

## VI. Special Statistical Considerations

The previous discussion has outlined the general statistical approach followed by this protocol, and has outlined planned analytical methods and inferential strategies for the mortality, questionnaire and physical examination study phases. This section provides an indepth consideration of some special statistical study aspects.

### A. False Reporting/Misrepresentation

Since concern for compensation could unconsciously or consciously influence symptom reporting, and since press reporting itself can stimulate anxiety-based symptom formation, a discussion of false reporting is indicated. A data pattern indicating overreporting has already been discussed in Section V. The goal here is to understand the effect of misrepresentation on estimates of relative risk and the odds ratio. Let  $S$  stand for presence of a symptom, and  $\bar{S}$  denote its absence. This false reporting may be represented as in Figure 9.

Figure 9

#### FALSE REPORTING/MISREPRESENTATION

		TRUE STATUS		Total
		S	$\bar{S}$	
REPORTED STATUS	S	A	B	A+B
	$\bar{S}$	C	D	C+D
		A+C	B+D	

The proportion of correctly classified positives is defined by  $A/(A+C)$  and is called the sensitivity of the classification scheme; the proportion of correctly classified negatives  $D/(B+D)$  is called the specificity.

When there is non-differential misrepresentation, that is, when the sensitivity and the specificity are the same among the exposed and nonexposed, the bias induced in the estimate of relative risk will be toward the null value. The situation is summarized by Figure 10.

Figure 10

MISREPRESENTATION IN RANCH HAND II

		TRUE STATUS					
		Exposed			Nonexposed		
		S	$\bar{S}$	TOTAL	S	$\bar{S}$	TOTAL
REPORTED STATUS	S	a	b	a + b	e	f	e + f
	$\bar{S}$	c	d	c + d	g	h	g + h
		a + c	b + d	n	e + g	f + h	n

Using this representation, the true relative risk is  $(a+c)/n \div (e+g)/n$ , and the apparent relative risk is  $(a+b)/n \div (e+f)/n$ . Figure 11 provides a graphic representation of how apparent relative risk varies as a function of specificity. For this curve, the true relative risk is 2 with the exposed population having a symptom incidence of 0.1 and the nonexposed population having a symptom incidence of 0.05 (Copeland et. al. 1977). The effect of nondifferential false reporting on the odds ratio is nearly as severe as that shown in Figure 11 for relative risk. A technique does exist for correcting the estimate of relative risk to account for false reporting, but the technique requires knowledge of the sensitivity and specificity of the classification scheme; knowledge that may not exist in this study. It should be noted that since the above remarks are concerned with relative risk, the number n of subjects in each group is irrelevant, as the results shown are independent of n.

If the false reporting is differential, an estimate of relative risk that is biased away from the null value can result. This will occur in situations in which the RANCH HAND personnel and controls do not misrepresent their symptoms in the same manner (Copeland et. al. 1977). Thus the "true" outcomes of herbicide exposure may be distorted depending upon the degree and direction of misrepresentation.

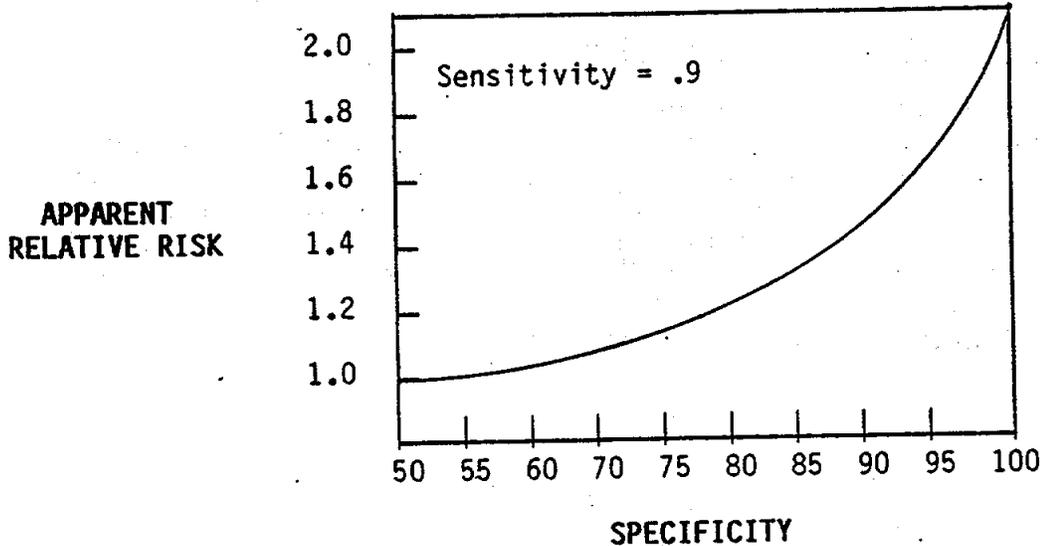
B. Adequacy of Sample Sizes

(1) Overview

The size of the RANCH HAND cohort is approximately 1000 individuals. It is clear that a lethal effect of herbicide which occurs in only 1 out of 2000 controls will be quite difficult to detect unless the herbicide effect is very strong. For example, at a rate of 1 in 2000, 0.5 affected controls are expected. If the basic rate is doubled by herbicide to 2 per 2000, one affected RANCH HAND individual would be expected. At a rate of 1 per 2000 for

Figure 11

APPARENT RELATIVE RISK VERSUS SPECIFICITY



controls and a rate of 2 per 2000 for RANCH HAND personnel, the probability of observing no affected individuals in both groups is

$$(1 - 1/2000)^{1000} (1 - 2/2000)^{1000} = .22$$

or, in other words, "there is a 22% chance" that no affected individuals will be found in this study. In a population of 100,000 exposed individuals, 100 cases would be expected, 50 of which would be due to herbicide. In short, since the size of the RANCH HAND group is fixed, this study has limited statistical power to define the relationship of herbicide to the rarer diseases.

The power (1-β) of a study design is the probability that a specified difference between populations will be detected if it in fact exists. In general, power is a direct function of sample size; that is, for a particular study design, the more subjects measured the larger the study power. It is understood that this protocol makes use of the entire known RANCH HAND population (and excludes ancillary exposed groups for reasons previously cited); the exposed sample size cannot be increased. Power augmentation, therefore, can only be accomplished by the less efficient procedure of increasing the control group size which has statistical limitations as well as staggering financial and logistic considerations. Hence, considerable effort has been made to correct loss to study issues (by replacement and other techniques to induce participation) and to use the most powerful statistical design concepts.

Essentially all previous animal and human studies concerning herbicide suffer from a lack of adequate consideration of study power. The following presents a preliminary analysis of study power for the case of continuous and dichotomous variables expected from the study. Also reviewed are alternative studies involving Marine samples.

(2) Power in Continuous Variable Case

Assume that blood cholesterol levels are being compared between RANCH HAND and control groups, and that the coefficient of variation for cholesterol in the control group is 0.1, where the coefficient of variation is the ratio  $\sigma_C/\mu_C$ . Assume  $\sigma_{RH} = \sigma_C$ . The symbol  $\alpha$  is the probability that the study will indicate an effect where none exists, and  $1-\beta$  is the power as defined before. Consider that the RANCH HAND mean cholesterol  $\mu_{RH}$  is shifted from the control mean  $\mu_C$ . A natural question is to inquire about the study power as a function of available pairs (n) and mean ratio  $\gamma = \mu_{RH}/\mu_C$ .

Table 11

POWER CALCULATIONS

ASSUMPTIONS:  $\alpha=0.05$ ,  $\sigma_C/\mu_C=0.1$ ,  $\gamma=\mu_{RH}/\mu_C$

<u>r</u>	<u><math>\gamma</math></u>	<u>Power = 1-<math>\beta</math></u>	
		<u>n=180</u>	<u>n=450</u>
.20	1.01	.20	.38
.20	1.02	.55	.88
.20	1.05	>.995	>.995
.70	1.01	.86	.995
.70	1.02	>.995	>.995
.70	1.05	>.995	>.995

Power calculations are displayed in Table 11. Study power in the case of a matched pair design is strongly dependent on the degree of positive correlation produced between the involved groups by the matching procedure. Of course, the degree of correlation can be expressed by the correlation coefficient r which can take values between -1 (negative correlation) and +1 (positive correlation), and two values of r have been employed in Table 11. From this table it is seen that if only 450 pairs are studied a 1% shift in mean (= 1.01) will not be reliably detected, but a 2% shift will be detected with a probability of 0.88 if r = 0.2 at least. From this calculation one can infer the need to examine at least 450 pairs to obtain the 2% shift, and to strive for more if possible.

(3) Power in the Dichotomous Variable Case

There is significant discussion in the mathematical statistics literature concerning the efficacy of paired designs in the setting of dichotomous responses (Billewicz, 1974; Ury, 1975; Miettinen, 1970; and several others). Table 12 shows a set of calculations which are applicable to the present study.

Table 12

POWER CALCULATIONS FOR THE DICHOTOMOUS VARIABLE CASE AS A FUNCTION OF EFFICACY OF PAIRED DESIGNS

				POWER = 1 - $\beta$			
$P_1$	$P_2$	Rel. Risk	$r$	$n=250$	$n=350$	$n=450$	
.05	.01	5	0	.77	.82	.92	↑
.04	.01	4	0	.61	.75	.85	
.03	.01	3	0	.40	.51	.59	
.10	.05	2	0	.61	.75	.85	
.20	.10	2	0	.87	.94	.97	
***							
.05	.01	5	.1	.89/.029	.94/.032	.98/.064	↑
.04	.01	4	.1	.72/.033	.87/.038	.88/.041	
.03	.01	3	.1	.38/.020	.68/.046	.71/.077	
.10	.05	2	.1	.76/.055	.85/.048	.88/.048	
.20	.10	2	.1	.94/.043	.98/.046	.99/.057	

\* $\alpha$  = .050

\*\* $\alpha$  = as indicated

In this table,  $r$  is again the correlation coefficient indicating the degree of correlation induced between the involved groups by the matching procedure. The probability of the disease among RANCH HAND personnel is symbolized as  $p_1$ , while  $p_2$  is the probability of the disease among the controls. Relative risk is the ratio  $p_1/p_2$ . With  $r = 0.1$ , sign test power tables were used as an exact version of McNemar's test, and therefore different  $\alpha$  levels are shown under each power number. Table 12 shows the positive influence of effective

pairing in the higher power levels noted. Also, it appears that for  $p_2 = 0.01$  and  $p_1 = 0.03$ , physical examination of 450 pairs (900 examinations) will disclose the three-fold relative risk with probability less than the minimum target .80. In other words, there is a greater than "20% chance" that a three-fold relative risk on a 1/100 disease state will go undetected in this study if only 350 pairs are examined and if low correlations occur. Once again the need to examine the maximum numbers of pairs in the study is seen.

To present these dichotomous power calculations more clearly, calculations in the context of actual disease states have been accomplished. The diseases considered are cardiovascular disease and cancer, corresponding to high and low rate illnesses for the age groups presently under investigation.

#### (a) Cardiovascular Disease

A logistic risk function was fitted to data from 17,455 autopsies gathered in a WHO collaborative study in Czechoslovakia, Sweden and the USSR. The function fitted has the form

$$P = [1 + \exp(\alpha + \beta(x-.5) + \gamma(y-.5))]^{-1}$$

where

$p$  = the probability of a complicated coronary lesion

$x$  = age scaled linearly so that  $x = 0$  is equivalent to 30 years, and  $x = 1$  is equivalent to 58 years (the age span of the current study)

$y = 1$  or  $0$  if the subject is exposed or not

and  $\alpha$  and  $\beta$  were obtained from the data. The function represents a fairly high rate disease in that at 40 years of age 7% of the group had the lesion, and at 60 years of age 20% had the lesion. The coefficient  $\gamma$ , represents the exposure effect. Power calculations for  $\gamma = \beta$  and  $\gamma = .8\beta$  are shown in Table 13. This table suggests that if, as a cell toxin, herbicide exposure accelerates cardiovascular disease, this study has a good chance of detecting that acceleration if the herbicide effect is comparable to the age effect. A slight beneficial effect of pairing is seen in this hypothetical example.

#### (b) Cancer

A logistic risk function was fitted to breast cancer data presented by Breslow and Day (1975). The function fitted represents a low rate disease in that at 35 years of age only .000336 of the group had the lesion while at 70 years of age .00676 of the group will have the lesion.

Using pairing to achieve a power of 0.80 in this setting, 1312 pairs would be needed, when the exposure effect is equal to the age effect. This exceeds the size of our RANCH HAND cohort, and reinforces the fact that herbicide exposure effects on rarer diseases will not have a high likelihood of being detected by this study, and again supports an attempt to examine as many pairs as possible.

Table 13

POWER CALCULATIONS AS A FUNCTION OF HERBICIDE EFFECT

ASSUMPTION:  $\alpha = 0.05$

Number of Pairs	$\gamma = \beta$		$\gamma = .88$	
	Power Neglecting Pairing	Power With Pairing	Power Neglecting Pairing	Power With Pairing
100	.93	.93	.64	.53 ( $\alpha = .036$ )
160	>.97	.98	.81	.82
200	>.99	>.995	.86	.87
250	>.99	>.995	.93	.95
300	>.99	>.995	.96	.97
350	>.99	>.995	.97	.98

(3) Alternative Studies Using Marine Cohorts

The GAO and the National Academy of Sciences have referred to specific Marine cohorts as candidates for a Herbicide Orange epidemiological study. In one suggested study configuration, 5900 marines who were within one half kilometer of a herbicide spray track on the day of spraying are called the exposed group, while 212,100 marines are considered unexposed. In a second suggested study configuration, 21,900 marines within one half kilometer of a spray path within 4 weeks of spraying are considered exposed, while a remaining 196,100 marines are considered unexposed. A mortality study was proposed in both of these study configurations. The mortality phase of this protocol involves approximately 1200 exposed and 6000 control individuals, so that, on the surface, the Marine studies would appear to be more powerful in a statistical sense due to larger numbers. However, in fact, two factors couple to render the marine studies less powerful than the RANCH HAND study detailed in this protocol. First, calculations show that a soldier standing directly

under a spray track at the exact time of spraying receives approximately 1/1000 the dose received by RANCH HAND individuals repeatedly disseminating the mixture throughout the usual RVN tour. Thus even if the unlikely event of being directly under a spray path were repeated 10 times during a marine's RVN tour, the marine's dose would still be only 1/100 that of the RANCH HANDERS. The second factor impacting the Marine study power is the difficulty imposed by the fact that troop positions are only very inexactly known. The available data provide only the battalion headquarter's position relative to herbicide spray paths. Thus troops considered to be exposed could be very far from spray paths, and in fact, be unexposed. On the other hand, troops deemed unexposed in terms of their battalion headquarter's position could in fact have been near spray paths on the day of spraying. Thus, the Marine studies are limited by the problem of misclassification in addition to the fact that the marines received a lesser herbicide exposure than RANCH HAND personnel.

It is possible to compare the RANCH HAND study described in this protocol with the Marine studies in a quantitative way. Results of such an analysis are set out in Tables 14 thru 17. In Table 14, the Marine study using 5900 exposed soldiers is contrasted with the RANCH HAND study considering a disease with an incidence of 0.001 in the control groups, and 0.004 in the RANCH HAND exposed cohort. With a relative risk of 4 against a control disease incidence of 0.001, RANCH HAND power is 0.87 while the Marine study power is much less for several combinations of Marine exposure and misclassification. The misclassification figures shown refers to the percentage inclusion of unexposed individuals into the exposed Marine group. For the calculations, disease incidence in the marine exposed group was assumed to be linearly related to exposure. Table 15 is strictly analogous to Table 14 except that the disease state studied has an incidence of 0.01 in the control groups and 0.02 in the RANCH HAND exposed cohort. Again the RANCH HAND study is seen to be significantly more powerful than the Marine study. Tables 16 and 17 directly parallel Tables 14 and 15, respectively, except that the Marine exposed group is considered to consist of 21,900 soldiers. Here again RANCH HAND study power is seen to be significantly superior.

Figure 12 shows the RANCH HAND mortality study power as a function of relative risk, and disease incidence in the control group. Figure 13 shows marine study power versus marine exposure for zero to 25% misclassification and a control disease incidence of 0.001 and RANCH HAND relative risk of 4. For this circumstance it is clear that the marine study becomes competitive with the RANCH HAND power only if one assumes that the marines received approximately one half of the RANCH HAND exposure dose. Figure 14 is the same as Figure 13 except that 21,900 marines are considered exposed. Again the Marine study becomes competitive with the RANCH HAND study only if one can assume the exposed marines received 0.2 or more of the RANCH HAND exposure, an assumption which is not supported by the available data.

### C. The Replacement Concept

In the mortality analysis, a randomly selected group of control individuals will be compared to the RANCH HAND group, and the data gathered will be analyzed for evidence of herbicide effect. In the questionnaire and

TABLE 14

**MORTALITY ANALYSIS****POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION  
CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE**

POWER TABLE

RANCH HAND POWER 1-B	% MISCLASSIFICATION	MARINE STUDY POWER			
		EXPOSURE LEVELS RELATIVE TO RANCH HAND			
		1/10	1/20	1/100	1/1000
.87	0	.18	.10	.06	.05
	10	.16	.09	.06	.05
	25	.15	.09	.06	.05

ASSUMPTIONS: RH STUDY POP. 1,200: 6,000 (1:5)  
 MARINE STUDY POP. 5,900: 212,100  
 NORMAL INCIDENCE OF DISEASE 0.001  
 DISEASE INCIDENCE IN RH 0.004  
 LINEAR DOSE - RESPONSE  
 MISCLASS. OF MARINE CONTROLS EXCLUDED

TABLE 15

## MORTALITY ANALYSIS

### POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE

POWER TABLE

RANCH HAND POWER 1-B	% MISCLASSIFICATION	MARINE STUDY POWER EXPOSURE LEVELS RELATIVE TO RANCH HAND			
		1/10	1/20	1/100	1/1000
		.92	0	.19	.10
	10	.17	.10	.06	.05
	25	.14	.09	.06	.05

ASSUMPTIONS: RH STUDY POP. 1,200: 6,000 (1:5)  
 MARINE STUDY POP. 5,900: 212,100  
 NORMAL INCIDENCE OF DISEASE = 0.01  
 DISEASE INCIDENCE IN RH = 0.02  
 LINEAR DOSE - RESPONSE  
 MISCLASS. OF MARINE CONTROLS EXCLUDED

TABLE 16

**MORTALITY ANALYSIS****POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE  
POPULATION CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE \***

POWER TABLE

RANCH HAND POWER 1-B	% MISCLASSIFICATION	MARINE STUDY POWER			
		EXPOSURE LEVELS RELATIVE TO RANCH HAND			
		1/10	1/20	1/100	1/1000
.87	0	.38	.17	.07	.05
	10	.33	.15	.06	.05
	25	.26	.13	.06	.05

ASSUMPTIONS: RH STUDY POP. 1,200; 6,000 (1:5)  
 MARINE STUDY POP. 21,900; 196,100  
 NORMAL INCIDENCE OF DISEASE = 0.001  
 DISEASE INCIDENCE IN RH = 0.004  
 LINEAR DOSE - RESPONSE  
 MISCLASS. OF MARINE CONTROLS EXCLUDED

\* INCORRECT POPULATION  
 NUMERICS BASED ON  
 ENVIRONMENTAL FATE  
 OF TCDD

TABLE 17

**MORTALITY ANALYSIS****POWER COMPARISON OF THE RANCH HAND STUDY TO THE MARINE POPULATION  
CONSIDERING MISCLASSIFICATION AND RELATIVE EXPOSURE \***

POWER TABLE

RANCH HAND POWER 1-B	% MISCLASSIFICATION	MARINE STUDY POWER			
		EXPOSURE LEVELS RELATIVE TO RANCH HAND			
		1/10	1/20	1/100	1/1000
.92	0	.41	.17	.07	.05
	10	.36	.16	.07	.05
	25	.28	.13	.06	.05

ASSUMPTIONS: RH STUDY POP. 1,200; 6,000 (1:5)  
 MARINE STUDY POP. 21,900; 196,100  
 NORMAL INCIDENCE OF DISEASE = 0.01  
 DISEASE INCIDENCE IN RH = 0.02  
 LINEAR DOSE - RESPONSE  
 MISCLASS. OF MARINE CONTROLS EXCLUDED

\* INCORRECT POPULATION  
 NUMERICS BASED ON  
 ENVIRONMENTAL FATE  
 OF TCDD

FIGURE 12

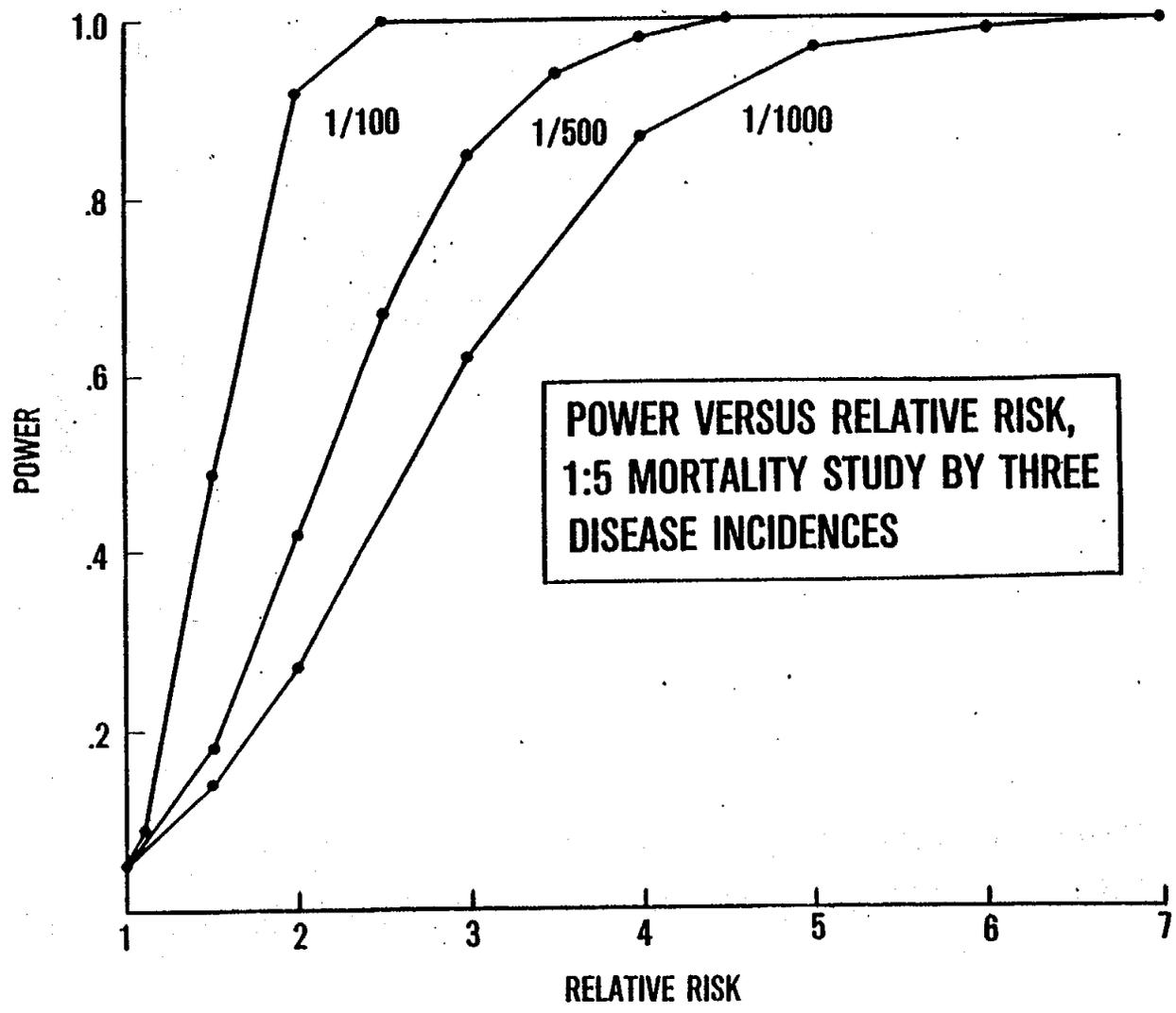


FIGURE 13

### POWER CURVES OF THE MARINE STUDY CONSIDERING RELATIVE EXPOSURE AND MISCLASSIFICATION OF THE STUDY POPULATION

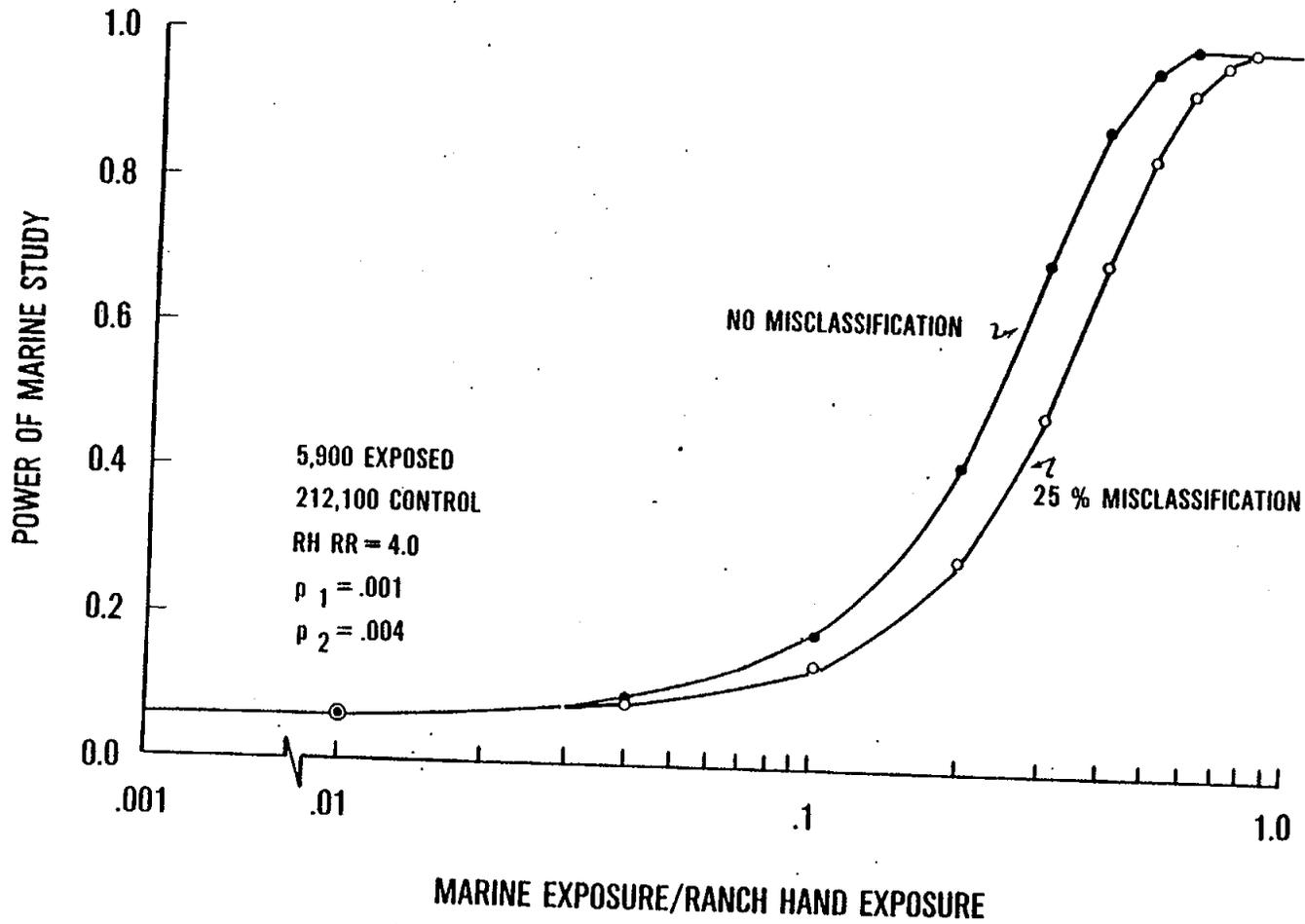
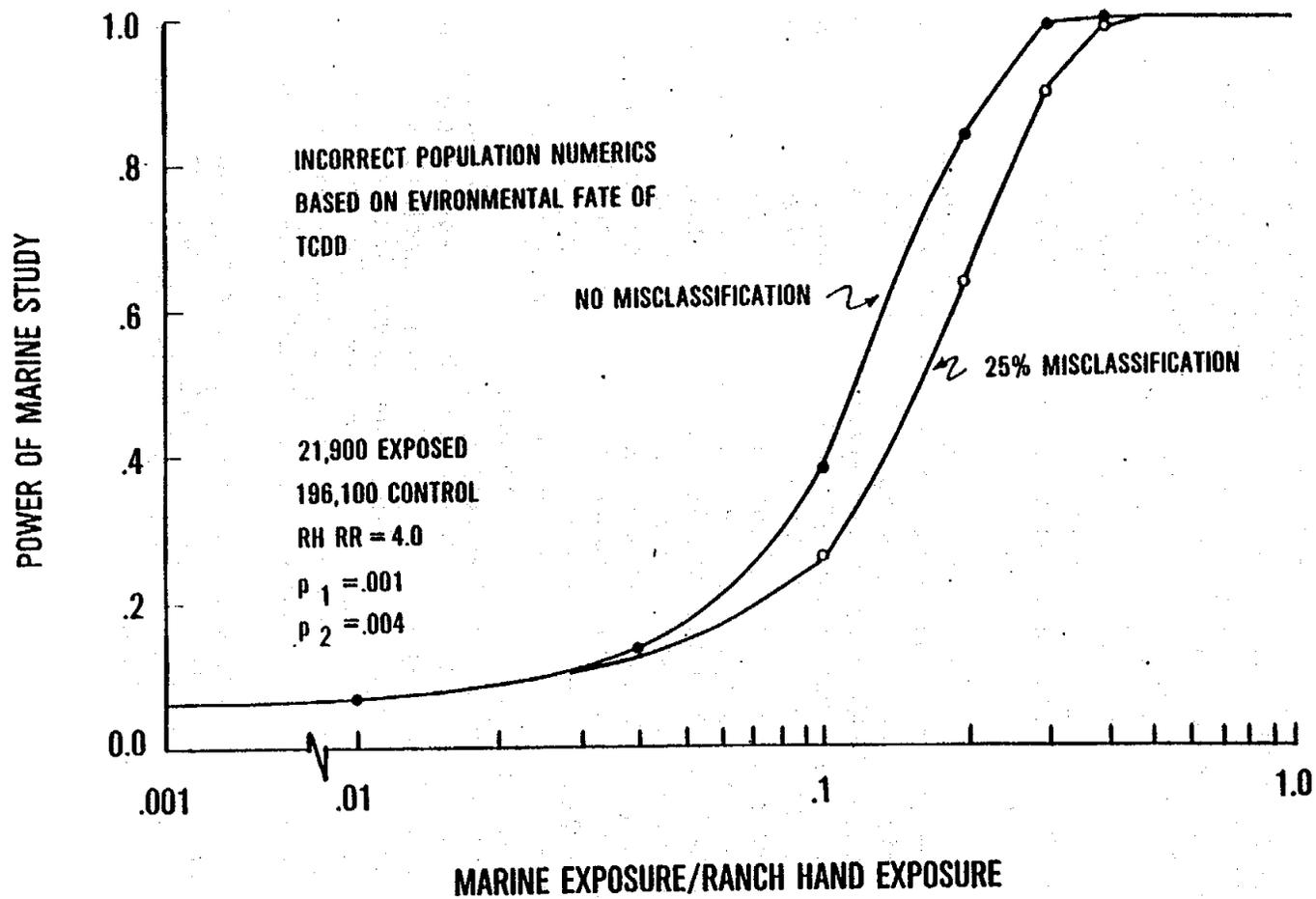


FIGURE 14

### POWER CURVES OF THE MARINE STUDY CONSIDERING RELATIVE EXPOSURE AND MISCLASSIFICATION OF THE STUDY POPULATION



physical examination phases of this study, one of the mortality controls will be randomly selected for each RANCH HAND individual. During the physical examination phase, we must anticipate a significant degree of unwillingness to participate, particularly on the part of control personnel. This loss to study can result in significant bias and loss in statistical power; thus the replacement concept has been developed to mitigate these consequences.

In this replacement strategy, we make use of the control individuals matched with each RANCH HAND person. As previously noted, this is accomplished using computerized data files and the matching parameters of age, AFSC, and race. With each RANCH HAND individual  $R_i$  there will be associated ten controls  $C_{j1}, C_{j2}, C_{j3}, \dots, C_{j10}$ . The first of these controls,  $C_{j1}$  will be employed in the questionnaire and physical examination phases of the study. If  $C_{j1}$  is alive, but unwilling to participate in the study, he will be replaced by another randomly selected participant with similar perception of health status. In order to avoid bias in morbidity analyses, no dead control will be replaced.

It is important to emphasize that all replacement controls will be carefully flagged so that they may be treated separately in the statistical analysis. These replacements will be carefully compared to the lost controls to develop indicators of comparability (e.g., morbidity and mortality experience). The initial analysis will be performed on the intact exposed/control pairs. Additional analysis will be conducted on all pairs, both those intact, and those with replaced controls. If we consider RANCH HAND individual  $R_i$ , with living control  $C_{j1}$ , we can calculate the probability that control  $C_{jk}$  will be available for the 1st, 2nd and 3rd physical examinations. To examine this question, a small computer Monte Carlo simulation was required. A short BASIC language computer program and glossary are included in Appendix Table A-8. This simulation examines the effect of non-participation expressed as two probabilities  $P_1$  and  $P_2$ . Figure A-2 displays the expected participation by the RANCH HAND population, and control group participation is expected to be somewhat less.  $P_1$  is the probability that when first asked to attend a physical examination, the control individual will not comply.  $P_2$  is the probability that a control individual who has agreed once to a physical examination, will not comply for a subsequent examination. In general,  $P_1$  may be greater than  $P_2$ . Note that the probabilities  $P_1$  and  $P_2$  must reflect all causes of non-compliance including morbidity and mortality. Table 18 displays a representative simulation run, which provides the number of controls required to find willing matches for 1000 RANCH HAND personnel.

The potential bias introduced by non-willingness in controls can be analyzed statistically. If  $P_C(x)$  is the probability density function for compliant individuals and  $P_{NC}(x)$  is the same function for non-compliant individuals, we have

Table 18

CONTROL DISTRIBUTIONS BY EXAMINATION  
MATCHING 1000 RANCH HAND PERSONNEL

( $P_1 = .70, P_2 = .25$ )

CONTROL COHORT	EXAMINATION NUMBER		
	1	2	3
C <sub>1</sub>	318	237	177
C <sub>2</sub>	211	188	156
C <sub>3</sub>	131	133	136
C <sub>4</sub>	96	101	97
C <sub>5</sub>	74	89	90
C <sub>6</sub>	49	68	77
C <sub>7</sub>	34	43	59
C <sub>8</sub>	25	39	52
C <sub>9</sub>	16	18	33
C <sub>10</sub>	13	20	35
Number of Matching Failures	33	64	88

$$p(x) = \alpha p_C(x) + \beta PNC(x)$$

where  $p(x)$  is the probability density function for the entire population and  $x$  is a vector of important health parameters available on each person. Since

$$\int p(x)dx = \int p_C(x)dx = \int PNC(x)dx = 1$$

it follows that

$$\alpha + \beta = 1$$

and  $\alpha$  and  $\beta$  may be viewed as coefficients which "mix" the two subpopulations.

If  $M_C$  and  $M_{NC}$  are the means of the compliant and non-compliant subpopulations respectively, it can be shown that

$$M = \alpha M_n + \beta M_{NC}$$

where  $M$  is the mean of the entire population. From this last equation, it is clear that as noncompliant individuals are lost (i.e.,  $\beta$  tends to zero,  $\alpha$  tends to one),  $M$  tends to  $M_C$ . Thus the maximum bias is the quantity  $M_C - M$ .

In this study we propose to replace non-compliant control individuals with matched RANCH HAND control individuals, that is with individuals drawn from a population with density equal to or at least similar to  $p_{NC}(x)$ . The resulting new density is  $p''(x)$  such that

$$p''(x) = \alpha'' p_n(x) + \beta'' \tilde{p}_{NC}(x)$$

where

$$\alpha'' + \beta'' = 1$$

$$M'' = \alpha'' M_n + \beta'' \tilde{M}_{NC}$$

and where  $\tilde{p}_{NC}(x)$  approximates  $p_{NC}(x)$ . If  $\beta''$  is chosen to be close to or equal to  $\beta$  above, it appears that  $M''$  can well approximate  $M$ , the true population mean. The difficulty in this approach will be to assure that the replacements are representative of the non-compliant individuals in all respects other than logistic factors impacting willingness to participate in the program.

Our proposed approach is to obtain sufficient data on the unwilling personnel so that a discrimination function of the form

$$D = f(h_1, \dots, h_j; l_1, \dots, l_j)$$

can be derived. This function is envisioned to have the following properties:

(a) larger values of D correspond to decreasing probabilities of compliance with the physical examination,

(b) the factors  $h_i$  relate to the subjects' health status, while the factors  $l_i$  relate to logistic difficulties (distance, job) which tend to preclude attendance at the physical. Factors to be considered in the formulation of this function are displayed in Table 19.

(c) D is an increasing function of each  $h_i$  and of each  $l_i$ ,

Table 19

FACTORS AFFECTING COMPLIANCE

<u>Health Status (<math>h_i</math>)</u>	<u>Logistic Difficulties (<math>l_i</math>)</u>
Subjective Health Assessment (good/poor)	Time Away from Family
Current Utilization of Long-Term Health Care (Yes/No)	Time Away from Job
Absenteeism Pattern (Greater Than/Less Than Ten Lost Days in Past Six Months)	Distance to Examination Site
	Active Pilot
	Income (Greater than/Less than \$17,000)

In the replacement scheme, controls substituted for noncompliant controls, should have identical health factors ( $h_i$ ) as those individuals they replace. The only significant differences should be in the logistic factors ( $l_i$ ). The replacement method should permit correction of non-compliance bias given that health factors  $h_i$  and logistic factors  $l_i$  are actually distinct. The determination of these two classes of factors will be made using data from the study itself. Specifically, the logistic factors  $l_i$  will be independent of health status to the degree testable by the quantity of data available in the study. This replacement strategy has two major advantages: selection bias reduction/estimation and cost reduction. Were replacements not employed, one would be compelled to start the morbidity study with a 4 to 1 or 5 to 1 design in order to insure an adequate number of participating controls on the third physical examination (see Table 18). Such a large control group for physical examination is very costly with little

corresponding gain in study power and with no correction of the selection bias.

#### D. Statistical Analysis of Large Data Sets

A large amount of data will be collected on each subject in this study. Testing at the 0.05  $\alpha$  level means that in 5 out of 100 instances where there has actually been no herbicide effect, a herbicide effect will be falsely inferred. This is the inverse of the power question which concerns the probability of detecting an event when it actually occurs. If 100 independent measures are taken from subjects one should expect, testing at the 0.05  $\alpha$  level, that five measures will be positive on the average. This awareness itself should help prevent over reaction to isolated findings. Further, the present protocol does not in fact have one hundred independent measures. Rather the data gathered are grouped into correlated batteries or systems of data. Findings with any given measure will be related to the values of other correlated variables to provide substantiation indicating an authentic finding.

#### E. Time-In-Study Effects

The study outlined in this protocol is expected to involve up to six examinations extending over a period of twenty years. It could be anticipated that participation in the study, by increasing the health awareness of the subjects, would tend to improve the health of the cohorts. The possibility of differential participation in the study by the exposed and control groups could bias against finding a herbicide effect if one exists. The control group could be less willing to participate in the study than will the exposed RANCH HAND personnel. Thus, if on the average, controls spend less time in the study than RANCH HANDERS, and under the supposition that increased time in study will correlate with better health, increased RANCH HAND participation would counterbalance any adverse herbicide health effect.

The corrector for this time-in-study effect is simply to study the relationship between health outcome and participation in the RANCH HAND study by regression or other analogous statistical methods. Participation can be quantitated by such metrics as (a) number of physical examinations attended (b) age at physical examinations attended or (c) pattern of physical examination attendance. Special study design features do not need to be incorporated to properly evaluate time-in-study effects on questionnaire and physical examination portions of the study. However, the effects of differential time-in-study on the mortality analysis must be carefully considered. In order to detect time-in-study effects on mortality, individuals whose mortality are being tracked should have been in the study for the same length of time (both exposed and control individuals), or the distribution of time spent in the study should be similar in both groups. Because of anticipated differential participation between the exposed and control groups, one cannot assume that both cohorts will have equal time in study distributions. Steps must be taken to insure that a proper time-in-study distribution occurs in the control mortality group. Control over this distribution is possible through placement of the mortality cohort in the structure of the control group with

respect to the replacement strategy. The following five designs have been considered:

- I. mortality subjects randomized over all ten control positions, and therefore called into the study randomly.
- II. mortality subjects in the first five control positions, and therefore called into the study first.
- III. mortality subjects in positions #1 and #2, with the three remaining subjects randomized into positions #3 through #10.
- IV. mortality subjects in positions #1, #2, #9, and #10, with the one remaining subject randomized in positions #3 through #8.
- V. mortality subjects in the first four positions and position #10.

For each of these five designs, certain quantities were calculated. For testing a physical examination effect on mortality, one would require adequate numbers of mortality subjects having had all six physical examinations, and adequate numbers having had none. Therefore, assuming 1200 RANCH HAND subjects,

E1 = expected number of mortality subjects having all six physical examinations.

E2 = expected number of mortality subjects never asked to take the physical examination.

E3 = expected number of mortality subjects having taken no physical examinations.

For testing or modeling time-in-study effects, one would want adequate numbers of mortality subjects having only one physical, having exactly two physicals, etc. Hence, we calculate, for  $J = 1, 2, 3, 4, 5, 6$ :

NJ = expected number of mortality subjects taking exactly J physicals (for example N3 is the number of mortality subjects who will have taken three physicals by the end of the study).

and

MJ = expected number of mortality subjects which will actually have taken examination J.

The values of E1, E2, E3, NJ, and MJ have been calculated for the five study designs outlined above using an adaptation of the Monte Carlo program

shown in Appendix Table A8. Best case and worst case situations were considered. In the worst case, it was assumed that when first asked to participate, 75% of the subjects refused, while when asked after having once participated, 50% of subjects refused further contact. In the best case, the first time refusal rate was assumed to be 50%, and the refusal rate for a subject who had participated in a prior examination was assumed to be only 15%. Table 20 shows the calculated results. In examining this table it is of interest to note that the calculated values are not strikingly dependent on study design configuration. However, for both the worst and best cases, design 2 where the mortality subjects are placed in the first five control positions, appears superior and will be used in this study.

Table 20. TIME-IN-STUDY EFFECTS

DESIGN PARAMETERS	WORST CASE					BEST CASE				
	1	2	3	4	5	1	2	3	4	5
E1	18	29	25	22	27	267	521	454	428	504
E2	765	194	580	865	493	3953	2409	3137	3478	2690
E3	4691	4548	4645	4716	4623	4977	4204	4569	4739	4345
N1	700	751	713	681	714	227	370	284	239	331
N2	340	374	347	328	355	185	307	284	239	331
N3	163	188	168	157	177	146	249	189	161	225
N4	71	84	76	72	79	116	203	157	136	188
N5	33	43	36	31	39	94	171	129	110	160
N6	18	29	25	22	27	267	521	454	428	504
M1	18	29				267	521			
M2	14	14				46	77			
M3	24	17				53	83			
M4	40	25				62	94			
M5	54	28				76	107			
M6	80	34				85	116			