

CHAPTER 4

STATISTICAL METHODS

This chapter summarizes statistical methods that were used for investigating relationships between serum dioxin measurements and health status of Ranch Hands and Comparisons. Current body burden dioxin levels were determined by the Centers for Disease Control (CDC) from serum samples taken from Ranch Hands and Comparisons. A variety of statistical procedures were applied to evaluate the relationships between specific health endpoints and dioxin, as measured from these serum samples.

MODELS AND ASSUMPTIONS

Prior Knowledge Regarding Dioxin

This study presents statistical analyses based on assumptions and models that were conceived in 1988 after the publication of the Ranch Hand dioxin pilot study and half-life substudy. At that time, available data regarding the elimination of dioxin in humans suggested that

- Measurements following the ingestion of dioxin by an individual showed that dioxin elimination appeared to be by first-order mechanisms (1).
- Air Force data on 36 Ranch Hand veterans with dioxin body burdens measured in blood drawn in 1982 and in 1987 produced a median half-life estimate of 7.1 years (2). The lack of correlation between individual half-lives and current dioxin levels supported the first-order elimination assumption.
- Assay results on 932 Ranch Hands and 888 Comparisons showed that the concentrations were lognormally distributed with the Ranch Hand distribution significantly shifted to the right of the Comparison distribution. The Comparison median was 4.2 ppt; the 98th percentile of the Comparison distribution was 10.17 ppt. The Ranch Hand median was 12.8 ppt and the 98th percentile was 168 ppt. Based on these data, levels at or below 10 ppt were considered background.

The term "elimination" denotes the overall removal of dioxin from the body. Some analyses in this report assume that the amount of dioxin in the body (C) decays exponentially with time according to the model $C = I \cdot \exp(-rT)$, where I is the initial level, $r = \log 2/H$, H is the half-life, and T is the time between the end of the Vietnam tour and the dioxin blood draw at the 1987 physical examination; this exponential decay law is termed first-order elimination in this report.

The first-order elimination assumption is not equivalent to assuming a one compartment model for dioxin distribution within the body. While a multicompartment model incorporating body composition and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) binding to tissue receptors would provide a detailed description of dioxin concentrations in different compartments, published multicompartment models for TCDD distribution within the body predict first-order elimination of TCDD, overwhelmingly due to fecal excretion (3). Direct

assessment of the first-order assumption with serial dioxin results taken over many years on a number of exposed individuals has not been, as yet, carried out.

The term "body burden" refers to the serum lipid-weight concentration of TCDD, expressed in parts per trillion (4, 5). The lipid-weight dioxin measurement, also called current dioxin body burden in this report, is a derived quantity calculated from the formula $ppt = ppq \cdot 102.6 / W$, where ppt is the lipid-weight concentration, ppq is the actual 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 (4).

The relationship 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 (6). 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% C.I.: [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 cannot 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.

Fundamental Limitations of the Serum Dioxin Data

There are two evident limitations to the available data:

- 1) While Ranch Hand and ingestion data do not appear to violate a first-order elimination assumption, no serially repeated dioxin assay results taken over many years are available yet with which to evaluate directly the adequacy of the first-order elimination model in humans.
- 2) At this time, it has not been determined whether Ranch Hands with dioxin burdens at or below 10 ppt were exposed and their body burdens had decayed to background levels since their duty in Vietnam or whether they were not exposed at all during their tour in Vietnam.

Health versus Dioxin in Ranch Hands

Because first-order elimination is suggested, but not validated directly in humans, the dioxin versus health relationship was assessed within Ranch Hands using two models. The first model directly depends upon the first-order elimination assumption; the second does not. In combination, these two models circumvent the first fundamental limitation by assessing the dioxin versus health relationship with and without first-order elimination. Table 4-1 shows these two models, their assumptions, advantages, and disadvantages for a continuously distributed health variable y.

In Table 4-1, the phrase "single dioxin dose" is a simplification of the process by which Ranch Hands accumulated dioxin during their tour of duty in Vietnam. This process, which undoubtedly varied from individual to individual, is unknown. However, the Ranch Hand tours generally were short (1 to 3 years) relative to the time elapsed since their tours. Hence, additional knowledge regarding the accumulation of dioxin during an individual Ranch Hand's tour, were it to become available, likely would not change conclusions drawn from any of the statistical analyses presented in this report.

Analyses based on model 1 are dependent directly on the first-order elimination assumption, while those based on model 2 are not. With model 1 one assumes that elimination is first-order and that the half-life is 7.1 years for all Ranch Hands. With model 2 one assumes nothing about the kinetics of dioxin elimination other than Ranch Hands received a dose in Vietnam and that their body burdens have decreased in an unspecified manner with time. Thus, with model 1 one assumes "everything" is known about dioxin elimination in Ranch Hands; with model 2 one assumes "nothing" about dioxin elimination in Ranch Hands. All health data were analyzed with both models to reduce the likelihood that an effect would be missed due to incorrect assumptions regarding dioxin elimination.

The introduction of the time-by-current dioxin interaction term ($b_3 T \log_2 [C]$) in model 2 allows investigation of the dioxin health relationship with respect to time. For example, such an effect would be detected by model 2 if there was no relationship between health and dioxin in the first few years after exposure and a strong positive relationship many years after exposure. In this case, if the effect were strong enough, it would be detected by the interaction coefficient (b_3) being significantly different from zero. Following that, analyses within time strata would find the coefficient (b_1) of $\log_2 (C)$ significantly different from zero and positive for large values of time (T); no significant difference between b_1 and 0 for small values of T would be found. It is important to note that a significant effect of this kind could be due to the passage of time or to a higher initial dioxin level received by Ranch Hands in the later time stratum or both of these.

Analyses based on models 1 and 2 were carried out both adjusted and unadjusted for covariates.

No additional data or other information exist to determine whether any of the Ranch Hands with background levels (≤ 10 ppt) of current dioxin ($n=345$) received a dose above background levels in Vietnam. To accommodate this lack of knowledge, all analyses based on models 1 and 2 were carried out with these Ranch Hands excluded. Additionally, since 10 ppt may be considered arbitrary or too conservative, all analyses based on models 1 and 2 were carried out with Ranch Hands having less than or equal to 5 ppt ($n=124$) excluded. With the second approach, it is assumed that Ranch Hands currently having more than 5 ppt (the approximate Comparison median) were exposed in Vietnam and those with less than 5 ppt were not. These two assumptions are termed "minimal" (Ranch Hands with more than 10 ppt were exposed in Vietnam) and "maximal" (Ranch Hands with more than 5 ppt were exposed in Vietnam).

TABLE 4-1.

**Models 1 and 2 for Assessing Health versus Dioxin in Ranch Hands Only:
Assumptions, Advantages, and Disadvantages**

Model 1: $y = \beta_0 + \beta_1 \log_2(I) + e$

where

y = health variable

I = extrapolated initial dose, assuming first-order elimination, $I = C \cdot \exp(\log_2 \cdot T/H)$

T = time between the end of the Vietnam Ranch Hand tour of duty and the 1987 physical examination

C = current dioxin body burden, determined in 1987

H = dioxin half-life in Ranch Hands assuming first-order elimination (7.1 years)

e = zero mean normal error

Assumptions: Ranch Hands received a single dioxin dose in Vietnam and background exposure thereafter.

Ranch Hands experienced first-order dioxin elimination with a constant known half-life of 7.1 years.

The error variance does not change with health status (y) or initial dioxin dose (I).

Advantages: Easily interpretable.

Most efficient if first-order elimination and constant half-life are valid assumptions and y is linearly related to $\log_2(I)$

Disadvantages: Will be biased if first-order elimination or constant half-life assumption is not valid.

Does not address time-related effects.

TABLE 4-1. (Continued)

**Models 1 and 2 for Assessing Health versus Dioxin in Ranch Hands Only:
Assumptions, Advantages, and Disadvantages**

Model 2: $y = \beta_0 + \beta_1 \log_2(C) + \beta_2 T + \beta_3 T \log_2(C) + e$

where

y = health variable

T = time between the end of the Vietnam Ranch Hand tour of duty and the 1987 physical examination

C = current dioxin body burden, determined in 1987

e = zero mean normal error

Assumptions: Ranch Hands received a single dioxin dose in Vietnam and background exposure thereafter.

Ranch Hand dioxin body burdens changed with time (T) in the same way for all individuals.

The dioxin versus health relationship may change with time (T).

The error variance does not change with values of the health variable (y), the current dioxin body burden (C), time (T), or the product of time and the logarithm of the current dioxin body burden ($T \log_2[C]$).

Advantages: Does not depend on any particular elimination law or half-life assumptions.

Assesses time-related effects.

Disadvantages: Less easily interpreted than model 1.

Less efficient than model 1 if first-order elimination and constant half-life are valid assumptions and y is linearly related to $\log_2(I)$.

Biased if any of the assumptions are violated.

In summary, to address the second fundamental limitation, two assumptions about Ranch Hands with current dioxin body burdens less than 10 ppt were made. These minimal and maximal assumptions are

- *Minimal assumption: Ranch Hands with less than or equal to 10 ppt were not exposed to dioxin in Vietnam*
- *Maximal assumption: Ranch Hands with less than or equal to 5 ppt were not exposed to dioxin in Vietnam.*

The terms minimal and maximal were given because fewer Ranch Hands were exposed under the minimal than under the maximal assumption. The numbers 5 and 10 correspond to the approximate median and 98th percentile of the Comparison current dioxin distribution. Based on this Comparison dioxin distribution, current dioxin levels less than 10 ppt are called background levels.

To assess the dioxin versus health relationship while addressing the second fundamental limitation, all analyses based on models 1 and 2 were carried out under the minimal and again under the maximal assumptions. Under the minimal assumption, Ranch Hands with less than or equal to 10 ppt were excluded from the analyses. Under the maximal assumption, Ranch Hands with less than or equal to 5 ppt were excluded from the analyses.

Table 4-2 shows counts of exposed Ranch Hands under the minimal and maximal assumptions with initial and current dioxin trichotomized for tabular presentation. Ranch Hands under the maximal assumption are termed the "maximal cohort"; those under the minimal assumption are termed the "minimal cohort." The time between the end of tour and the 1987 physical examination is dichotomized at 18.6 years (corresponding approximately to the year 1969), the approximate median of the maximal cohort. The cutpoints for stratifying dioxin levels (I and C) were the approximate 25th and 75th percentiles and were specific to a particular cohort.

Health versus Dioxin in Ranch Hands and Comparisons

Finally, an assessment of the health consequences of current dioxin body burdens above background was carried out with a third model (model 3) that required no assumptions about when or how increased dioxin body burdens were attained and was applied to both Ranch Hand and Comparison data. This model assessed health versus categorized current dioxin body burden (D) with four levels, found in Table 4-3.

The cutpoint between the low and high categories, 33.3 ppt, is the approximate median dioxin level of Ranch Hands having more than 15 ppt. Ranch Hands having between 10 ppt and 15 ppt were excluded from these categorized dioxin analyses in an attempt to avoid misclassification of Ranch Hands to the unknown and low categories due to various sources of variation in the dioxin measurement.

Table 4-4 shows counts of participants within each level of categorized current dioxin. The relationship between current health and categorized dioxin body burden was based on the model shown in Table 4-5.

TABLE 4-2.

Ranch Hand Sample Sizes Under the Minimal and Maximal Assumptions

| Assumption | Stratum Name | Initial Dioxin (I) | | Current Dioxin (C) | | |
|------------|--------------|--------------------|-------|--------------------|-----------------|-----------------|
| | | Stratum | Count | Stratum | T≤18.6 Count | T>18.6 Count |
| Minimal | Low | 52<I≤93 | 130 | 10<C≤14.65 | 72 | 58 |
| | Medium | 93<I≤292 | 260 | 14.65<C≤45.75 | 128 | 132 |
| | High | 292<I | 131 | 45.75<C | 54 | 77 |
| | | Total | 521 | | 254 | 267 |
| Maximal | Low | 25<I≤56.9 | 185 | 5<C≤9.01 | 106 | 79 |
| | Medium | 56.9<I≤218 | 371 | 9.01<C≤33.3 | 191 | 179 |
| | High | 218<I | 186 | 33.3<C | 83 | 104 |
| | | Total | 742 | | 380 | 362 |

TABLE 4-3.

Current Dioxin Body Burden (D) Categorized in Ranch Hands and Comparisons for Model 3

| Value | Definition |
|------------|--|
| Background | Comparisons with up to 10 ppt |
| Unknown | Ranch Hands with up to 10 ppt |
| Low | Ranch Hands with more than 15 and up to 33.3 ppt |
| High | Ranch Hands with more than 33.3 ppt |

TABLE 4-4.

Counts of Participants by Level of Categorized Current Dioxin (D)

| Level | Count |
|------------|-------|
| Background | 786 |
| Unknown | 345 |
| Low | 196 |
| High | 187 |
| Total | 1,514 |

TABLE 4-5.

**Model 3 for Assessing Health versus Categorized Current Dioxin
Body Burden in Ranch Hands and Comparisons**

Model 3: $y = \beta_0 + \beta_1 D + e$

where

- y = health variable
- D = categorized current dioxin
- e = zero mean normal error

Assumptions: Dioxin body burden has accumulated with time.
The error variance does not change with categorized current dioxin body burden (D).

Advantage: Requires no assumption regarding the time course of dioxin accumulation or elimination.

Disadvantages: Makes no use of prior belief that Ranch Hands received an unusually large dioxin dose in Vietnam.
Does not address time-related effects.

In addition to assessing the overall mean change in the health variable (y) with levels of categorized current dioxin (D), the mean values of y within the unknown, low, and high categories were contrasted with the mean values of y within the background category.

Figure 4-1 summarizes the current dioxin levels used in models 1, 2, and 3.

Data Error

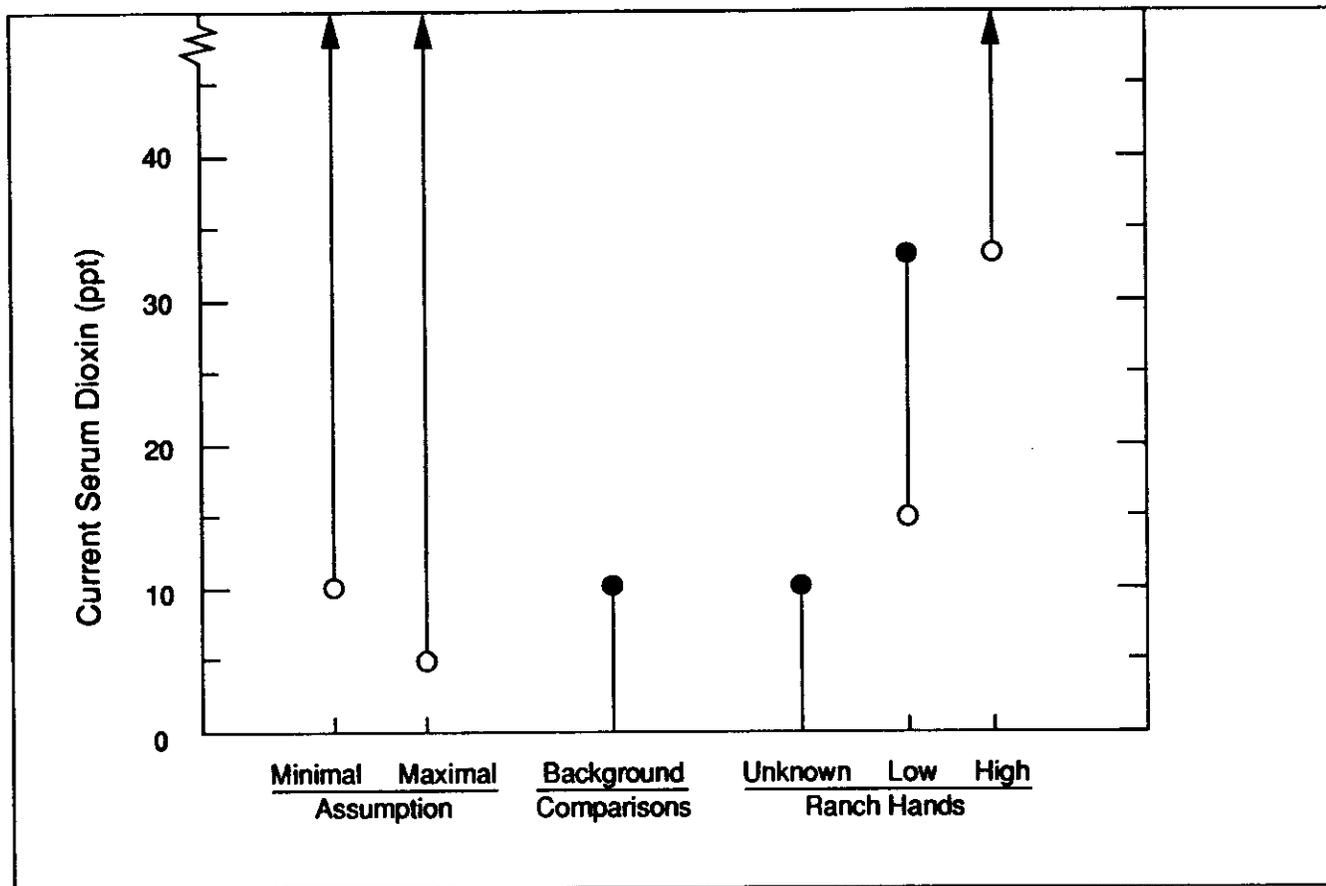
After the serum dioxin analyses were well underway, an error was discovered with respect to the race of one Comparison. The participant (subject 36410) was listed in the data base as a non-Black when in fact he was a Black. The Comparison was a 49-year-old at the Baseline examination and he was a member of the enlisted groundcrew cohort. His current serum dioxin value was 3.97 ppt as determined from the assay performed on the 1987 examination serum sample. The following abnormal medical conditions were noted for this individual: hepatomegaly, reported and verified hypertension, hyperpigmentation, and acne. The data error was corrected for the cardiovascular, malignancy, and dermatology assessments. Because the individual was a Comparison only the model 3 analyses of the other clinical area assessments were affected.

Bias Calculations

In any epidemiologic study, investigators must be concerned with avoiding spurious conclusions that are attributable to limitations in study design or analysis. The introduction of the dioxin assay as the measure of exposure in this study has provided the best available information regarding dioxin exposure in Ranch Hands and Comparisons. Uncertainties remain, however, regarding the choice of statistical models with which to assess the relationship between dioxin and health.

Biased results will be produced if the assumptions underlying any of the three statistical models are violated. Of the three models, model 1 is the most vulnerable to this kind of bias, since it depends directly on two unvalidated assumptions: (a) that dioxin elimination is first-order and (b) all Ranch Hands eliminate dioxin at the same rate (all Ranch Hands have the same dioxin half-life of 7.1 years). Air Force investigators currently are gathering additional data to evaluate both assumptions. The original half-life study on 36 Ranch Hands is being expanded to approximately 500 Ranch Hands. Assuming that dioxin elimination is first-order, this larger study will allow an assessment of half-life variability with weight changes, percent body fat changes, and disease since exposure. Additionally, the Air Force is collaborating with the CDC and Italian health authorities to assay serum collected periodically from people exposed in the Seveso accident. These data will consist of five dioxin measurements taken over a period of 10 years on 20 males who were adults at the time of the accident and will allow, for the first time, a direct assessment of the first-order elimination assumption in humans.

Until the Ranch Hand half-life study is expanded, the only available information regarding half-life variation in Ranch Hands is that derived from the smaller cohort of 36 subjects. Unpublished analyses of half-life heterogeneity among those 36 Ranch Hands suggest that half-life varies with relative weight changes between 1982 and 1987. With relative weight changes dichotomized at the median (2.7%), the 18 Ranch Hands below the median have an estimated half-life of 9.7 years (95% C.I.: [6.8,17.3]) and the 18 Ranch



764P

FIGURE 4-1. Ranges of Current Serum Dioxin Levels Used in Different Analysis Models

Hands above the median have an estimated half-life of 6.2 years (95% C.I.: [5.0,8.0]). The analysis showed a significant difference between these two half-lives ($p=0.02$). The two confidence intervals overlap because they are not derivable from the test for equality of half-lives. "Apparent" half-life decreases may be due to weight gain because of dilution of the body burden when it is redistributed to the new adipose tissue. Conversely, when there has been weight loss, the body burden may be redistributed in less adipose tissue and the serum concentration increases.

If these results are generalized to all Ranch Hands, statistical inference based on model 1 will be biased. For example, if the first-order elimination assumption is valid, but the constant half-life assumption is not, and there is no misclassification with regard to health status, odds ratios expressing the relationship between health and dioxin based on model 1 will be biased toward unity. That is, a misspecification of a constant half-life when, in fact, half-life changes with weight changes, will lead to misclassification with regard to dioxin level and therefore reduce our ability to detect an association between health and dioxin. To evaluate this possibility, the bias induced in the odds ratio under the maximal assumption and the computation of initial dioxin body burden assuming a constant half-life of 7.1 years (when in fact 50 percent of Ranch Hands have a dioxin half-life of 6 years and the other 50 percent have a dioxin half-life of 10 years) was calculated (7). In carrying out this calculation, it was assumed that initial dioxin had been dichotomized to high and low, with Ranch Hands assigned to the high category if their calculated initial dioxin level was greater than 218 ppt and assigned to the low category if their level was less than 218 ppt. The sample sizes of the real maximal cohort were used in the calculation; 186 Ranch Hands had a high initial dose and 556 had a low initial dose. With these assumptions, 76.3 percent of Ranch Hands assigned to the high category and 6.1 percent assigned to the low category truly had an initial dose above 218 ppt. The resultant bias in the odds ratio due to this misclassification depends on the true value of the odds ratio and the disease prevalence in the low category. For example, if the true odds ratio is 2.0 and the disease prevalence in the low initial dioxin category is 5 percent, this misclassification will produce an odds ratio of 1.7. Table 4-6 shows other values of the biased odds ratio produced by this misclassification for true odds ratios from 1 to 3 and the disease prevalence in the low initial dioxin category held fixed at 5 percent. There is no bias under assumptions if there is no association between initial dioxin and disease (true odds ratio equal to 1.0).

Model 2 also may be biased if, as suggested by the weight change analysis on the 36 Ranch Hands in the half-life study, 50 percent of Ranch Hands are fast dioxin eliminators (having a short half-life) and 50 percent of Ranch Hands are slow eliminators (with a longer half-life). If this attribute is not taken into account in the analysis (such as through adjustment for relative weight change), then the odds ratio relating disease to dioxin exposure will be biased toward unity. Again, disease status is assumed to be determined without error. For example, if slow eliminators experience an effect that does not become expressed until 20 years after exposure, if fast eliminators do not experience the effect, and if the analysis is not adjusted for relative weight change, then the ability of the model to detect the effect will be attenuated by the lack of adjustment. The extent of this bias toward the null depends on the nature of the four-factor interaction between health, current dioxin, time, and relative weight change, as well as upon the disease prevalence among Ranch Hands with low dioxin levels at each combination of categories of time and relative weight change. Bias calculations for this scenario, therefore, are more complicated and speculative than those presented for model 1 and were not pursued further.

TABLE 4-6.

Biased Odds Ratios Produced by a Misspecification of the Half-Life in the Calculation of the Initial Dioxin Body Burden in Model 1, Assuming a Disease Prevalence of 5 Percent in Ranch Hands Having a Low Calculated Initial Dose

| True Odds Ratio | Biased Odds Ratio |
|-----------------|-------------------|
| 1.0 | 1.0 |
| 1.5 | 1.3 |
| 2.0 | 1.7 |
| 2.5 | 2.0 |
| 3.0 | 2.2 |

Model 3 requires fewer assumptions than models 1 or 2, but is susceptible to bias due to misclassification or incorrect modeling. Biased results most likely are to occur with model 3 due to the failure to adjust for an important covariate. Every attempt, however, has been made in this report to adjust for all known important covariates.

The Correlation Between Initial Dioxin and Current Dioxin

The extrapolated initial dioxin dose is correlated highly with current dioxin level (correlation coefficient >0.98 for both the minimal and maximal cohorts). The same high correlation is, of course, seen between the logarithms of these quantities. The reason for the high correlation is that the initial dioxin dose is the current dioxin body burden multiplied by 2 raised to the power $T/7.1$. This high correlation is simply an expression of the fact that if the first-order model is valid and if dioxin half-life is constant, then models 1 and 2 nearly are redundant because the variation of time (T) is relatively small (see Figure 4-2).

FACTORS DETERMINING ANALYTICAL METHOD

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

- Dependent Variable Form — Continuous or discrete
- Serum Dioxin Estimate — Initial dioxin, current dioxin and time since tour, or categorized current dioxin incorporating group membership
- Analysis Type — Unadjusted, adjusted, or longitudinal

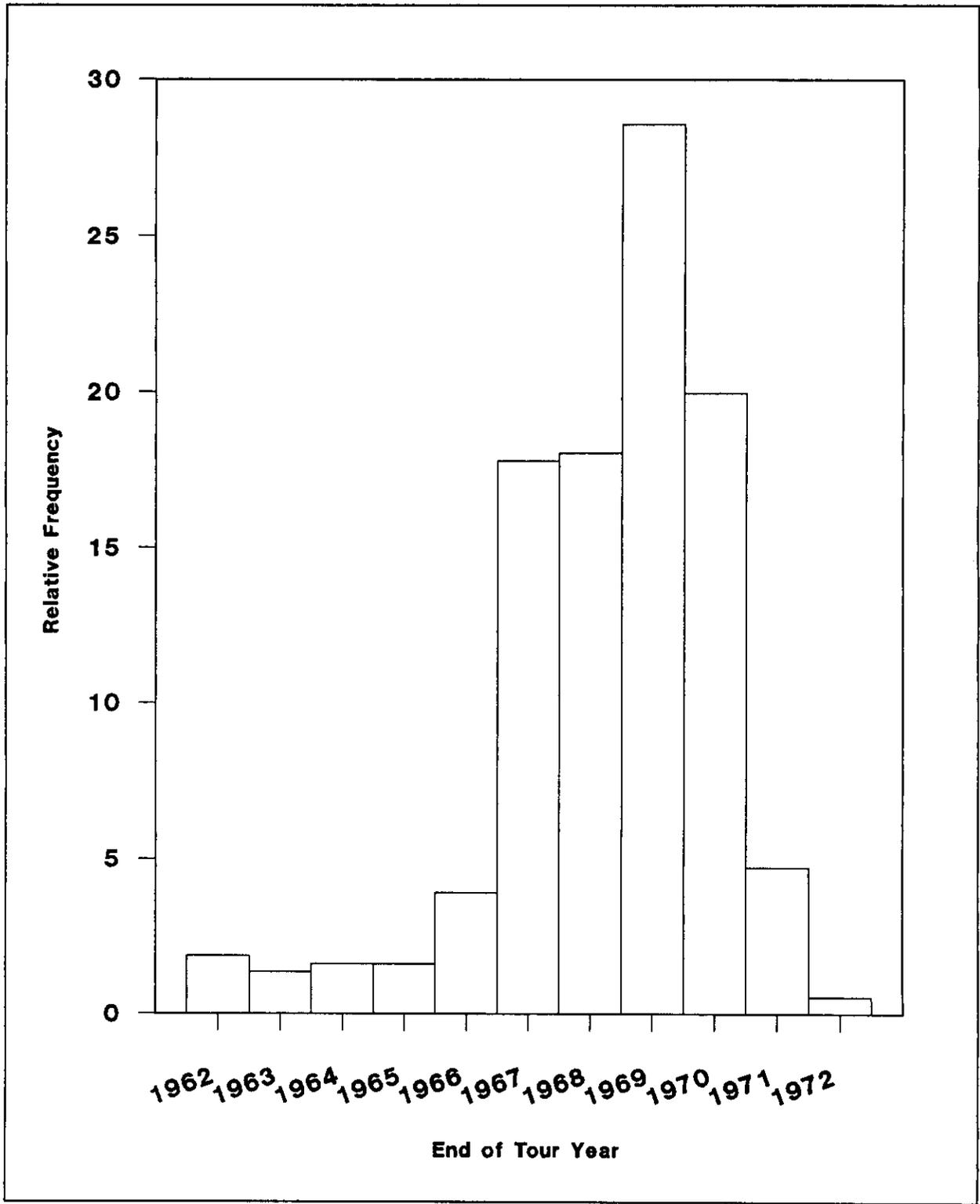


FIGURE 4-2. Relative Frequency Distribution of End of Tour Year in Ranch Hands Under the Maximal Assumption (N=742)

- Analysis Cohort(s) — Ranch Hands: minimal assumption, Ranch Hands: maximal assumption, and defined subsets of Ranch Hands and Comparisons for the categorized current dioxin variable.

Appendix Table C-1 specifies 30 separate analysis situations based on dependent variable form, serum dioxin estimate, analysis type, and analysis cohort. For each of the 30 situations, the statistical method is specified.

ANALYSIS METHODOLOGIES

As in previous Air Force Health Study reports, current health dependent variables can be either continuous or discrete. For the former case, the general linear model approach is the basis for applying such techniques as simple and multiple linear regression, analysis of variance, analysis of covariance, and repeated measures analysis. This approach permits model fitting of the dependent variable as a function of dioxin, relevant covariates, dioxin-by-covariate interactions, and interactions between covariates. As part of the previous analyses of 1987 data, the health variables were examined to ensure that assumptions underlying statistical methods were met. Transformations used to enhance normality for specific continuous health variables in the previous analyses of 1987 data also were used for the serum dioxin analysis. For these continuous analyses, SAS® GLM (8) was used. When a "best" model was fitted, tests of significance for a dioxin effect were made. Associations with a p-value less than or equal to 0.05 are described as significant, and associations with a p-value greater than 0.05 but less than or equal to 0.10 are termed marginally significant or borderline significant. If there was a significant interaction between the dioxin variable and any covariate, the dioxin effect was assessed using stratification by different levels of the covariate(s) involved in the interaction.

Discrete dependent variables were analyzed by methods parallel to those used for continuous variables. For dichotomous variables, logistic regression was performed using BMDP®-LR (9). For polychotomous dependent variables, log-linear modeling was performed using BMDP®-4F (9) by incorporating the full k-factor interaction term involving the k covariates used in the model. For the log-linear modeling approach, covariate information must be categorized. Because of this required categorization of the covariate(s), the marginals were fixed in the log-linear model (10), effectively converting the log-linear model into a logit model. For the log-linear model, the significance of the relative risk for a particular categorized dioxin variable (i.e., categorized initial dioxin, categorized current dioxin and categorized time, or categorized current dioxin for specified subsets of Ranch Hands and Comparisons) was determined by examination of the appropriate model, as determined by the model that includes all statistically significant effects and a dioxin measure, or by examination of the significant interactions. Adjusted relative risks were derived from the coefficients of the appropriate model.

Selected longitudinal analyses were performed investigating changes in health status between 1982 and 1987, for each of the three dioxin analysis models. The variables selected for longitudinal study were chosen prior to all 1987 examination data analyses. 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

1987 examination. Analysis was performed in this manner to investigate any temporal effects of dioxin in the subgroup at risk (i.e., those participants who could become abnormal over the time span). The rate of abnormalities under this restriction approximates an incidence rate between 1982 and 1987. The dependent variable in this type of analysis was the health of participants at the 1987 examination whose health was normal in 1982. The independent variable(s) were the appropriate dioxin measures.

For some variables, measurements in 1985 were substituted for 1982 measurements because the variable was not analyzed at the 1982 examination or inherently was different from the 1987 variable. For example, to enhance comparability, the longitudinal analyses for the neurological assessment were based on changes between 1985 and 1987 because SCRF conducted both of these examinations.

Both the general linear model and the logistic regression model approaches were applied using covariate information in either the discrete or the continuous form. Table 4-7 provides a summary of the basic statistical methods for the serum dioxin analyses.

MODELING STRATEGY

In each clinical category, many covariates were considered for inclusion in the statistical models relating specific health endpoints and dioxin. The large number of covariates, consequent interaction terms, and resulting difficulties of interpretation obligated the adoption of a strategy for identifying a moderately simple model using a stepwise strategy, as defined below. Interpretation of possible dioxin relationships was then made in the context of this simpler model.

In general, based on one of the adjusted analysis models described in Appendix Table C-1, an initial model was constructed containing any requisite two or three-factor interaction terms. As a first step, screening was performed at the 0.15 significance level to eliminate unnecessary two- and three-factor interactions. A hierarchical stepwise deletion strategy was applied at the 0.15 significance level on the set of main effect covariates (to address possible confounding effects between the covariates and dioxin) and at the 0.05 significance level for interactions. In general, the only effects not subject to the deletion strategy were the serum dioxin variables of interest (i.e., initial dioxin; current dioxin, time since tour, and current dioxin-by-time interaction; categorized current dioxin). With the objective of producing the simplest model, other lower-order effects were retained in the model only if involved in significant higher-order interactions. Significant interactions between covariates were retained as terms in the model.

The modeling strategy was refined slightly for adjusted statistical analyses of discrete dependent variables for particular clinical areas where a large number of covariates and/or sparse number of abnormalities were encountered. In these situations, the starting model included all main effects and excluded all interactions. Main effects were stepped out of the model if the associated p-value was greater than 0.15 and interactions were entered into the model if the associated p-value was less than or equal to 0.05. The alternative strategy was used to avoid overspecification of the model and minimize collinearity among terms that can lead to imprecise parameter and standard error estimates.

TABLE 4-7.

Summary of Statistical Procedures

Chi-square Contingency Table Test

The chi-square test of independence (11) is calculated for a contingency table by the following formula:

$$\chi^2 = \sum(f_o - f_e)^2 / f_e$$

where the sum is taken over all cells of the contingency table and

f_o = observed frequency in a cell

f_e = expected frequency under the hypothesis of independence.

Large values indicate deviations from the null hypothesis and are tested for significance by comparing the calculated χ^2 to the tables of the chi-square distribution.

Fisher's Exact Test

Fisher's exact test (11) is a randomization test of the hypothesis of independence for a 2 x 2 contingency table. This technique was used for small samples and sparse cells. This is a permutation test based on the exact probability of observing the particular set of frequencies, or of one more extreme.

Correlation Coefficient (Pearson's Product-Moment)

The population correlation coefficient (12), ρ , measures the strength of the linear relationship between two random variables X and Y. A commonly used sample-based estimate of this correlation coefficient is

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{[\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2]^{1/2}}$$

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 521 (the size of the Ranch Hand cohort under the minimal assumption), a sample correlation coefficient of ± 0.086 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 for the sample size of 742 under the maximal assumption, a sample coefficient of ± 0.072 is sufficient.

TABLE 4-7. (Continued)
Summary of Statistical Procedures

General Linear Models Analysis

The form of the general linear model (13) for two independent variables is

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \epsilon$$

where

- Y = dependent variable (continuous)
- α = 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
- β_1, β_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.
- β_{12} = coefficient reflecting the linear interaction of X_1 and X_2 , adjusted for linear main effects
- ϵ = 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 and interaction terms 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 (11, 14) enables a dichotomous dependent variable to be modeled in a regression framework with continuous and/or discrete independent variables. For two risk factors, such as dioxin and age, the logistic regression model would be

$$\text{logit } P = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \epsilon$$

TABLE 4-7. (Continued)
Summary of Statistical Procedures

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:

α = log odds for the disease when $X_1 = 0$ and $X_2 = 0$

β_1 = coefficient indicating the dioxin effect adjusted for age

β_2 = coefficient indicating the age effect adjusted for dioxin

β_{12} = coefficient indicating the interaction between dioxin and age, adjusted for linear main effects

ϵ = error term.

In the absence of an interaction ($\beta_{12} = 0$) for a dichotomous risk factor (e.g., Comparisons, Ranch Hands), $\exp(\beta_1)$ reflects the adjusted odds ratio for individuals in group 1 ($X_1 = 1$) relative to group 0 ($X_1 = 0$). If the probability of disease is small, the odds ratio will be approximately equal to the relative risk. In the absence of an interaction for a continuous risk factor (e.g., initial dioxin in its continuous form), $\exp(\beta_1)$ reflects the adjusted odds ratio for a unit increase in the risk factor. 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, the adjusted odds ratios will be referred to as adjusted relative risks. Correspondingly, in the absence of covariates (i.e., unadjusted analysis), the odds ratios will be referred to as estimated relative risks.

This technique will also be used for longitudinal analyses of dichotomous dependent variables to examine changes in health status between 1982 (or 1985) and 1987 in relation to the dioxin measures.

TABLE 4-7. (Continued)
Summary of Statistical Procedures

Log-linear Analysis

Log-linear analysis (11) is a statistical technique for analyzing cross-classified data or contingency tables. A saturated log-linear model for a three-way table is

$$\ln (Z_{ijk}) = U_0 + U_{1(i)} + U_{2(j)} + U_{3(k)} + U_{12(ij)} + U_{23(jk)} + U_{13(ik)} + U_{123(ijk)}$$

where

Z_{ijk} = expected cell count

$U_{1(i)}$ = specific one-factor effect

$U_{12(ij)}$ = specific two-factor effect or interaction

$U_{123(ijk)}$ = three-factor effect or interaction.

The simplest models are obtained by including only the significant U-terms. Adjusted relative risks are derived from the estimated U-terms from an adequately fitting model.

In the analysis of a particular health variable, when no dioxin-by-covariate interactions were significant at the 0.05 level, adjusted means (15) or relative risks were presented. If a dioxin-by-covariate interaction was significant at the 0.05 level, the behavior of the dioxin variable was explored for different levels (categories) of the covariate to identify subpopulations for which a dioxin relationship might exist. Further, for illustrative purposes, if any dioxin-by-covariate interaction was significant at a level between 0.01 and 0.05, the adjusted means or relative risks also were presented, after dropping the interaction terms from the model.

In some instances a followup model also was performed that excluded a highly significant interaction ($p < 0.01$). This optional model was run at the discretion of the analyst in an attempt to simplify the interpretation that may be complicated by an interaction difficult to explain from a clinical perspective.

For all models that included a dioxin-by-covariate interaction, the stratified results presented in the appendices display adjusted relative risks, confidence intervals, and associated p-values determined from a model that included the interaction term. However, in the model 2 analyses the p-values for the stratified current dioxin-by-time since tour interaction terms were determined from separate models for each covariate stratum; similarly in the model 3 analyses, the overall p-values were determined from separate models.

The adjusted models assessed the statistical significance of interactions between dioxin and the covariates to determine whether the relationship between dioxin and the dependent variable (health-related endpoint) differed across levels of the covariate. In many instances the clinical importance of a statistically significant dioxin-by-covariate interaction is unknown or uncertain. The clinical relevance of a statistically significant interaction would be strengthened if the same interaction persisted among related endpoints. It is recognized that due to the large number of dioxin-by-covariate interactions that were examined for approximately 300 variables, some of the dioxin-by-covariate interactions judged significant at the 0.05 level might be spurious; i.e., chance occurrences not of biological/clinical relevance. This should be considered when significant dioxin-by-covariate interactions are interpreted. It is important that the size of the p-value associated with each dioxin-by-covariate interaction be weighed carefully; for this reason, if the p-value for a dioxin-by-covariate interaction was between 0.01 and 0.05, the adjusted means or relative risks (omitting the interaction) were reported.

For the neurology, cardiovascular, renal, and endocrine clinical assessments, additional analyses were performed when certain covariates were retained in the final model. These covariates were variables that may have been affected by dioxin exposure and included diabetic class (neurology and renal), percent body fat (cardiovascular and endocrine), and cholesterol (cardiovascular). Due to the association between these covariates and dioxin, both the statistical and clinical interpretation of other health variables can be affected. Analyses were consequently performed with these covariates in the final model, and with the covariates removed from the model. Tabular results with these covariates in the model are given in the body of the clinical chapter; results with these covariates removed are given in the associated chapter appendix.

POWER

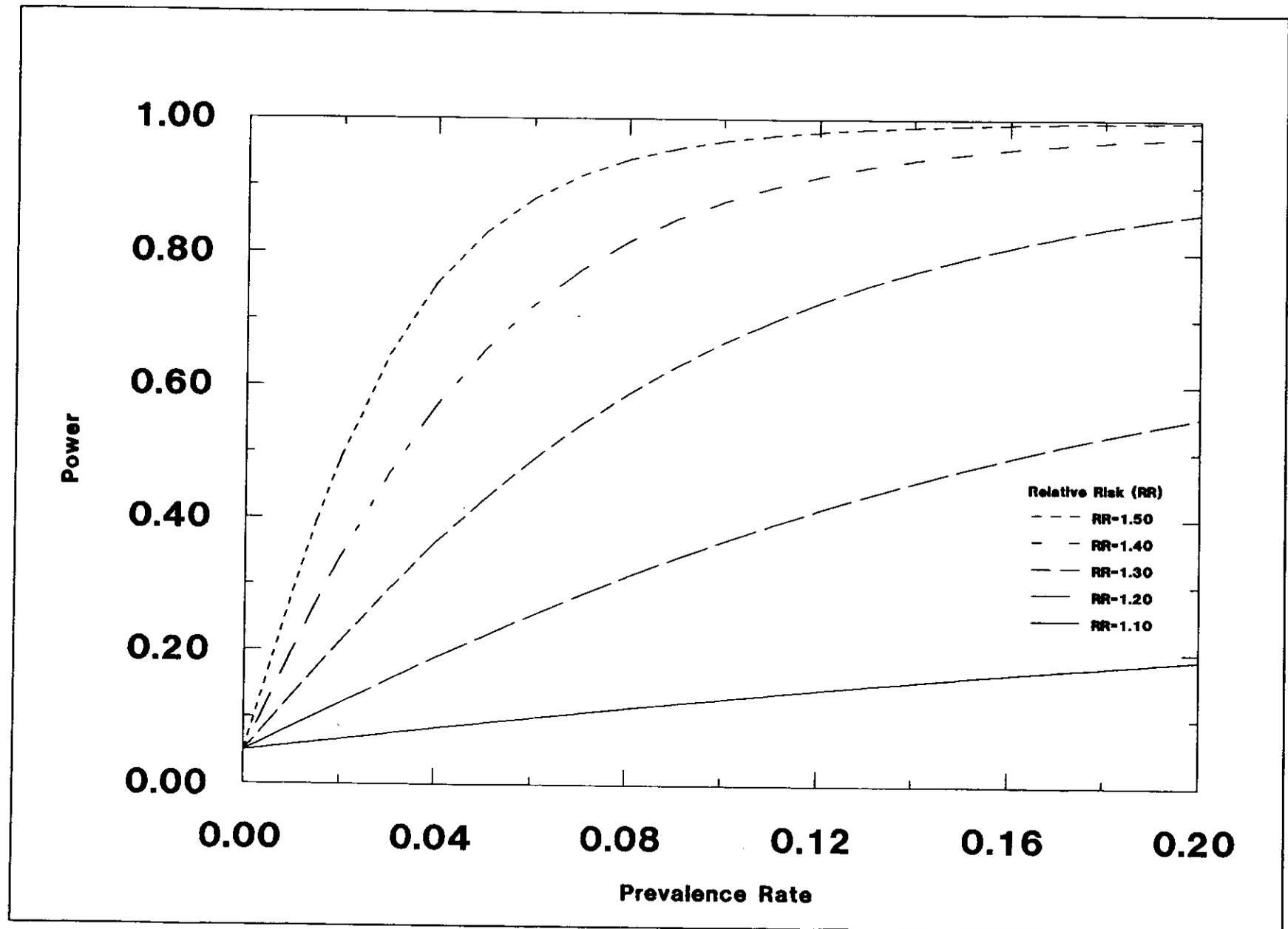
Conducting a statistical test using a type I error, also called alpha level, of 0.05 means that, on the average in 5 cases out of 100, a false conclusion would be made that an association (dioxin effect) exists when, in reality, there is no association. The other possible inference error (called a type II error) is the failure to detect an association when one actually exists. The probability of a type II error for a statistical test is 1 minus the power of the test. The power of the test is the probability that the test will reject the hypothesis of no dioxin effect when an effect does in fact exist. The power of a test depends on the distribution of the dioxin data, the sample size, the disease prevalence rate, and the true dioxin effect measured in terms of the relative risk.

Table 4-8 contains the approximate power for detecting specified relative risks for a given prevalence rate (discrete dependent variable), using initial dioxin in its continuous form and an alpha level of 0.05 for a two-sided test under the minimal assumption ($n=521$). The corresponding power under the maximal assumption is slightly higher. Figure 4-3 presents a graphical display of the power at different prevalence rates, where the different curves represent relative risks of 1.1, 1.2, 1.3, 1.4, and 1.5. Power calculations were performed using the logarithm (base 2) of initial dioxin, and consequently the relative risk is for a twofold increase in initial dioxin. These calculations also assume approximate prevalences at the mean \log_2 (initial dioxin) value of 7.49, corresponding to an initial dioxin level of 180 ppt.

TABLE 4-8.

Power to Detect an Initial Dioxin Effect Based on the Minimal Assumption at a 5 Percent Significance Level (Discrete Dependent Variable)

| Prevalence Rate of Disease | Relative Risk | | | | | | |
|----------------------------|---------------|------|------|------|------|------|------|
| | 1.10 | 1.20 | 1.30 | 1.40 | 1.50 | 1.75 | 2.00 |
| 0.005 | 0.05 | 0.07 | 0.09 | 0.12 | 0.17 | 0.33 | 0.54 |
| 0.01 | 0.06 | 0.09 | 0.13 | 0.20 | 0.29 | 0.56 | 0.80 |
| 0.02 | 0.07 | 0.12 | 0.21 | 0.34 | 0.49 | 0.82 | 0.96 |
| 0.03 | 0.08 | 0.16 | 0.29 | 0.46 | 0.64 | 0.93 | 0.99 |
| 0.04 | 0.08 | 0.19 | 0.36 | 0.57 | 0.75 | 0.97 | 1.00 |
| 0.05 | 0.09 | 0.22 | 0.43 | 0.65 | 0.83 | 0.99 | 1.00 |
| 0.10 | 0.13 | 0.36 | 0.66 | 0.88 | 0.97 | 1.00 | 1.00 |
| 0.15 | 0.16 | 0.47 | 0.79 | 0.95 | 0.99 | 1.00 | 1.00 |
| 0.20 | 0.18 | 0.55 | 0.86 | 0.97 | 1.00 | 1.00 | 1.00 |



**FIGURE 4-3. Power to Detect an Initial Dioxin Effect
(Discrete Dependent Variable)**

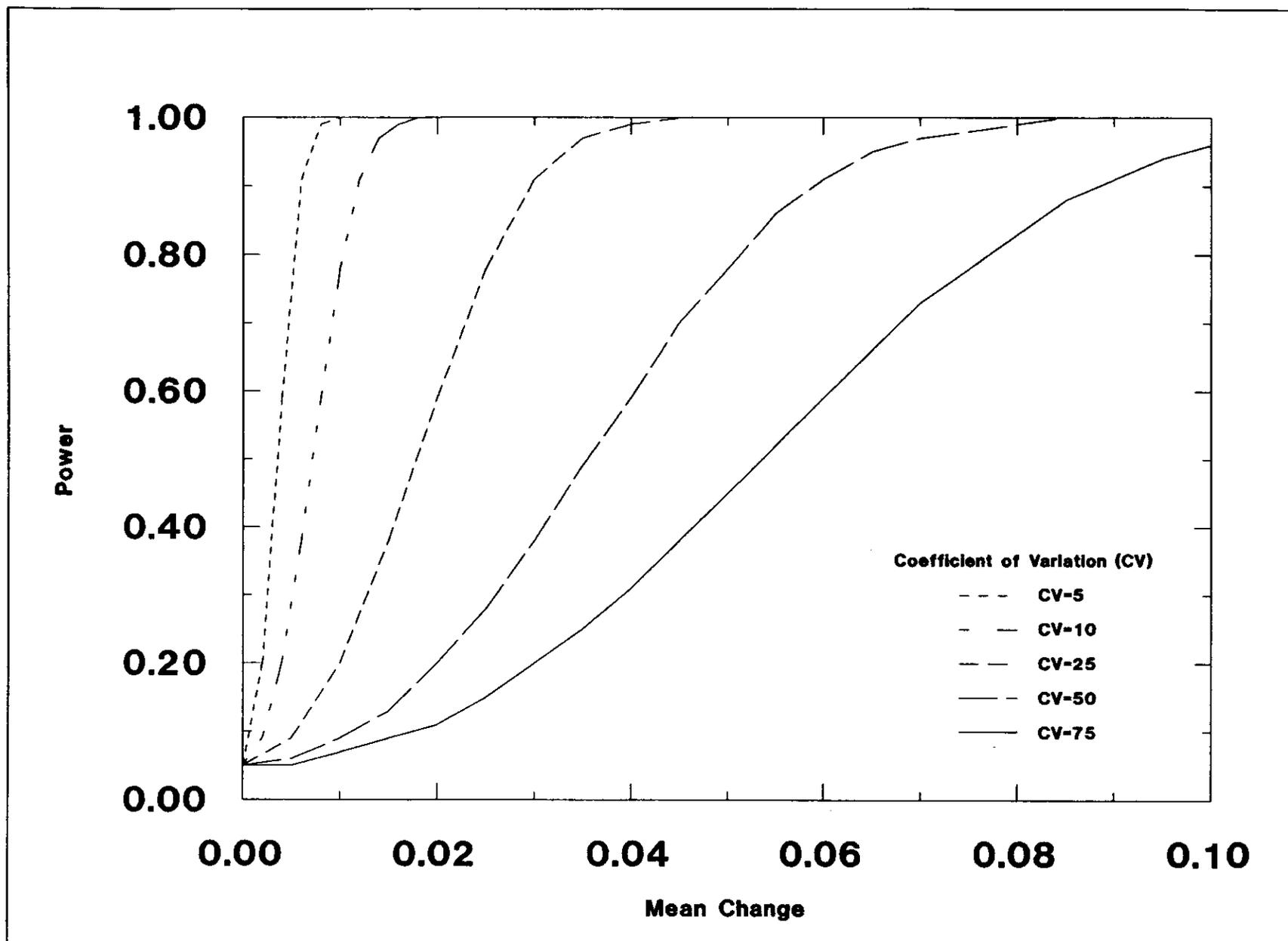
As an example, using age-adjusted incidence rates for all U.S. males (based on data from the Surveillance Epidemiology and End Results program of the National Cancer Institute), prevalence rates for all cancers, non-Hodgkin's lymphoma (NHL), and soft tissue sarcoma (STS) were estimated as 0.07, 0.002, and 0.001, respectively. Thus, Table 4-8 shows at least a power of 0.80 to detect a relative risk of 1.5 or greater given an estimated prevalence of 0.07 for all cancers. For the estimated prevalences of NHL and STS, the power to detect a relative risk of 2.0 would be less than 0.50.

Table 4-9 provides the same information for continuous variables in terms of coefficients 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 proportion mean change in this table 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. These mean changes are evaluated at the mean \log_2 (initial dioxin) value of 7.49, corresponding to an initial dioxin level of 180 ppt. The proportion mean change corresponds mathematically to the slope of the initial dioxin variable divided by the dependent variable mean, assuming no transformation of the dependent variable. An analogous quantity can be derived based on transformed statistics. Figure 4-4 shows a graphical display of the power at a given proportion mean change, where the different curves represent coefficients of variation of 5, 10, 25, 50, and 75. In this study, continuously distributed laboratory results were subject to a laboratory-error coefficient of variation of less than 3 percent.

TABLE 4-9.

Power to Detect an Initial Dioxin Effect Based on the Minimal Assumption at a 5 Percent Significance Level (Continuous Dependent Variable)

| Mean Change | Coefficient of Variation (σ/μ) | | | | |
|-------------|---|------|------|------|------|
| | 5 | 10 | 25 | 50 | 75 |
| 0.005 | 0.78 | 0.28 | 0.09 | 0.06 | 0.05 |
| 0.01 | 1.00 | 0.78 | 0.20 | 0.09 | 0.07 |
| 0.02 | 1.00 | 1.00 | 0.59 | 0.20 | 0.11 |
| 0.03 | 1.00 | 1.00 | 0.91 | 0.38 | 0.20 |
| 0.04 | 1.00 | 1.00 | 0.99 | 0.59 | 0.31 |
| 0.05 | 1.00 | 1.00 | 1.00 | 0.78 | 0.45 |
| 0.10 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 |



**FIGURE 4-4. Power to Detect an Initial Dioxin Effect
(Continuous Dependent Variable)**

TABLE 4-10.

Location of Table Results from Different Analysis Models

| Subpanel in Table | Dioxin Estimate | Type of Analysis | Assumption |
|-------------------|----------------------------|------------------|------------|
| a | initial ^a | unadjusted | minimal |
| b | initial ^a | unadjusted | maximal |
| c | initial ^a | adjusted | minimal |
| d | initial ^a | adjusted | maximal |
| e | current, time ^a | unadjusted | minimal |
| f | current, time ^a | unadjusted | maximal |
| g | current, time ^a | adjusted | minimal |
| h | current, time ^a | adjusted | maximal |
| i | current ^b | unadjusted | - - |
| j | current ^b | adjusted | - - |

^aRanch Hands only.

^bCategorized current dioxin, Ranch Hands and Comparisons.

EXPLANATION OF TABLES

This section introduces the reader to the contents of the tables that are used to report the results of the analyses for continuous and discrete dependent variables (two levels and more than two levels). Selected results from the statistical analysis methods applied in the hematology assessment (see Chapter 13, Hematologic Assessment) will be referenced throughout this discussion. The contents of each summary table depend on the form of the health status endpoint (i.e., whether the dependent variable under analysis is a continuous or discrete variable). Generally, the results of the various analyses will be summarized in subpanels within each table as specified in Table 4-10. The subpanel specifications may be slightly different when adjusted analyses are not performed. This section also provides an explanation of the information contained in these tables.

Continuous Variables

Table 13-3 presents an example of the results of analysis when the dependent variable is continuous. Subpanels (a) and (b) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association between the dependent variable and initial dioxin without adjusting for covariate information. Immediately below the specified assumption, the aggregate sample size (n) and the coefficient of determination (R²) associated with the simple linear regression of the continuous dependent variable on log₂ (initial dioxin) are presented. Sample sizes also are presented for low, medium, and high categories of initial dioxin. The numerical values defining these categories are specified in a table footnote. The low, medium, and high categories are based on the lower 25th percent, the 25th to 75th percent, and the upper 25th percent of the initial dioxin estimates for the cohort corresponding to the specified assumption. Means of the dependent variable (transformed to the original units, if necessary) are calculated from the data and are

presented for the low, medium, and high initial dioxin categories. Based on the simple linear regression analysis, the estimated slope and its associated standard error are reported for each assumption. 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 estimated slope is equal to zero also is presented under both assumptions.

Based on analyses that incorporate covariate and interaction information, subpanels (c) and (d) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association between the dependent variable and initial dioxin. Immediately below the specified assumption, the aggregate sample size (n) and the multiple coefficient of determination (R^2) are presented for a multiple linear regression of the continuous dependent variable on \log_2 (initial dioxin) including covariate and interaction effect terms in the adjusted model. Similar to the unadjusted analyses, sample sizes are also presented for low, medium, and high categories of initial dioxin. The numerical values defining these categories are specified in a table footnote. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information. Adjusted means of the dependent variable (transformed to the original units, if necessary) also are presented. The adjusted means are presented for the low, medium, and high initial dioxin categories. Based on the multiple linear regression analysis, the adjusted slope for the \log_2 (initial dioxin) term and its associated standard error are reported for each assumption. 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 adjusted slope is equal to zero also is presented under both assumptions.

Covariates with p-values less than or equal to 0.15 and interactions with p-values or equal to 0.05 retained in the multiple regression model after implementing the modeling strategy are presented under covariate remarks, along with the associated p-values. If the multiple regression model contains a significant initial dioxin-by-covariate interaction with an associated p-value less than or equal to 0.01, then the adjusted means, adjusted slope, standard error, and p-value generally are not reported. The entries for these statistics are reported as four asterisks (****) and are identified by a table footnote. Covariates and interactions retained in the model are, however, reported under covariate remarks. For some clinical assessments, an analyst may exercise discretion and report the adjusted means, adjusted slope, standard error, and a p-value from a model that excludes the interaction having a p-value less than 0.01. When these discretionary followup analyses are performed, the results are reported along with three asterisks (***) and are explained by a table footnote. If the multiple regression model contains a significant initial dioxin-by-covariate interaction with an associated p-value between 0.01 and 0.05, then the adjusted means, adjusted slope, standard error, and p-value are reported from a model that excludes that interaction. The entries for these statistics are reported along with two asterisks (**) accompanied by a table footnote. In either case (i.e., $p \leq 0.01$ or $0.01 < p \leq 0.05$), stratified analyses are undertaken and the results are reported in an associated appendix for each individual clinical area.

Subpanels (e) and (f) of Table 13-3, for example, report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association of the dependent variable with current dioxin and time since tour without adjusting for covariate information.

Multiple regression techniques are used to generate the statistics provided in both panels. In the multiple regression model, current dioxin is included as a continuous variable and time since tour as a discrete variable. The interaction of current dioxin and time since tour also is included. For these models, time since tour is dichotomized and separate statistics are presented on the association between the dependent variable and current dioxin within each time stratum. For each subpanel, the aggregate sample size (n) and the coefficient of determination (R^2) are presented, under each specified assumption, for the multiple linear regression model. For presentation purposes, current dioxin and time since tour both are categorized. The numerical values defining the current dioxin categories are specified in a table footnote. The low, medium, and high categories are based on the lower 25th percent, the 25th to 75th percent, and the upper 25th percent of the current dioxin estimates for the cohort corresponding to the specified assumption. The value of 18.6 years for time since tour corresponds to approximately the median value of time since tour in the Ranch Hand cohort. The means of the dependent variable (transformed to the original units, if necessary) are calculated from the data and are presented, along with sample size, for the combinations of trichotomized current dioxin and dichotomized time since tour. The first p-value within each subpanel evaluates the interaction term of the multiple regression using current dioxin in continuous form and time since tour in discrete form. The p-value for the interaction term provides a test of the equality of the slopes for the two time strata. For each time stratum, a simple linear regression model of the dependent variable on current dioxin (\log_2 scale) provides an estimated slope, associated standard error, and p-value for testing the significance of the slope. If the dependent variable was transformed for regression analysis, the means, slope, and standard error are footnoted and the transformation identified in the footnote.

Incorporating covariate and current dioxin-by-time-by-covariate interaction information into the analysis, subpanels (g) and (h) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association of the dependent variable with current dioxin, time since tour, and the current dioxin-by-time interaction. Multiple linear regression techniques are used to generate the statistics provided. In the overall multiple regression model, current dioxin is included as a continuous variable and time since tour as a discrete variable. The interaction of current dioxin and time since tour also is included. The test of the interaction of current dioxin and time since tour (i.e., the first p-value in each subpanel) determines whether the adjusted slopes of the two time strata differ significantly.

Immediately below the specified assumption, the aggregate sample size (n) and the multiple coefficient of determination (R^2) are presented for the multiple linear regression of the continuous dependent variable on current dioxin (\log_2 scale), time since tour, the current dioxin-by-time interaction, covariates, and other interactions retained in the model. For each time stratum (≤ 18.6 years or > 18.6 years), separate statistics relating the dependent variable to current dioxin (\log_2 scale) are presented. In particular, based on the multiple linear regression analysis, the adjusted slope for the current dioxin term (\log_2 scale), its associated standard error, and a p-value for testing the significance of the slope are reported.

Sample sizes also are presented for combinations of low, medium, and high categories of current dioxin and dichotomized time since tour. The numerical values defining these categories are specified in a table footnote. Sample sizes for corresponding panels of unadjusted and adjusted analyses may differ because of missing covariate information.

Adjusted means of the dependent variable (transformed to the original units, if necessary) are presented. The adjusted means are presented for the combinations of trichotomized current dioxin and dichotomized time since tour. 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.

Covariates (p-values less than or equal to 0.15) and interactions (p-values less than or equal to 0.05) retained in the multiple regression model after implementing the modeling strategy are presented under covariate remarks, along with the associated p-values. If the multiple regression model contains a significant current dioxin-by-time-by-covariate interaction term with an associated p-value less than or equal to 0.01, then the adjusted means, adjusted slope, standard error, and p-value generally are not reported. The entries for these statistics are reported as four asterisks (****) and are identified by a table footnote. Covariates and interactions retained in the model are, however, reported under covariate remarks. For some clinical assessments, an analyst may exercise discretion and report adjusted means, adjusted slope, standard error, and a p-value from a model that excludes the interaction having a p-value less than 0.01. When these discretionary followup analyses are performed, the results are reported along with three asterisks (***) and are explained by a table footnote. If the multiple regression model contains a significant current dioxin-by-time-by-covariate interaction with an associated p-value between 0.01 and 0.05, then the adjusted means, adjusted slope, standard error, and p-value are reported from a model that excludes that interaction. The entries for these statistics are reported along with two asterisks (**) accompanied by a table footnote. In either case, interactions are investigated within strata of the covariate and reported in an associated appendix for each clinical area.

Subpanels (i) and (j) of Table 13-3, for example, show the results of unadjusted and adjusted analyses that compare the means of a continuous dependent variable for Ranch Hands with high, low, and unknown current dioxin levels and for Comparisons having background current dioxin levels. The note at the bottom of the table defines the four current dioxin categories. Sample sizes for each category and across the four categories are reported. The coefficient of determination (R^2) also is presented.

For the unadjusted analysis, dependent variable means are presented for each category. If the dependent variable was transformed for the analysis, the means of the transformed values are converted to the original scale and the column heading is footnoted. A test of the simultaneous equality of the four category means is evaluated by the first p-value cited. If the analysis was performed on a transformed scale, the p-value column is footnoted to indicate that the p-value is based on the difference of means on a transformed scale. For the individual contrasts of the three Ranch Hand categories versus Comparison background category, differences in means are reported on the original scale. If the analyses were performed on a transformed scale, 95 percent confidence intervals on the differences of means are not presented and the column is footnoted. A p-value also is reported to determine whether a difference in means for a specified contrast is significantly different from zero.

For an adjusted analysis, the table is modified to include adjusted means, differences in adjusted means (reported on the original scale), 95 percent confidence intervals on the differences in adjusted means (if the analysis was performed on the original scale), and any

covariates and interactions retained in the adjusted model along with their associated p-values.

Discrete Variables

Discrete Variable With Two Categories

Table 13-4 presents an example of the results of analysis when the dependent variable is discrete and dichotomous in form. Subpanels (a) and (b) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association between the dependent variable and initial dioxin without adjusting for covariate information. Immediately below the specified assumption, the aggregate sample size (n) associated with the simple logistic regression of the continuous dependent variable on \log_2 (initial dioxin) is presented. Sample sizes also are presented for low, medium, and high categories of initial dioxin. The numerical values defining these categories are specified in a table footnote. The low, medium, and high categories are based on the lower 25th percent, the 25th to 75th percent, and the upper 25th percent of the initial dioxin estimates for the cohort corresponding to the specified assumption. The percentage of Ranch Hands with the specified dichotomous characteristic (as cited in the column heading) is calculated from the data and presented for the low, medium, and high initial dioxin categories. Based on the simple logistic regression model, an estimated relative risk and its associated 95 percent confidence interval are reported for each assumption. The p-value associated with testing whether the relative risk is equal to one also is presented for both assumptions. The relative risk, confidence interval, and p-value are based on \log_2 (initial dioxin) in its continuous form.

Results may exhibit a significant ($p \leq 0.05$) p-value associated with testing whether the relative risk is equal to 1.00, while the corresponding 95 percent confidence interval on the relative risk contains the number 1.00. These results occur because the BMDP®-LR procedure uses a normal distribution in calculating an approximate 95 percent confidence interval and a chi-square distribution based on a likelihood ratio statistic (9) in the determination of a p-value. Similarly, the results may exhibit a 95 percent confidence interval of a relative risk that does not contain the number 1.00, while the corresponding p-value is not significant ($p > 0.05$) for the reasons stated above.

Incorporating covariate and interaction information, subpanels (c) and (d) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association between the discrete dependent variable and initial dioxin. Immediately below the specified assumption, the aggregate sample size (n) is presented for a multiple logistic regression of the discrete dependent variable on \log_2 (initial dioxin) including covariate and interactions in the adjusted model. Based on the multiple logistic regression model, the adjusted relative risk for the \log_2 (initial dioxin) term and its associated 95 percent confidence interval are reported for each assumption. The p-value for testing whether the adjusted relative risk is equal to 1 also is presented under both assumptions. Covariates (p-values less than or equal to 0.15) and interactions (p-values less than or equal to 0.05) retained in the multiple regression model after implementing the modeling strategy are presented under covariate remarks, along with the associated p-values. If the multiple logistic regression model contains a significant initial dioxin-by-covariate interaction with an associated p-value less than or equal to 0.01, then the adjusted relative risk, 95 percent confidence interval, and associated p-value generally are not reported. The entries for these statistics are reported

as four asterisks (****) and are identified by a table footnote. Covariates and interactions retained in the model are, however, reported under covariate remarks. For some clinical assessments, an analyst may exercise discretion and report an adjusted relative risk, 95 percent confidence interval, and an associated p-value from a model that excludes the interaction having a p-value less than 0.01. When these discretionary followup analyses are performed, the results are reported along with three asterisks (***) and are explained by a table footnote. If the multiple logistic regression model contains a significant initial dioxin-by-covariate interaction with a p-value between 0.01 and 0.05, then the adjusted relative risk, 95 percent confidence interval, and associated p-value are reported from a model that excludes that interaction. The entries for these statistics are reported along with two asterisks (**) accompanied by a table footnote. In either case (i.e., $p \leq 0.01$ or $0.01 < p \leq 0.05$), stratified analyses are undertaken and the results are reported in an appropriate appendix.

Subpanels (e) and (f) of Table 13-4, for example, report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association of the discrete dependent variable with current dioxin and time since tour without adjusting for covariate information. Multiple logistic regression techniques are used to generate the statistics provided in both panels. In the multiple logistic regression model, current dioxin is treated as a continuous variable and time since tour as a discrete variable. The interaction of current dioxin and time since tour also is included in the model. For the logistic regression model, time since tour is dichotomized and separate statistics are presented for the association between the dependent variable and current dioxin within each time stratum. For each subpanel, the aggregate sample size (n) is presented under each specified assumption for the multiple logistic regression model. For presentation purposes, current dioxin and time since tour both are categorized. The numerical values defining the current dioxin categories are specified in a table footnote. The low, medium, and high categories are based on the lower 25th percent, the 25th to 75th percent, and the upper 25th percent of the measured current dioxin for the cohort corresponding to the specified assumption. The value of 18.6 years for time since tour corresponds to approximately the median value in the Ranch Hand cohort. The percentage of Ranch Hands with the specified dichotomous characteristic (as cited in the column heading) is calculated from the data and presented, along with sample size, for the combinations of trichotomized current dioxin and dichotomized time since tour. Each panel also contains a p-value (i.e., the first p-value in each subpanel) for the interaction of the multiple logistic regression using current dioxin in continuous form and time since tour in discrete form. The p-value for the interaction term provides a test of the equality of the relative risks for the two time strata. For each time stratum, the logistic regression on current dioxin (\log_2 scale) provides an estimated relative risk, associated 95 percent confidence interval, and p-value for testing the significance of the relative risk.

Incorporating covariate and interaction information into the analysis, subpanels (g) and (h) report summary statistics (for the minimal and maximal assumptions, respectively) assessing the association of the discrete dependent variable with current dioxin, time since tour, and the current dioxin-by-time interaction. Multiple logistic regression techniques are used to generate the statistics provided. In the multiple logistic regression model, current dioxin is included as a continuous variable and time since tour as a discrete variable. The interaction of current dioxin and time since tour also is included. The test of the interaction of current dioxin and time since tour (i.e., the first p-value in each subpanel) determines whether the adjusted relative risks of the two time strata differ significantly.

Immediately below the specified assumption, the aggregate sample size (n) is presented for the multiple logistic regression of the continuous dependent variable on \log_2 (current dioxin), time since tour, the current dioxin-by-time interaction, covariates, and other interactions retained in the model. For each time stratum (≤ 18.6 years or > 18.6 years), separate statistics relating the dependent variable to current dioxin (\log_2 scale) are presented. Based on the multiple logistic regression analysis, the adjusted relative risk for the \log_2 (current dioxin) term, its associated 95 percent confidence interval, and a p-value for testing the significance of the adjusted relative risk are reported.

Covariates (p-values less than or equal to 0.15) and interactions (p-values less than or equal to 0.05) retained in the multiple logistic regression model after implementing the modeling strategy are presented under covariate remarks, along with the associated p-values. If the multiple logistic regression model contains a significant current dioxin-by-time-by-covariate interaction term such that the associated p-value is less than or equal to 0.01, then the adjusted relative risk, associated 95 percent confidence interval, and p-value generally are not reported. The entries for these statistics are reported as four asterisks (****) and are identified by a table footnote. Covariates and interactions retained in the model, however, are reported under covariate remarks. For some clinical assessments, an analyst may exercise discretion and report an adjusted relative risk, 95 percent confidence interval, and an associated p-value from a model that excludes the interaction having a p-value less than 0.01. When these discretionary followup analyses are performed, the results will be reported along with three asterisks (***) and are explained by a table footnote. If the multiple logistic regression model contains a significant current dioxin-by-time-by-covariate interaction such that the interaction lies between 0.01 and 0.05, then the adjusted relative risk, 95 percent confidence interval, and p-value are reported from a model that excludes that interaction. The entries for these statistics are reported along with two asterisks (**) accompanied by a table footnote. In either case ($p \leq 0.01$ or $0.01 < p \leq 0.05$), stratified analyses are undertaken and reported in the appropriate appendix.

Subpanels (i) and (j) of Table 13-4, for example, show the results of unadjusted and adjusted analyses that compare Ranch Hands with high, low, and unknown current dioxin levels and Comparisons having background current dioxin levels on the relative frequency for a specified discrete dependent variable (e.g., percent of participants in a current dioxin category with an abnormal condition). The note at the bottom of the table defines the four categories. Sample sizes for each category and across the four categories are reported.

For the unadjusted analysis, a relative frequency is presented for each current dioxin category. The simultaneous equality of the four category relative frequencies is evaluated by the first p-value cited. For the individual contrasts of the three Ranch Hand categories versus Comparison background category, relative risks, associated 95 percent confidence intervals for the relative risks, and p-values to evaluate if the risks differ significantly from 1 are presented.

Results may exhibit a significant ($p \leq 0.05$) p-value associated with testing whether the relative risk is equal to 1.00, while the corresponding 95 percent confidence interval on the relative risk contains the number 1.00. Similarly, the results may exhibit a 95 percent confidence interval of a relative risk that does not contain the number 1.00, while the

corresponding p-value is not significant ($p > 0.05$). These patterns are due to the use of the normal distribution in calculating an approximate 95 percent confidence interval and the use of Fisher's exact test for unadjusted analyses in the determination of the corresponding p-values in the event of sparse data.

For an adjusted analysis, the table presents adjusted relative risks, 95 percent confidence intervals on the adjusted relative risks, and covariates and interactions retained in the adjusted model along with their associated p-values.

Discrete Variable With More Than Two Categories

Log-linear analysis techniques were used to analyze discrete dependent variables having more than two levels (e.g., low, normal, high—see Table 13-6). For the unadjusted and adjusted analyses relating such discrete dependent variables to initial dioxin, summary tables present sample sizes, relative frequencies, relative risks, 95 percent confidence intervals for the relative risks, and associated p-values. For the adjusted analyses, any covariates and interactions retained in the model along with their associated p-values also are presented. One difference between the table presentations for dichotomous dependent variables and discrete dependent variables with more than two levels is that relative frequencies of Ranch Hands belonging to each of the dependent variable categories are summarized with respect to each initial dioxin category (i.e., low, medium, and high initial dioxin). Therefore, for each initial dioxin level, the relative frequencies sum to 100 percent across the dependent variable categories. Also, for specified pairs of dependent variable levels (e.g., low and normal or high and normal for the discrete dependent variable), contrasts for high initial dioxin versus low initial dioxin, and medium initial dioxin versus low initial dioxin, are constructed with relative risks, 95 percent confidence intervals, and associated contrast p-values. Contrasts are based on a categorized form (i.e., low, medium, and high) of initial dioxin rather than \log_2 (initial dioxin). A p-value for an overall test of independence between the dependent variable and initial dioxin also is reported.

Similar to the log-linear analysis using initial dioxin, unadjusted and adjusted analyses of discrete dependent variables with more than two categories were performed using current dioxin and time since tour. For the unadjusted analysis, sample sizes, relative frequencies (within each current dioxin level), current dioxin contrasts for specified pairs of dependent variable levels with relative risks, 95 percent confidence intervals on the relative risks, and associated contrast p-values were reported for each time since tour stratum. For these analyses a categorized form of current dioxin (i.e., low, medium, and high), rather than the continuous form of \log_2 (current dioxin), is used. For the adjusted analysis, contrast-specific adjusted relative risks with 95 percent confidence intervals, associated contrast p-values, and covariates and interactions retained in the model along with associated p-values are presented. For both the unadjusted and the adjusted analyses, a p-value is provided that tests the significance of the interaction between current dioxin and time since tour and, for each time stratum, another p-value is reported as an overall test of independence between the discrete dependent variable and current dioxin.

For log-linear analyses of initial dioxin, and those concerning current dioxin and time since tour, the cutpoints between the three dioxin categories (i.e., between low and medium dioxin, and between medium and high dioxin) are the same under both the minimal and

maximal assumptions. The actual cutpoints are relevant for log-linear analyses, and this standardization was done to permit a more valid comparison of category contrasts between the minimal and maximal assumptions.

Unadjusted and adjusted analyses comparing relative frequencies for discrete dependent variables of more than two categories also were performed to compare the four current dioxin categories. For the unadjusted analysis, sample sizes, relative frequencies (within each of the four categories), Ranch Hand versus Comparison contrasts for specified pairs of dependent variable levels with relative risks, 95 percent confidence intervals on the relative risks, and associated contrast p-values were reported. For the adjusted analysis, sample sizes, contrast-specific adjusted relative risks with 95 confidence intervals, associated contrast p-values, and covariates and interactions retained in the model along with associated p-values are presented. For both the unadjusted and the adjusted analyses, an all categories p-value is provided that tests the independence of the categories and the discrete dependent variable.

GRAPHICS

The analytic activities for the serum dioxin analyses were supplemented by two sets of graphic displays: data plots/histograms and interaction plots/histograms. These graphics were produced using the SYSTAT® graphics procedure (16).

Data Plots/Histograms

As part of the serum dioxin analyses, graphic displays were produced describing the relationship between each dependent variable and serum dioxin level, as well as relevant covariates and serum dioxin level. Evaluations of the relationships between dioxin and the covariates were carefully made because such relationships particularly are important in the interpretation of dioxin effects for this study (see Chapter 5, Covariate Associations). Initial and current dioxin levels were used in continuous form. Transformations used in statistical analyses also were incorporated into the graphic presentations.

For initial dioxin, dependent variable and covariate relationships were displayed separately for Ranch Hands under the minimal and maximal assumptions. In addition, graphic relationships between dependent health variables and current dioxin level, as well as relevant covariates and current dioxin level, were presented separately for all Comparisons and Ranch Hands.

For continuous dependent variables, bivariate scatterplots were produced. For binary or categorical dependent variables, bar charts with percentages of participants classified as abnormal for common interval groupings of dioxin were generated for each of the clinical areas. For the covariate associations section, relative frequency histograms were produced for each level of the covariate.

Figure 4-5 presents an illustration of the bar charts seen in the appendix for each clinical area. Figures 4-5(a), (b), and (c) display a positive relationship, no relationship, and a negative relationship between the percentage of participants classified as abnormal and dioxin. These displays were generated assuming equal sample sizes for each bar; inference based on unequal sample sizes is not straightforward. Figures 4-6(a), (b), and (c) illustrate

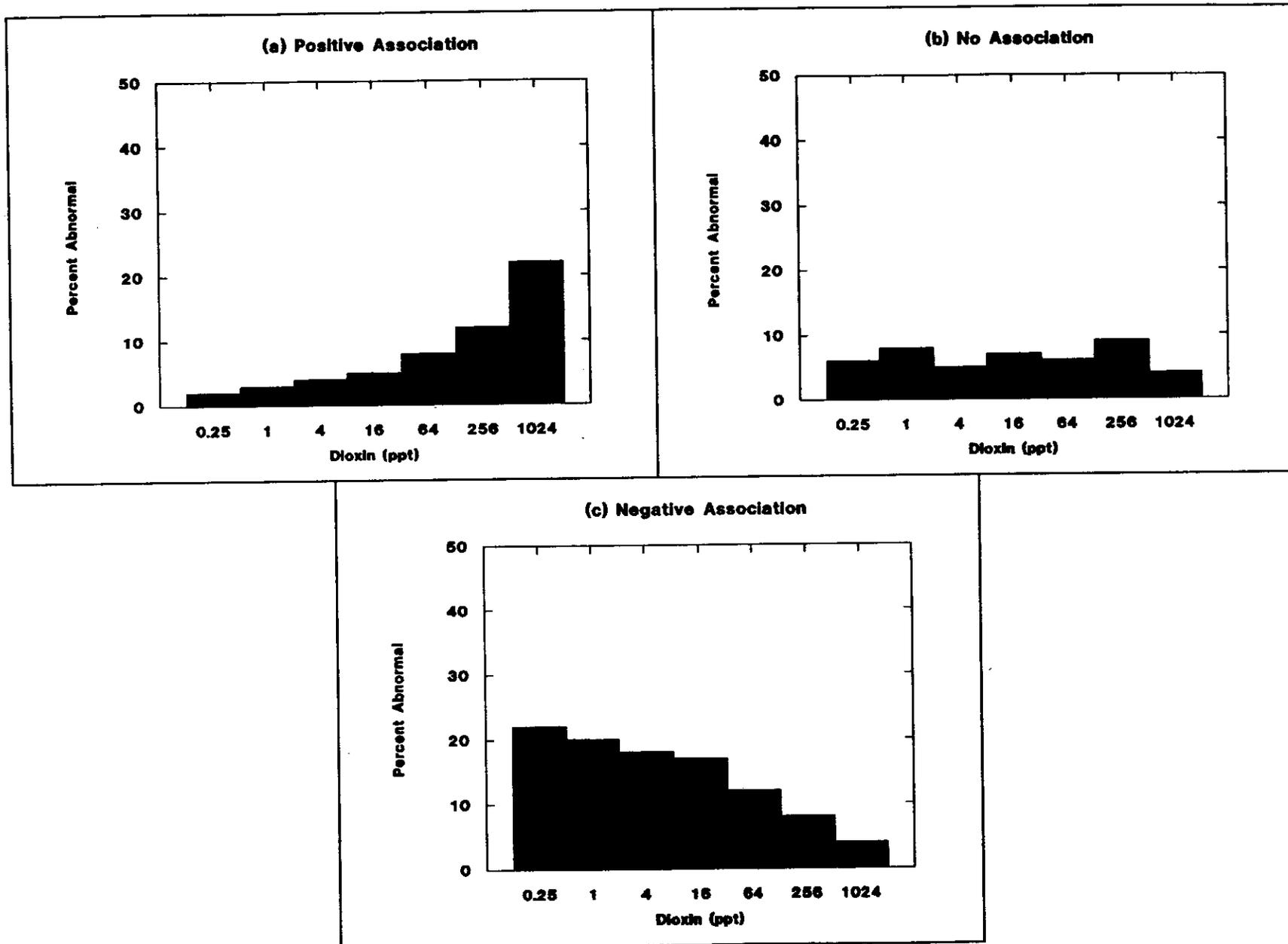


FIGURE 4-5. Hypothetical Data (Discrete Dependent Variable) versus Dioxin

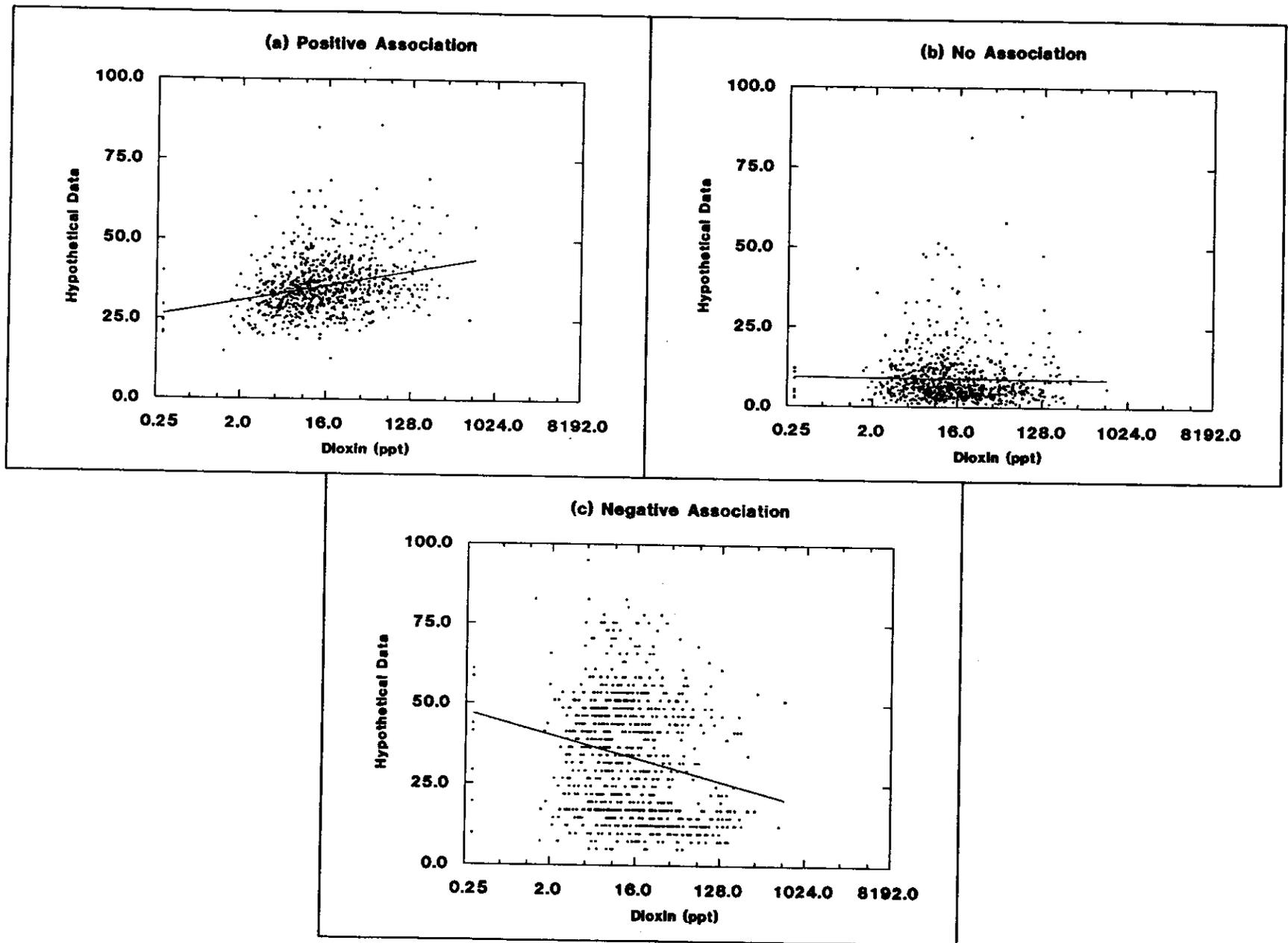


FIGURE 4-6. Hypothetical Data (Continuous Dependent Variable) versus Dioxin

examples of a positive relationship, no relationship, and a negative relationship between a dependent health variable and dioxin.

Interaction Plots/Histograms

Dioxin-by-covariate interactions also were investigated through appropriate graphic displays. Analogous to the data plots/histograms, transformations were used in the presentations when appropriate. If the dependent variable was continuous (e.g., blood urea nitrogen), a significant interaction between dioxin level (e.g., initial dioxin) and a covariate (e.g., age) was presented as a set of bivariate scatterplots (dependent variable versus initial dioxin) for each level of a categorized covariate. For a discrete dependent variable (e.g., kidney disease: yes versus no), a significant interaction between initial dioxin and a covariate was displayed using bar charts at each level of a categorized covariate. The bar charts contrasted percentages of participants classified as abnormal for common interval groupings of initial dioxin.

Statistical Analysis Protocol

Except for changes suggested by the Advisory Committee (deleting conditional analyses and moving fasting glucose from Chapter 10, Gastrointestinal Assessment to Chapter 15, Endocrine Assessment), all statistical analyses summarized in this report were carried out as specified in an analytical plan (17) written in July 1989 and the contract Statement of Work; the analyses began in October 1989 and concluded in November 1990. The analytical plan specified statistical methods, dependent variables, covariates, and exclusions. These analyses did not deviate from those specified in the plan. In certain cases, clarification analyses were carried out, however. Strict adherence to the plan was maintained to avoid the possibility that some analyses might be conducted based on the observation of significant results. Such analyses are called "post hoc" and are known to be biased (18).

CHAPTER 4

REFERENCES

1. Poiger, H., and C. Schlatter. 1986. Pharmacokinetics of 2,3,7,8-TCDD in man. *Chemosphere* 15:1489-94.
2. Pirkle, J.L., W.H. Wolfe, D.G. Patterson, L.L. Needham, J.E. Michalek, J.C. Miner, M.R. Peterson, and D.L. Phillips. 1989. Estimates of the half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Vietnam veterans of Operation Ranch Hand. *J. of Toxicol. and Environ. Health* 27:165-71.
3. 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.
4. 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.
5. Patterson, Jr., 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* 59:2000-05.
6. 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 Environ. Toxicol.* 17:139-43.
7. Kleinbaum, D.G., L.L. Kupper, and H. Morgenstern. 1982. *Epidemiologic research: Principles and quantitative methods*. London: Lifetime Learning Publications.
8. *SAS/STAT guide for personal computers*. 6th ed. 1985. Cary, North Carolina: SAS Institute, Inc.
9. *BMDP statistical software manual*. 1985. Berkeley, California: University of California Press.
10. Feinberg, S.E. 1981. *The analysis of cross-classified data*. Cambridge: MIT Press.
11. Bishop, Y.M.M., S.E. Feinberg, and P.W. Holland. 1975. *Discrete multivariate analysis: Theory and practice*. Cambridge: MIT Press.
12. Kleinbaum, D.G., and L.L. Kupper. 1978. *Applied regression analysis and other multivariable methods*. North Scituate, Massachusetts: Duxbury Press.
13. Neter, J., and W. Wasserman. 1974. *Applied linear statistical models*. Homewood, Illinois: Richard D. Irwin, Inc.
14. Fleiss, J.L. 1981. *Statistical methods for rates and proportions*. 2d ed. New York: John Wiley & Sons.
15. Searle, S.R., F.M. Speed, and G.A. Milliken. 1980. Population marginal means in the linear model: An alternative to least squares means. *The American Statistician* 34:216-21.
16. Wilkinson, Leland. 1988. *SYGRAPH*. Evanston, IL: SYSTAT, Inc.

17. Science Applications International Corporation. 1989. Analytical plan for the blood serum dioxin analysis. McLean, VA: Science Applications International Corporation.
18. Freedman, David A. 1983. A note on screening regression equations. *The American Statistician* 37:152-55.