study, the dioxin measurements were taken many years after exposure and are subject to some measurement error. Thus, in most analyses, the values of  $R^2$  were small.

### 7.5.5 Clinical Interpretation of Discrete versus Continuous Data

Small but significant mean differences in a continuously measured health variable (e.g., alkaline phosphatase) between exposed and unexposed groups when there are no corresponding differences in the percentage of abnormal tests are difficult to interpret in any study. In this study, significant differences in the means between exposed and unexposed groups sometimes are observed without a corresponding difference between the groups in the percentage of participants with an abnormal measurement. Such contrasting situations may be interpreted as spurious outcomes of no clinical consequence, or as a subclinical dioxin effect. Significant trends in the mean with increasing levels of dioxin were interpreted as a dioxin-related effect if a corresponding trend was seen in the proportion above or below the normal range or if the trend was consistent with other findings.

### 7.5.6 Power

A type I error is making a false conclusion that an association (group or dioxin effect) exists when there is no association. The other possible inference error, a type II error, is the failure to detect an association when one actually exists. The power of a statistical test is 1 minus the probability of a type II error. The power of the test is the probability that the test will reject the hypothesis of no group or dioxin effect when an effect does in fact exist.

The fixed size of the Ranch Hand cohort limits the ability of this study to detect some group or dioxin associations if they exist. This limitation is most obvious for specific types of cancer, such as soft tissue sarcoma (STS) and non-Hodgkin’s lymphoma (NHL). These conditions are so uncommon that fewer than two cases are expected in this study, indicating that there is virtually no statistical power to detect low-to-moderate associations between dioxin and cancer. In an attempt to overcome the lack of power to detect group differences for specific types of systemic cancer, for example, all types of systemic cancer were combined into a single variable. It is still possible, however, that an increased risk could exist for a particularly rare type of cancer, allowing that increased risk to be missed in this study.

Table 7-7 and Appendix Tables E-1 through E-3 contain the approximate power at a significance level of 0.05 to detect specified relative risks for a given prevalence rate of a discrete dependent variable. Table 7-7 presents power calculations for Model 1 (group), and Appendix Tables E-1 through E-3 present power calculations for Model 2 (initial dioxin), Model 3 (categorized dioxin—low plus high Ranch Hand versus Comparison contrast), and Model 4 (lipid-adjusted 1987 dioxin). Power calculations were performed using the logarithm (base 2) of dioxin in Models 2 and 4, and consequently, the relative risk is for a twofold increase in dioxin. The power of a test for a discrete variable depends on the significance level, actual relative risk, prevalence of the condition, and the Ranch Hand and Comparison sample sizes (for Models 1 and 3) or the distribution of the dioxin data (for Models 2 and 4).

As an example, using age-adjusted incidence rates for all U.S. males (based on data from the Surveillance Epidemiology and End Results program of the National Cancer Institute), prevalence rates for all cancers, NHL, and STS were estimated as 0.07, 0.002, and 0.001, respectively. Thus, Table 7-7 shows a power less than 0.21 to detect a relative risk of 2.0 for the estimated prevalences of NHL and STS. For a disease with a prevalence of 0.05, the power to detect a relative risk of 1.5 would be 0.54.

**Table 7-7. Approximate Power To Detect a Group Effect at a 5-Percent Level of Significance (Discrete Dependent Variable)**

Prevalence of Condition	Relative Risk								
	1.10	1.20	1.30	1.40	1.50	1.75	2.00	10.00	20.00
0.005	0.05	0.06	0.07	0.09	0.10	0.15	0.21	0.92	0.97
0.01	0.06	0.07	0.09	0.12	0.16	0.26	0.36	1.00	1.00
0.02	0.06	0.09	0.14	0.19	0.26	0.45	0.62	1.00	1.00
0.03	0.07	0.11	0.18	0.26	0.36	0.60	0.79	1.00	1.00
0.04	0.07	0.13	0.22	0.33	0.45	0.72	0.89	1.00	1.00
0.05	0.08	0.15	0.26	0.40	0.54	0.81	0.94	1.00	1.00
0.10	0.10	0.24	0.44	0.64	0.80	0.97	1.00	1.00	1.00
0.15	0.12	0.32	0.58	0.79	0.92	1.00	1.00	1.00	1.00
0.20	0.14	0.38	0.67	0.87	0.96	1.00	1.00	1.00	1.00

Table 7-8 and Appendix Tables E-4 through E-6 provide the same information on power as Table 7-7 and Appendix Tables E-1 through E-3 for a continuous dependent variable at a significance level of 0.05. The power calculations are defined in terms of the coefficient of variation (100 times the standard deviation of the dependent variable divided by the mean of the dependent variable) and the proportion mean change. The coefficient of variation relates the spread of the data relative to the magnitude of the data. In general, the power of a test is greater when the coefficient of variation is smaller. Table 7-8 presents power calculations for Model 1 (group), and Appendix Tables E-4 through E-6 present power calculations for Model 2 (initial dioxin), Model 3 (categorized dioxin—low plus high Ranch Hand versus Comparison contrast) and Model 4 (lipid-adjusted 1987 dioxin). Power calculations were performed using the logarithm (base 2) of dioxin in Models 2 and 4, and consequently, the relative risk is for a twofold increase in dioxin. The power of a test for a continuous variable depends on the significance level, actual difference in the true dependent variable means or slope of the dioxin coefficient, variation in the dependent variable data, sample size, and the distribution of the dioxin data if dioxin is the exposure estimate.

The proportion mean change in Table 7-8 and Appendix Table E-5 is defined as the difference in the true Ranch Hand and Comparison means, relative to the combined average of the two groups, assuming no transformation of the dependent variable. The proportion mean change in Appendix Tables E-4 and E-6 is defined as the change in the expected value (mean) of the dependent variable for a twofold increase in initial dioxin, relative to the dependent variable mean. The proportion mean change in Appendix Tables E-4 and E-6 corresponds mathematically to the slope of initial or 1987 dioxin divided by the dependent variable mean, assuming no transformation of the dependent variable. Analogous quantities can be derived based on transformed statistics. As an example, white blood cell count (on the natural logarithm scale) for all participants has a coefficient of variation of approximately 15 percent. With this coefficient of variation, for the 870 Ranch Hands and 1,251 Comparisons in Model 1, the power is approximately 0.86 for detecting a 2-percent increase in the mean white blood cell count of Ranch Hands relative to the mean white blood cell count of Comparisons (mean change = 0.02).

**Table 7-8. Approximate Power To Detect a Group Effect at a 5-Percent Level of Significance (Continuous Dependent Variable)**

Mean Change	Coefficient of Variation ( $100\sigma/\mu$ )					
	5	10	15	25	50	75
0.005	0.62	0.21	0.12	0.07	0.06	0.05
0.01	0.99	0.62	0.33	0.15	0.08	0.06
0.02	1.00	0.99	0.86	0.44	0.15	0.09
0.03	1.00	1.00	0.99	0.78	0.27	0.15
0.04	1.00	1.00	1.00	0.95	0.44	0.23
0.05	1.00	1.00	1.00	0.99	0.62	0.33
0.10	1.00	1.00	1.00	1.00	0.99	0.86

In summary, this study has good power to detect relative risks of 2.0 or more with respect to diseases, such as heart disease and basal cell carcinoma, occurring at a prevalence of at least 5 percent in unexposed populations. In addition, the study size is sufficient to detect small mean shifts in the continuously distributed variables. The detection of significant mean shifts without a corresponding indication of increased Ranch Hand abnormalities or disease may be an artifact of multiple testing, could represent a subclinical effect, or could be of little or no medical importance.

## 7.6 EXPLANATION OF TABLES

This section explains the contents of the tables used to report the results of the analyses for continuous and discrete dependent variables (two levels and more than two levels). Selected tables from the General Health Assessment (Chapter 9) and the Hematology Assessment (Chapter 15) will be referenced throughout this discussion. The contents of each table depend on the form of the health status endpoint (i.e., whether the dependent variable under analysis is a continuous or discrete variable). A discussion of the contents of exposure analysis tables is discussed first, followed by an explanation of the longitudinal analysis tables.

### 7.6.1 Exposure Analysis

The results of the exposure analysis are displayed in subpanels within each table as specified in Table 7-9. The specification of the subpanels is applicable whether the dependent variable is continuous or discrete.

**Table 7-9. Location of Table Results from Different Exposure Analysis Models**

Model	Exposure Estimate	Subpanel in Table	Type of Analysis
1	Group <sup>a</sup>	a	Unadjusted
		b	Adjusted
2	Initial Dioxin <sup>b</sup>	c	Unadjusted
		d	Adjusted
3	Categorized Dioxin <sup>a</sup>	e	Unadjusted
		f	Adjusted
4	1987 Dioxin <sup>b</sup>	g	Unadjusted
		h	Adjusted

<sup>a</sup> Ranch Hands and Comparisons.

<sup>b</sup> Ranch Hands only.

#### 7.6.1.1 Continuous Variables

Table 9-8 in the General Health Assessment chapter presents an example of the results of the analysis when the dependent variable was continuous. Subpanels (a) and (b) show the results of unadjusted and adjusted Model 1 analyses that compared the Ranch Hand and Comparison means of a dependent variable. Contrasts between Ranch Hands and Comparisons also are presented within each occupational category (i.e., officer, enlisted flyer, and enlisted groundcrew).

For the unadjusted analysis in subpanel (a), a sample size (n) and a mean are presented for all occupational categories combined and separately for each occupational category. If the dependent variable was transformed for the analysis, the means of the transformed values were converted to the original scale and the column heading is footnoted. For each contrast of Ranch Hands versus Comparisons, the difference of means on the original scale and the associated 95-percent confidence interval are reported. The 95-percent confidence interval was constructed by adding and subtracting 1.96 multiplied by the standard error (for the upper and lower bounds, respectively) to the estimated mean. If the analysis was performed on a transformed scale, the 95-percent confidence interval on the differences of means is not presented and the column is footnoted. When presenting results from analyses of means based on log-transformed (or square root-transformed) data, means were converted back to original units. Conversion of the standard deviation from log units to original units is not recommended (25); therefore, confidence intervals for mean differences in original units are not presented. A p-value also is reported to determine whether a difference in means on the scale used for analysis for a specified contrast was equal to zero. The confidence interval and p-value for each occupational category were determined using analysis of variance techniques from a group-by-occupation interaction in the model. The group-by-occupation interaction was used to determine the model coefficients and standard errors simultaneously for officers, enlisted flyers, and enlisted groundcrew. The respective coefficients and standard errors from the group and group-by-interaction terms in the model, along with the covariances between the estimates, were combined as appropriate to construct the confidence intervals and p-values for the three occupational strata.

For an adjusted Model 1 analysis, subpanel (b) includes a sample size, an adjusted mean, a difference of Ranch Hand and Comparison adjusted means on the original scale, the associated 95-percent confidence interval (if the analysis was performed on the original scale), and a p-value for each contrast. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. The confidence interval and p-value for each occupational category were determined using analysis of covariance techniques from a group-by-occupation interaction in the model.

Subpanel (c) of Table 9-8 reports summary statistics from the analysis that assessed the association between the continuous dependent variable and initial dioxin (Model 2) without adjusting for covariate information. The sample size and mean of the dependent variable (transformed to the original units, if necessary) are presented for low, medium, and high categories of initial dioxin. The low, medium, and high categories were determined by dividing all Ranch Hands with initial dioxin estimates into three approximately equal-sized categories based on their initial dioxin estimate. The numerical values defining these categories are specified in a table subpanel footnote. Means of the dependent variable, adjusted for percent body fat at the time of the blood measurement of dioxin, also are presented for the low, medium, and high categories of initial dioxin. Based on a linear regression analysis, adjusted for percent body fat at the time of the blood measurement of dioxin, the coefficient of determination ( $R^2$ ), the estimated slope, and its associated standard error are reported. If the dependent variable was transformed for the regression analysis, the transformation is identified in the footnote. The p-value associated with testing whether the slope was equal to zero also is presented. The summary statistics that are reported were based on initial dioxin divided into three categories, whereas the  $R^2$ , slope, standard error, and p-value were based on  $\log_2$  (initial dioxin) in its continuous form.

Based on analyses that incorporate covariate information, subpanel (d) reports summary statistics from the analysis that assessed the association between the continuous dependent variable and initial dioxin (Model 2). Similar to the unadjusted analysis, a sample size and adjusted mean of the dependent variable (transformed to the original units, if necessary) are presented for low, medium, and high categories of initial dioxin. The numerical values defining these categories are specified in a table subpanel footnote.

Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. Based on the multiple linear regression of the dependent variable on  $\log_2$  (initial dioxin) and covariate effects, including percent body fat at the time of the blood measurement of dioxin, the coefficient of determination ( $R^2$ ), the adjusted slope for  $\log_2$  (initial dioxin), and its associated standard error are reported. If the dependent variable was transformed for the regression analysis, the adjusted means, adjusted slope, and standard error are footnoted and the transformation is identified in the footnote. The p-value for testing whether the slope was equal to zero also is presented.

Subpanels (e) and (f) of Table 9-8 show the results of unadjusted and adjusted Model 3 analyses that contrasted the means of a continuous dependent variable for Ranch Hands with background, low, high, and low plus high dioxin levels with Comparisons having lipid-adjusted dioxin levels less than or equal to 10 ppt. The low and high Ranch Hand categories were determined by dividing all Ranch Hands with lipid-adjusted dioxin estimates greater than 10 ppt into two approximately equal-sized categories based on their initial dioxin estimate. The low plus high Ranch Hand category is a combination of the low and high categories. The note at the bottom of the table subpanels defines the dioxin categories. The mean for the low plus high category is a weighted average (transformed to the original units, if necessary) of the low Ranch Hand and high Ranch Hand categories' means on the scale used for transformation, where the weights were based on the low and high Ranch Hand categories' sample sizes. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

For the unadjusted analysis in subpanel (e), a sample size and dependent variable mean are presented for each category. If the dependent variable was transformed for the analysis, the means of the transformed values were converted to the original scale and the column heading is footnoted. The mean of the dependent variable adjusted for percent body fat at the time of the blood measurement of dioxin also is presented for each dioxin category. For each individual contrast of the Ranch Hand category versus the Comparison category, the difference of means on the original scale and the associated 95-percent confidence interval are reported. If the analysis was performed on a transformed scale, the 95-percent confidence interval on the differences of means is not presented and the column is footnoted. A p-value also is reported to determine whether a difference in means for a specified contrast was equal to zero. The p-value was based on the difference of means on the scale used for analysis. The adjusted mean, confidence interval, and p-value for each contrast was determined from an analysis of covariance model with adjustment for percent body fat at the time of the blood measurement of dioxin.

For the adjusted analysis in subpanel (f), the table includes a sample size, an adjusted mean (adjusted for percent body fat at the time of the blood measurement of dioxin and covariates), a difference in adjusted means on the original scale, and a 95-percent confidence interval on the difference in adjusted means (if the analysis was performed on the original scale). The p-value for testing whether the difference in adjusted means for a specified contrast was equal to zero also is presented.

Subpanel (g) of Table 9-8 reports summary statistics from Model 4 analyses, which assessed the association between the continuous dependent variable and 1987 dioxin without adjusting for covariate information. The sample size and mean of the dependent variable (transformed to the original units, if necessary) are presented for low, medium, and high categories of 1987 dioxin. The low, medium, and high categories were determined by dividing all Ranch Hands with 1987 dioxin levels into three approximately equal-sized categories based on their 1987 dioxin measurement. The numerical values defining the low, medium, and high categories of 1987 dioxin are specified in a table subpanel footnote. Based on a linear regression of the dependent variable on  $\log_2$  (1987 dioxin + 1), the coefficient of determination ( $R^2$ ), the estimated slope, and its associated standard error are reported for each model. A

value of 1 was added to each measurement because of the presence of 1987 dioxin measurements of 0 ppt. If the dependent variable was transformed for the regression analysis, the means, slope, and standard error are footnoted and the transformation is identified in the footnote. The p-value associated with testing whether the slope was equal to zero also is presented.

Based on analyses that incorporate covariate information, subpanel (h) reports summary statistics for Model 4 analyses that assessed the association between the continuous dependent variable and 1987 dioxin. The sample size and adjusted mean of the dependent variable (transformed to the original units, if necessary) are presented for low, medium, and high categories of 1987 dioxin. The numerical values defining these categories are specified in a table subpanel footnote. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. Based on the multiple linear regression of the dependent variable on  $\log_2(1987 \text{ dioxin} + 1)$  and covariates, the coefficient of determination ( $R^2$ ), the adjusted slope for  $\log_2(1987 \text{ dioxin} + 1)$ , and its associated standard error are reported for each model. If the dependent variable was transformed for the regression analysis, the adjusted means, adjusted slope, and standard error are footnoted and the transformation is identified in the footnote. The p-value for testing whether the slope was equal to zero also is presented.

### 7.6.1.2 *Discrete Variables*

#### 7.6.1.2.1 *Discrete Variable with Two Categories*

Table 9-3 in the General Health Assessment chapter presents an example of the results of analysis when the dependent variable is discrete and dichotomous. Subpanels (a) and (b) display the results of unadjusted and adjusted Model 1 analyses that compared the percentage of Ranch Hands and Comparisons that were considered abnormal for the dependent variable of interest (the abnormal classification for self-perception of health in Table 9-3 is “fair or poor”). Contrasts between Ranch Hands and Comparisons also are presented within each occupational category (i.e., officer, enlisted flyer, and enlisted groundcrew). For the unadjusted analysis in subpanel (a), a sample size and the number and percentage of participants considered abnormal are presented for each group within each occupational category. For the contrasts of Ranch Hands versus Comparisons, an estimated relative risk, an associated 95-percent confidence interval on the relative risk, and a p-value for testing whether the risk was equal to 1.0 are presented. The normal distribution was used to calculate an approximate 95-percent confidence interval. Results for each occupational category were determined from a group-by-occupation interaction that was included in the model.

For the adjusted analysis of Model 1, as presented in subpanel (b), the table presents an adjusted relative risk, a 95-percent confidence interval on the relative risk, and a p-value for testing whether the risk was equal to 1.0. The adjusted relative risk, confidence interval, and p-value were determined from a multiple logistic regression model that used the appropriate covariates for the clinical area and dependent variable of interest. Results for each occupational category were determined from a group-by-occupation interaction that was included in the model.

Subpanel (c) of Table 9-3 reports summary statistics for analyses that assessed the association between the dependent variable and initial dioxin (Model 2) without adjusting for covariate information. Sample sizes are presented for low, medium, and high categories of initial dioxin. The numerical values defining these categories are specified in a table footnote. The number and percentage of Ranch Hands considered abnormal are presented for the low, medium, and high initial dioxin categories. Based on a logistic regression model, adjusted for percent body fat at the time of the blood measurement of dioxin, an estimated relative risk and its 95-percent confidence interval are reported. The p-value associated

with testing whether the relative risk was equal to 1.0 also is presented. The normal distribution was used to determine an approximate 95-percent confidence interval. The summary statistics that are reported were based on initial dioxin divided into three categories, whereas the relative risk, confidence interval, and p-value were based on  $\log_2$  (initial dioxin) in its continuous form.

Subpanel (d) of Table 9-3 reports summary statistics for analyses that assessed the association between the discrete dependent variable and initial dioxin (Model 2), adjusted for percent body fat at the time of the blood measurement of dioxin and covariate information. The sample size given is based on a multiple logistic regression of the discrete dependent variable on  $\log_2$  (initial dioxin), percent body fat at the time of the blood measurement of dioxin, and covariates. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. The adjusted relative risk for  $\log_2$  (initial dioxin) and its associated 95-percent confidence interval are reported and are based on this multiple logistic regression model. The normal distribution was used to determine an approximate 95-percent confidence interval. The p-value for testing whether the relative risk was equal to 1.0 also is presented.

Subpanels (e) and (f) of Table 9-3 show the results of unadjusted and adjusted Model 3 analyses that contrasted Ranch Hands having background, low, high, and low plus high dioxin levels with Comparisons having lipid-adjusted dioxin levels less than or equal to 10 ppt. The percentage of participants that were considered abnormal for the dependent variable of interest was contrasted between the four categories of Ranch Hands and Comparisons. The low and high Ranch Hand categories were determined by dividing all Ranch Hands with lipid-adjusted dioxin estimates greater than 10 ppt into two approximately equal-sized categories based on their initial dioxin estimate. The low plus high Ranch Hand category is a combination of the low and high Ranch Hand categories. The note at the bottom of the table subpanel defines the dioxin categories. The percentage of Ranch Hands in the low plus high category is a weighted average of the low Ranch Hand and high Ranch Hand categories, where the weights are based on the low category and high category sample sizes. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

For the Model 3 unadjusted analysis in subpanel (e), the sample size and the number and percentage of participants considered abnormal is presented for each dioxin category. For the individual contrasts of the Ranch Hand categories versus Comparisons, an estimated relative risk, a 95-percent confidence interval for the relative risk, and a p-value associated with testing whether the risk was equal to 1.0 are presented. The relative risk, confidence interval, and p-value were determined from a logistic regression model, adjusted for percent body fat at the time of the blood measurement of dioxin. The normal distribution was used to determine an approximate 95-percent confidence interval.

For the Model 3 adjusted analysis, subpanel (f) of the table presents an adjusted relative risk, a 95-percent confidence interval for the relative risk, and a p-value associated with testing whether the risk was equal to 1.0 for the individual contrasts of the Ranch Hand categories with Comparisons. The normal distribution was used to determine an approximate 95-percent confidence interval.

Subpanels (g) and (h) of Table 9-3 present summary statistics from Model 4, which assessed the association between the dependent variable and 1987 dioxin. For the unadjusted analysis, the sample size and the number and percentage of participants considered abnormal is presented for each 1987 dioxin category. The low, medium, and high categories were determined by dividing all Ranch Hands with 1987 dioxin levels into three approximately equal-sized categories. The numerical values defining these categories are specified in a table footnote. Based on a logistic regression model, an estimated relative risk and its 95-percent confidence interval are reported. The p-value associated with testing

whether the relative risk was equal to 1.0 also is presented. The normal distribution was used to determine an approximate 95-percent confidence interval. The summary statistics are reported for 1987 dioxin divided into three categories, whereas the relative risk, confidence interval, and p-value were based on  $\log_2(1987 \text{ dioxin} + 1)$  in its continuous form.

Incorporating covariate information, subpanel (h) reports summary statistics from analyses that assessed the association between the dichotomous dependent variable and 1987 dioxin. The sample size is presented for a multiple logistic regression of the discrete dependent variable on  $\log_2(1987 \text{ dioxin} + 1)$  including covariates in the final adjusted model. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. Based on the multiple logistic regression model, the adjusted relative risk for  $\log_2(1987 \text{ dioxin} + 1)$  and its associated 95-percent confidence interval are reported. The normal distribution was used to determine an approximate 95-percent confidence interval. The p-value for testing whether the relative risk was equal to 1.0 also is presented.

#### *7.6.1.2.2 Discrete Variable with More Than Two Categories*

Polytomous regression techniques were used to analyze discrete dependent variables having more than two levels (e.g., abnormal low, normal, abnormal high—see Table 15-4 in the Hematology Assessment chapter). Results were presented in a similar fashion to discrete variables with only two categories; however, the number and percentage of participants for each dependent variable category (including normal) are given. Therefore, the relative frequencies sum to 100 percent across the dependent variable categories and the number of participants in each of the dependent variable categories adds to the total number of participants in each exposure group or dioxin category. In addition, a relative risk, a 95-percent confidence interval, and a p-value were presented for each contrast with the normal level of the dependent variable (e.g., abnormal low versus normal and abnormal high versus normal).

In Table 15-4, subpanels (a) and (b) display the results of unadjusted and adjusted Model 1 analyses that compared Ranch Hands and Comparisons on the relative frequencies of each abnormal level for a specified discrete dependent variable. For example, the percentage of participants with an abnormally high red blood cell count was contrasted to participants with a normal red blood cell count, and the percentage of participants with an abnormally low red blood cell count was contrasted to participants with a normal red blood cell count. Contrasts between Ranch Hands and Comparisons also are presented within each occupational category (i.e., officer, enlisted flyer, and enlisted groundcrew). For the unadjusted analysis in subpanel (a), a sample size is presented for each exposure group (Ranch Hand, Comparison) across all occupational categories and within each occupational category. The number and percentage of participants are presented for each level of the dependent variable for each group. For the contrasts of Ranch Hands versus Comparisons, an estimated relative risk, a 95-percent confidence interval for the relative risk, and a p-value associated with testing whether the risk was equal to 1.0 are presented for each contrast against the normal level of the dependent variable (e.g., abnormal low versus normal and abnormal high versus normal). The normal distribution was used to calculate an approximate 95-percent confidence interval. Results for each occupational category were determined from the group-by-occupation interaction that was included in the model.

For a Model 1 analysis adjusted for covariate information and shown in subpanel (b), the table presents an adjusted relative risk, a 95-percent confidence interval on the relative risk, and a p-value associated with testing whether the risk was equal to 1.0 for each occupational category and each contrast. The normal distribution was used to calculate an approximate 95-percent confidence interval. Results for

each occupational category were determined from the group-by-occupation interaction that was included in the model.

Subpanels (c) and (d) of Table 15-4 summarize the unadjusted and adjusted Model 2 analyses relating discrete dependent variables having more than two categories to initial dioxin. Both unadjusted and adjusted analyses are adjusted for percent body fat at the time of the blood measurement of dioxin. In subpanel (c), the sample size and the number and percentage of Ranch Hands in each category of the dependent variable are presented for each initial dioxin category (i.e., low, medium, and high initial dioxin). The relative risk, the 95-percent confidence interval for the relative risk, and the p-value associated with testing whether the risk was equal to 1.0 are presented for each abnormal level of the dependent variable (e.g., abnormal low versus normal and abnormal high versus normal). The summary statistics that are reported were based on initial dioxin divided into three categories, whereas the relative risk, confidence interval, and p-value were based on  $\log_2$  (initial dioxin) in its continuous form.

In subpanel (d), after adjustment for covariate information, the sample size, the adjusted relative risk, the 95-percent confidence interval for the relative risk, and the p-value associated with testing whether the risk was equal to 1.0 are presented for each abnormal level of the dependent variable. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

Subpanels (e) and (f) of Table 15-4 present unadjusted and adjusted Model 3 analyses of categorized dioxin versus a discrete dependent variable having more than two categories. Both unadjusted and adjusted analyses are adjusted for percent body fat at the time of the blood measurement of dioxin. Results are presented in a similar fashion to the group analysis (Model 1), except that contrasts involve the four Ranch Hand categories (background, low, high, and low plus high) versus Comparisons, and contrasts are not performed for each occupation. For the unadjusted analysis, a sample size is presented for each dioxin category. The low plus high Ranch Hand category is a combination of the low and high Ranch Hand categories. The percentage of Ranch Hands in the low plus high category is a weighted average of the low Ranch Hand and high Ranch Hand categories, where the weights are based on the low category and high category sample sizes. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

The number and percentage of participants for each level of the dependent variable are presented for each dioxin category in subpanel (e). For each contrast of a Ranch Hand category versus the Comparison group, an estimated relative risk, a 95-percent confidence interval for the relative risk, and a p-value associated with testing whether the risk was equal to 1.0 are presented. These results are given for each contrast against the normal level of the dependent variable (e.g., abnormal low versus normal and abnormal high versus normal). For an adjusted Model 3 analysis in subpanel (f), the table presents an adjusted relative risk, a 95-percent confidence interval on the relative risk, and a p-value for each contrast of Ranch Hands versus Comparisons for each abnormal level of the dependent variable.

Similar to the polytomous regression analysis using initial dioxin, unadjusted and adjusted analyses of discrete dependent variables with more than two categories were performed using 1987 dioxin in Model 4. In Table 15-4, summaries of the analyses are given in subpanels (g) and (h). For the unadjusted analysis in subpanel (g), sample sizes are presented for each 1987 dioxin category (i.e., low, medium, and high 1987 dioxin). The number and percentage of Ranch Hands for each dependent variable category for each 1987 dioxin category are presented. An estimated relative risk, a 95-percent confidence interval on the relative risk, and an associated contrast p-value are reported for each abnormal level of the dependent variable (e.g., abnormal low vs. normal and abnormal high vs. normal). The summary statistics that are

reported were based on 1987 dioxin divided into three categories, whereas the relative risk, confidence interval, and p-value were based on  $\log_2(1987 \text{ dioxin} + 1)$  in its continuous form.

Adjusted analysis results in subpanel (h) include a total sample size, an adjusted relative risk, a 95-percent confidence interval on the relative risk, and an associated contrast p-value for each abnormal level of the dependent variable. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

### 7.6.2 Longitudinal Analysis

The results of the longitudinal analysis are displayed in subpanels within each table as specified in Table 7-10. The specification of the subpanels is applicable whether the dependent variable is continuous or discrete.

**Table 7-10. Location of Table Results from Different Longitudinal Analysis Models**

Model	Exposure Estimate	Subpanel in Table
1	Group <sup>a</sup>	a
2	Initial Dioxin <sup>b</sup>	b
3	Categorized Dioxin <sup>a</sup>	c

<sup>a</sup> Ranch Hands and Comparisons.

<sup>b</sup> Ranch Hands only.

Most of the longitudinal analyses in this report are based on a comparison of data from the 1982 baseline examination and the 1997 follow-up examination, and the discussion of tables below is based on the comparison of the 1982 and 1997 examinations. Some analyses, however, are based on a comparison of data from the 1985 follow-up examination and the 1997 follow-up examination (e.g., neurological indices in Chapter 11, Neurological Assessment, or Doppler pulses in Chapter 14, Cardiovascular Assessment). The 1985 follow-up examination data were used because of methodological differences in the measurements between the 1982 baseline examination and the 1985 follow-up examination, or because the measurement was not obtained at the 1982 baseline examination. In addition, spirometry measurements were not taken at the 1985 follow-up examination, and Doppler pulse measurements were not made at the 1987 follow-up examination; therefore, summary statistics based on data from the respective examinations are not provided for these variables.

#### 7.6.2.1 Continuous Variables

Table 9-15 in the General Health Assessment chapter presents an example of a longitudinal analysis when the dependent variable was continuous. In subpanel (a), a mean and a sample size (n) are provided for all occupational categories combined and separately for each occupational category (i.e., officer, enlisted flyer, and enlisted groundcrew). The mean and sample size are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended both examinations. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination,

and the respective follow-up examination that was summarized. The summary statistics for the 1985, 1987, and 1992 follow-up examinations are provided for reference purposes. If the dependent variable was transformed for the analysis, the means of the transformed values were converted to the original scale and the transformation is specified in a footnote.

Subpanel (a) shows the Ranch Hand and Comparison difference in means between the 1997 follow-up examination and 1982 baseline examination. The Ranch Hand and Comparison difference in means between the 1997 follow-up examination and 1982 baseline examination is presented for all occupations combined and separately for each occupational category. The difference between Ranch Hands and Comparisons in the change between the 1997 follow-up examination mean and the 1982 baseline examination mean also is reported in subpanel (a). The p-value that was used to determine whether the difference in the examination mean change between Ranch Hands and Comparisons was equal to zero is given. This p-value was based on the difference in Ranch Hand and Comparison examination mean changes on the scale used for analysis. The p-value for each occupational category was determined using analysis of covariance techniques from a group-by-occupation interaction in the model. The longitudinal analysis performed in subpanel (a) was adjusted for the 1982 measurement of the dependent variable and age at the 1997 physical examination.

Subpanel (b) of Table 9-15 reports summary statistics on the continuous dependent variable of interest. The sample size and mean of the dependent variable (transformed to the original units, if necessary) are presented for low, medium, and high categories of initial dioxin. The low, medium, and high categories were determined by dividing all Ranch Hands with initial dioxin estimates into three approximately equal-sized categories based on their initial dioxin estimate. The numerical values defining these categories are specified in the table subpanel footnote. The mean and sample size are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended both examinations. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination, and the respective follow-up examination that was summarized. If the dependent variable was transformed for the analysis, the transformation is specified in a footnote.

For each participant who attended both the 1982 and 1997 physical examinations, a difference between the dependent variable as measured at the 1997 follow-up examination and as measured at the 1982 baseline examination was created. The difference in these two measurements was on the scale used for analysis. The association between the difference in the examination measurements and initial dioxin was determined and adjusted for the 1982 measurement of the dependent variable, age at the 1997 physical examination, and percent body fat at the time of the blood measurement of dioxin. The estimated slope, its associated standard error, and the p-value associated with testing whether the slope was equal to zero are reported in subpanel (b). If the dependent variable was transformed for the regression analysis, the transformation is identified in the footnote. The summary statistics that are reported were based on initial dioxin divided into three categories, whereas the slope, standard error, and p-value were based on  $\log_2$  (initial dioxin) in its continuous form.

Subpanel (c) of Table 9-15 shows the results of Model 3 analyses that contrasted the means of a continuous dependent variable for Ranch Hands with background, low, high, and low plus high dioxin levels with Comparisons having lipid-adjusted dioxin levels less than or equal to 10 ppt. The low and high Ranch Hand categories were determined by dividing all Ranch Hands with lipid-adjusted dioxin estimates greater than 10 ppt into two approximately equal-sized categories based on their initial dioxin estimate. The low plus high Ranch Hand category is a combination of the low and high categories. The

note at the bottom of the table subpanel defines the dioxin categories. The mean for the low plus high category is a weighted average (transformed to the original units, if necessary) of the low Ranch Hand and high Ranch Hand category means on the scale used for transformation, where the weights were based on the low and high Ranch Hand category sample sizes.

In subpanel (c), a mean and a sample size are provided for all Ranch Hand and Comparison dioxin categories. The mean and sample size are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended the 1982 baseline examination and the 1997 follow-up examination. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination, and the respective follow-up examination that was summarized. The summary statistics for the 1985, 1987, and 1992 follow-up examinations are provided for reference purposes. If the dependent variable was transformed for the analysis, the means of the transformed values were converted to the original scale and the transformation is specified in a footnote.

Subpanel (c) shows the Ranch Hand and Comparison difference in dioxin category means between the 1997 follow-up examination and 1982 baseline examination. The Ranch Hand and Comparison difference in dioxin category means between the 1997 follow-up examination and 1982 baseline examination is presented for all occupations combined and separately for each occupational category. The difference between Ranch Hands and Comparisons in the change between the 1997 follow-up examination mean and the 1982 baseline examination mean also is reported in subpanel (c). The p-value that was used to determine whether the difference in the examination mean change between the Ranch Hand dioxin category and Comparisons was equal to zero is given. This p-value was based on the difference in Ranch Hand and Comparison examination mean changes on the scale used for analysis. The p-value for each occupational category was determined using analysis of covariance techniques. The longitudinal analysis performed in subpanel (c) was adjusted for the 1982 measurement of the dependent variable, age at the 1997 physical examination, and percent body fat at the time of the blood measurement of dioxin.

#### 7.6.2.2 *Discrete Variables with Two Categories*

Table 9-10 in the General Health Assessment chapter presents an example of the longitudinal analysis when the dependent variable was discrete and dichotomous. In subpanel (a), the number and percentage of participants defined as abnormal and a sample size (n) are provided for all occupational categories combined and separately for each occupational category (i.e., officer, enlisted flyer, and enlisted groundcrew). The summary statistics are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended both examinations. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination, and the respective follow-up examination that was summarized. The summary statistics for the 1985, 1987, and 1992 follow-up examinations are provided for reference purposes.

Subpanel (a) also shows the number of Ranch Hands and Comparisons and the number and percentage of participants considered abnormal at the 1997 examination (the abnormal classification for self-perception of health in Table 9-10 is “fair or poor”). These summary statistics are presented for all occupations combined and separately for each occupational category, and are restricted to participants that were considered normal in 1982 (the normal classification for self-perception of health in Table 9-10 is

“excellent or good”). For the contrasts of Ranch Hands versus Comparisons, a relative risk, an associated 95 percent confidence interval on the relative risk, and a p-value for testing whether the risk was equal to 1.0 are presented. The normal distribution was used to calculate an approximate 95-percent confidence interval. Results for each occupational category were determined from the group-by-occupation interaction that was included in the logistic regression model. The longitudinal analysis performed in subpanel (a) was adjusted for age at the 1997 physical examination.

Subpanel (b) of Table 9-10 reports the number and percentage of participants defined as abnormal and a sample size for low, medium, and high categories of initial dioxin. The low, medium, and high categories were determined by dividing all Ranch Hands with initial dioxin estimates into three approximately equal-sized categories based on their initial dioxin estimate. The numerical values defining these categories are specified in the table subpanel footnote. The summary statistics are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended both examinations. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination, and the respective follow-up examination that was summarized.

Based on a logistic regression model adjusted for age at the 1997 physical examination and percent body fat at the time of the blood measurement of dioxin, the association between the dichotomous dependent variable and initial dioxin was determined. The analysis was restricted to participants that were considered normal in 1982. The relative risk and its 95-percent confidence interval are reported in subpanel (b), along with the p-value associated with testing whether the relative risk was equal to 1.0. The summary statistics that are reported were based on initial dioxin divided into three categories, whereas the relative risk, confidence interval, and p-value were based on  $\log_2$  (initial dioxin) in its continuous form.

Subpanel (c) of Table 9-10, for example, shows the sample size and the number and percentage of participants considered abnormal for Ranch Hands with background, low, high, and low plus high dioxin levels and Comparisons having lipid-adjusted dioxin levels less than or equal to 10 ppt. The low and high Ranch Hand categories were determined by dividing all Ranch Hands with lipid-adjusted dioxin estimates greater than 10 ppt into two approximately equal-sized categories based on their initial dioxin estimate. The low plus high Ranch Hand category is a combination of the low and high categories. The note at the bottom of the table subpanel defines the dioxin categories. The percentage of Ranch Hands in the low plus high category is a weighted average of the low Ranch Hand and high Ranch Hand categories, where the weights are based on the low category and high category sample sizes.

The summary statistics in subpanel (c) are provided for data from the 1982 baseline examination and the 1985, 1987, 1992, and 1997 follow-up examinations. Summary statistics for the 1982 baseline examination and the 1997 follow-up examination were based on participants that attended both examinations. Summary statistics for the 1985, 1987, and 1992 follow-up examinations were based on participants that attended the 1982 baseline examination, the 1997 follow-up examination, and the respective follow-up examination that was summarized.

Subpanel (c) also shows the number of Comparisons and Ranch Hands in each of the dioxin categories for the 1997 physical examination, and the number and percentage of participants considered abnormal at the 1997 examination. The analysis was restricted to participants that were considered normal in 1982. The relative risk and its 95-percent confidence interval are reported, along with the p-value associated with testing whether the relative risk was equal to 1.0. The normal distribution was used to calculate an

approximate 95-percent confidence interval. The longitudinal analysis was based on a logistic regression model and was adjusted for age at the 1997 physical examination and percent body fat at the time of the blood measurement of dioxin.

#### *7.6.2.2.1 Discrete Variable with More Than Two Categories*

An example of a longitudinal analysis on a discrete variable with more than two categories is provided in Table 15-26 in the Hematology Assessment chapter. The statistics provided in this table are identical to the statistics provided for a discrete variable with two categories (e.g., Table 9-10). The tables for a discrete variable with more than two categories have a separate subpanel for each abnormal level of the dependent variable. For example, in Table 15-26, platelet count has three levels: abnormal low, normal, and abnormal high. Subpanels (a1), (b1), and (c1) contrast abnormal low levels of platelet count with normal levels for Models 1, 2, and 3, respectively. Subpanels (a2), (b2), and (c2) contrast abnormal high levels of platelet count with normal levels for Models 1, 2, and 3, respectively. As with the longitudinal analysis on a dichotomous dependent variable, analyses are restricted to participants that were normal in 1982.

## REFERENCES

---

1. Lathrop, G. D., W. H. Wolfe, R. A. Albanese, and P. M. Moynahan. 1984. The Air Force Health Study: An epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: Baseline morbidity study results. NTIS: AD A 138 340. USAF School of Medicine, Brooks Air Force Base, Texas.
2. Lathrop, G. D., S. G. Machado, T. G. Karrison, W. D. Grubbs, W. F. Thomas, W. H. Wolfe, J. E. Michalek, J. C. Miner, and M. R. Peterson. 1987. Epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: First followup examination results. NTIS: AD A 188 262. USAF School of Medicine, Brooks Air Force Base, Texas.
3. Thomas, W. F., W. D. Grubbs, T. G. Karrison, M. B. Lustik, R. H. Roegner, D. E. Williams, W. H. Wolfe, J. E. Michalek, J. C. Miner, and R. W. Ogershok. 1990. An epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: 1987 followup examination results, May 1987 to January 1990. NTIS: AD A 222 573. USAF School of Medicine, Brooks Air Force Base, Texas.
4. Grubbs, W. D., W. H. Wolfe, J. E. Michalek, D. E. Williams, M. B. Lustik, A. S. Brockman, S. C. Henderson, F. R. Burnett, R. G. Land, D. J. Osborne, V. K. Rocconi, M. E. Schreiber, J. C. Miner, G. L. Henriksen, and J. A. Swaby. 1995. The Air Force Health Study: An epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: Final Report. 1992 Followup Examination Results. NTIS: AD A 304 306, 304 308-316. USAF School of Medicine, Brooks Air Force Base, Texas.
5. Roegner, R. H., W. D. Grubbs, M. B. Lustik, A. S. Brockman, S. C. Henderson, D. E. Williams, W. H. Wolfe, J. E. Michalek, and J. C. Miner. 1991. The Air Force Health Study: An epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: Serum Dioxin Analysis of 1987 Examination Results. NTIS: AD A 237 516-24. USAF School of Medicine, Brooks Air Force Base, Texas.
6. Lathrop, G. D., W. H. Wolfe, R. A. Albanese, and P. M. Moynahan. 1982. Epidemiologic investigation of health effects in Air Force personnel following exposure to herbicides: Study Protocol. NTIS: AD A 122 250. USAF School of Aerospace Medicine, Brooks Air Force Base, Texas.
7. Michalek, J. E., J. L. Pirkle, S. P. Caudill, R. C. Tripathi, D. G. Patterson, and L. L. Needham. 1996. Pharmacokinetics of TCDD in veterans of Operation Ranch Hand: 10-year follow-up. *Journal of Toxicology and Environmental Health* 47:209-20.
8. Poiger, H., and C. Schalatter. 1986. Pharmacokinetics of 2,3,7,8-TCDD in man. *Chemosphere* 15:1489-94.
9. Kissel, J. C., and G. M. Robarge. 1988. Assessing the elimination of 2,3,7,8-TCDD from humans with a physiologically based pharmacokinetic model. *Chemosphere* 17:2017-27.
10. Akins, J. R., K. Waldrep, and J. T. Bernert Jr. 1989. The estimation of total serum lipids by a completely enzymatic "summation" method. *Clinica Chimica Acta* 184:219-26.

11. Patterson, D. G., L. Hampton, C. R. Lapeza, W. T. Belser, V. Green, L. Alexander, and L. Needham. 1987. High resolution gas chromatographic/high-resolution mass spectrometric analysis of human-serum on a whole-weight and lipid basis for 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Annals of Chemistry* 59:2000-5.
12. Patterson, D. G., L. L. Needham, J. L. Pirkle, D. W. Roberts, J. Bagby, W. A. Garret, J. S. Andrews Jr., H. Falk, J. T. Bernert, E. J. Sampson, and V. N. Houk. 1988. Correlation between serum and adipose tissue levels of 2,3,7,8-tetrachlorodibenzo-p-dioxin in 50 persons from Missouri. *Archives of Environmental Toxicology* 17:139-43.
13. Knapik, J. J., A. R. L. Burse, and J. A. Vogel. 1983. Height, weight, percent body fat, and indices of adiposity for young men and women entering the Army. *Aviation, Space, and Environmental Medicine* 54:223-31.
14. *SAS/STAT User's Guide*. Version 6, 4th ed. 1990. Cary, North Carolina: SAS Institute, Inc.
15. *SAS/STAT software: Changes and enhancements through release 6.11*. 1996. Cary, North Carolina: SAS Institute, Inc.
16. Feinberg, S. E. 1981. *The analysis of cross-classified data*. Cambridge: MIT Press.
17. Bishop, Y. M. M., S. E. Feinberg, and P. W. Holland. 1975. *Discrete multivariate analysis: Theory and Practice*. Cambridge: MIT Press.
18. Neter, J., and W. Wasserman. 1974. *Applied linear statistical models*. Homewood, Illinois: Richard D. Irwin, Inc.
19. Cox, D. R. 1984. *Analysis of survival data*. London: Chapman and Hall.
20. Hosmer, D., and S. Lemeshow. 1989. *Applied logistic regression*. New York: John Wiley & Sons.
21. Agresti, A. 1990. *Categorical data analysis*. New York: John Wiley & Sons.
22. Mausner, J. S., and A. K. Bahn. 1974. *Epidemiology—An introductory text*. Philadelphia: W. B. Saunders Company.
23. Fleiss, J. L. 1986. *The design and analysis of clinical experiments*. New York: John Wiley & Sons.
24. Snedecor, G. W., and W. G. Cochran. 1980. *Statistical methods*. Ames, Iowa: The Iowa State University Press.
25. Steel, R. G. D., and J.H. Torrie. 1960. *Principles and Procedures of Statistics*. New York: McGraw-Hill Book Company, Inc.