

**Application of the SUDAAN<sup>®</sup> Software Package  
to Clustered Data Problems:**

***Pharmaceutical Research***

*by*

***Gayle S. Bieler  
Rick L. Williams  
Research Triangle Institute***

*Prepared for:  
New Jersey Chapter of the ASA  
May 12, 1997*

*Originally presented to:  
US Food and Drug Administration  
Washington, DC  
February 9 and June 21, 1996*

***1996 Joint Statistical Meetings  
Computer Technology Workshop  
Chicago, Illinois  
August, 1996***

*Application of the SUDAAN Software Package to Clustered Data Problems in  
Pharmaceutical Research* was written by **Gayle S. Bieler and Rick L. Williams**

Copyright 1997 by Research Triangle Institute  
P.O. Box 12194  
Research Triangle Park, NC 27709-2194

All rights reserved. No part of this publication may be reproduced or transmitted by any means without permission from the publisher.

SUDAAN and RTI are trademarks of the Research Triangle Institute.  
SAS is a trademark of SAS Institute, Inc.

# SUDAAN

*Software for the Statistical Analysis of Correlated Data*

## **Application of the SUDAAN<sup>®</sup> Software Package to Clustered Data Problems:**

### ***Pharmaceutical Research***

***Gayle S. Bieler***  
*gbmac@rti.org*

***Rick L. Williams***  
*willy@rti.org*  
*Research Triangle Institute*

*Prepared for:*  
***New Jersey Chapter of the ASA***  
*May 12, 1997*

*Originally presented to:*  
***US Food and Drug Administration***  
*February 9 and June 21, 1996*

***1996 Joint Statistical Meetings***  
*Chicago, Illinois*  
*August, 1996*



Research Triangle Institute  
PO Box 12194  
Research Triangle Park, NC 27709-2194

# ***Application of the SUDAAN Software Package to Clustered Data Problems: Pharmaceutical Research***

## **Table of Contents**

---

Presentation Notes .....	1-63
References .....	64-68
Appendix I: Abstracts from Related Papers .....	69-76
<b>Appendix II: Examples Using the SUDAAN Software Package .....</b>	<b>79-185</b>
<b>Example 1. Teratology:</b> Effect of the Compound DEHP on Fetal Death .....	79-100
<b>Example 2. Clinical Trials:</b> Effect of Treatment on Repeated Exercise Times to Angina Pectoris .....	101-130
<b>Example 3. Clinical Trials:</b> Comparison of Two Inhaler Devices in a Cross-Over Clinical Trial .....	131-160
<b>Example 4. Teratology:</b> Effect of Boric Acid on Fetal Body Weight .....	161-185
Appendix III: Comparisons to Other Methods .....	188-194

## ABSTRACT

In the pharmaceutical sciences, researchers often encounter data which are observed in clusters. Individual responses may represent multiple outcomes from the same patient (such as sets of teeth, pairs of eyes, or longitudinal outcomes on the same individual) or from multiple patients within a larger cluster, such as a physician clinic or an animal litter. Intracluster correlation, or the potential for clustermates to respond similarly, poses special problems for statistical analysis. This occurs because experimental units from the same cluster are not statistically independent. Failure to account for the cluster effect in the statistical analysis can result in underestimated standard errors and false-positive test results. In addition, cross-over clinical trials will not yield the associated increase in statistical power if the design is ignored in the analysis.

This seminar will describe the statistical methods used in SUDAAN and demonstrate its use via a series of examples. The concept underlying the statistical methods is to fit *marginal* or *population-averaged* regression models (linear, logistic, multinomial logistic, and proportional hazards models) via Generalized Estimating Equations, treating the intracluster correlation as a nuisance parameter. Robust variances estimators (also known as sandwich estimators) ensure consistent variance estimates and valid inferences even when the correlation structure has been misspecified. The methods can also be used for descriptive data analysis. Examples from clinical trials, teratology, and developmental toxicology experiments will be presented.

## Workshop Outline

---

- Clustered Data Applications
  
- Problems associated with statistical analysis of clustered data
  - Why SUDAAN?
  - When SUDAAN?
  
- Summarization of the statistical techniques
  
- SUDAAN Capabilities
  
- Demonstration of the cluster sample techniques using SUDAAN: Examples from clinical and toxicologic research

## SUDAAN Procedures

---

### DESCRIPTIVE PROCEDURES

#### CROSSTAB

Computes frequencies, percentage distributions, odds ratios, relative risks, and their standard errors (or confidence intervals) for user-specified cross-tabulations, as well as chi-square tests of independence and the Cochran-Mantel-Haenszel chi-square test for stratified two-way tables.

#### DESCRIPT

Computes estimates of means, totals, proportions, percentages, geometric means, quantiles, and their standard errors; also computes standardized estimates and tests of single degree-of-freedom contrasts among levels of a categorical variable.

#### RATIO

Computes estimates and standard errors of generalized ratios of the form  $\Sigma y / \Sigma x$ , where  $x$  and  $y$  are observed variables; also computes standardized estimates and tests single-degree-of-freedom contrasts among levels of a categorical variable.

### REGRESSION PROCEDURES

#### REGRESS

Fits linear regression models and performs hypothesis tests concerning the model parameters; uses GEE to efficiently estimate regression parameters, with robust and model-based variance estimation.

#### LOGISTIC

Fits logistic regression models to binary data and computes hypothesis tests for model parameters; also estimates odds ratios and their 95% confidence intervals for each model parameter.

#### MULTILOG

Fits multinomial logistic regression models to ordinal and nominal categorical data and computes hypothesis tests for model parameters; estimates odds ratios and their 95% confidence intervals for each model parameter; uses GEE to efficiently estimate regression parameters, with robust and model-based variance estimation.

#### SURVIVAL

Fits discrete and continuous proportional hazards models to failure time data; also estimates hazard ratios and their 95% confidence intervals for each model parameter.

## Clustered Data Applications

---

### Pharmaceutical Research

#### Toxicology / Preclinical Studies

- **Developmental toxicity**  
Presence of malformations and death recorded on fetuses clustered within litters (*Cluster = litter*)
- **Neurobehavioral toxicity**  
Recurrent failure times recorded over a series of trials on each animal (*Cluster = animal*)

#### Clinical Trials

- **Periodontal / Dental trials**  
Multiple teeth per patient (*Cluster = patient*)
- **Ophthalmology trials**  
Pairs of eyes per patient (*Cluster = patient*)
- **Repeated measures studies**  
Recurrent events per patient, such as illnesses or adverse events (*Cluster = patient*)

#### ***Example***

Repeated ordinal responses of pain relief over an 8-hour period in a randomized clinical trial of acute pain relief comparing placebo with 2 analgesics (Gansky, *et al*, 1994, *Journal of Biopharmaceutical Statistics*)



## Clustered Data Applications

---

### Pharmaceutical Research

#### Clinical Trials (continued)

- **Cross-Over Studies**

Patients receive each treatment in sequence  
(*Cluster = patient*)

***Example:***

3-period, 3 treatment cross-over study (Snapinn and Small, 1986, *Biometrics*):

Investigational drug, aspirin, and placebo administered in sequence to headache sufferers; Patients rated each drug on scale of 1-4 according to amount of pain relief.

## Why Did We Bother Developing SUDAAN?

---

### Intra-Cluster Correlation

- Potential for clustermates to respond similarly (genetic and environmental influences)
- Experimental units from the same cluster are not statistically independent
- Usually results in **overdispersion**, or extra-variation in the responses beyond what would be expected under independence
- Negative correlations have the opposite effect *i.e.*, **underdispersion**, or reduction in variance below what would be expected under independence
- Other standard statistical packages (*e.g.*, SAS<sup>®</sup>, SPSS<sup>®</sup>) do not uniformly address the correlated data problem in all analytical procedures

**SAS** mainly uses correlated data methods for discrete (GENMOD) and continuous (MIXED, GENMOD) outcomes in regression models, but not for descriptive data analysis

**SUDAAN** also uses correlated data methods for:

- Means and percentages
- Medians and percentiles
- Odds ratios and relative risks
- Chi-square tests of independence
- Cochran-Mantel-Haenszel tests

- Ratios of random variables

## Impact on Statistical Analysis

---

Failure to account for the cluster effect *usually* leads to:

- Underestimated standard errors for parameters of interest (means, proportions, regression coefficients)
- Test statistics with inflated Type I error rates (false positive tests of treatment effects)

## Implications for Safety and Efficacy

### Safety

- False positives
- Erroneously declaring compounds unsafe

### Efficacy

- False positives
- Erroneously declaring new drugs efficacious
- Reverse effects for cross-over designs:
  - Loss of Power
  - Failure to detect effective treatments

## Multivariate Responses (Clustered Data)

### Notation

$i = \text{cluster}$

$= 1, \dots, n$

$j = \text{observation within the cluster}$

$= 1, \dots, m_i$

### Data

$(y_{ij}, \mathbf{x}_{ij})$ ,  $j = 1, \dots, m_i$

$i = 1, \dots, n$

$N = \sum_i m_i = \text{total sample size}$

### Responses

$\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{im_i})$

### Covariates

$\mathbf{x}_{ij} = (x_{ij1}, x_{ij2}, \dots, x_{ijp})$

This is the clustered data situation covered by SUDAAN

## Cluster-Correlated Data Structures

---

### Example: Teratology Study

$y_{ij}$  = malformation status (present vs. absent)

$x_{ij}$  = exposure level (dosage)

CLUSTER $i$	OBSERVATION $j$				
Patient	Organs (e.g., eyes, teeth, lungs)				
Patient	Recurrent Events				
Litter	Fetuses				
School	Students				
Hospital / Clinic	Patients				
	1	2	3	4	etc.
1	$y_{11}, x_{11}$	$y_{12}, x_{12}$	$y_{13}, x_{13}$	$y_{14}, x_{14}$	
2	$y_{21}, x_{21}$	$y_{22}, x_{22}$			
3	$y_{31}, x_{31}$	$y_{32}, x_{32}$	$y_{33}, x_{33}$		
4	$y_{41}, x_{41}$	$y_{42}, x_{42}$	$y_{43}, x_{43}$	$y_{44}, x_{44}$	
etc.					$y_{ij}, x_{ij}$

- May have unbalanced data (unequal cluster sizes)

## Repeated Measures Data Structures

---

### Example: Longitudinal Study of Pain Relief

$y_{ij}$  = pain relief (none, slight, moderate, complete)

$x_{ij}$  = treatment and/or dosage administered

CLUSTER $i$ Person Animal	OCCASION $j$				
	Time of Measurement (in hours, days, etc.)				
	1	2	3	4	etc.
1	$y_{11}, x_{11}$	$y_{12}, x_{12}$	$y_{13}, x_{13}$	$y_{14}, x_{14}$	
2	$y_{21}, x_{21}$	$y_{22}, x_{22}$	.	.	
3	$y_{31}, x_{31}$	$y_{32}, x_{32}$	$y_{33}, x_{33}$	.	
4	$y_{41}, x_{41}$	$y_{42}, x_{42}$	$y_{43}, x_{43}$	$y_{44}, x_{44}$	
etc.					$y_{ij}, x_{ij}$

- May have unbalanced data (missing values)
- Evaluate the effects of Treatment, Time, and any other (possibly time-dependent) covariates.

## Examples: Data Set 1

---

### Developmental Toxicity Study (EPA, Butler 1988)

- 5 experimental groups
- 25-30 pregnant mice per group, ave 12.4 pups / litter
- Exposure to DEHP (Diethylhexyl phthalate, a plasticizing agent) daily during gestation

0 ppm (Control group)

250 ppm

500 ppm

1000 ppm

1500 ppm

- Outcomes in Fetuses (within litters)

Fetal Death (yes/no)

Malformations (yes/no)

Fetal Body Weight

- Focus here on fetal death: Clustered Binary Data

$$y_{ij} = \begin{cases} 0, & \text{if fetus alive} \\ 1, & \text{if fetus dead} \end{cases}$$

**Question:** Does the incidence of fetal death (and/or malformation) increase with dosage?

## Examples: Data Set 1

---

### Descriptive Statistics for Fetal Death in the DEHP Experiment

Dose Group	Number Litters	Number Fetuses	Percent Dead	Standard Error		Variance Ratio
				Cluster	Independent	
Control	30	396	16.67	4.17	1.87	4.97
250 ppm	26	320	10.00	1.55	1.68	0.85
500 ppm	26	319	13.17	1.87	1.89	0.98
1000 ppm	24	276	50.36	7.57	3.01	6.32
1500 ppm	25	308	83.77	4.73	2.10	5.07



**SUDAAN**



**Standard  
Packages:  
Too Small**

Cluster: SUDAAN  
Independent: Standard Statistical Packages  
(*e.g.*, SAS PROC FREQ)

Source: Bieler and Williams (1995), *Biometrics* **51**, 764-776.



## Examples: Data Set 2

---

### Cross-Over Clinical Trial:

Repeated Exercise Times to Angina Pectoris (Crouchley and Pickles, *Biometrics*, 1993)

- Double-blind randomized cross-over design  
(not enough info to test carry-over effects)
- 21 male patients (clusters) with coronary heart disease
- Tested 4 times on each of two consecutive days  
(Cluster size = 8)

Just before drug administration

1 hr post

3 hrs post

5 hrs post

- One day:       Active treatment (isosorbide dinitrate)  
Other day:     Placebo
- Outcome at each of 8 time points:

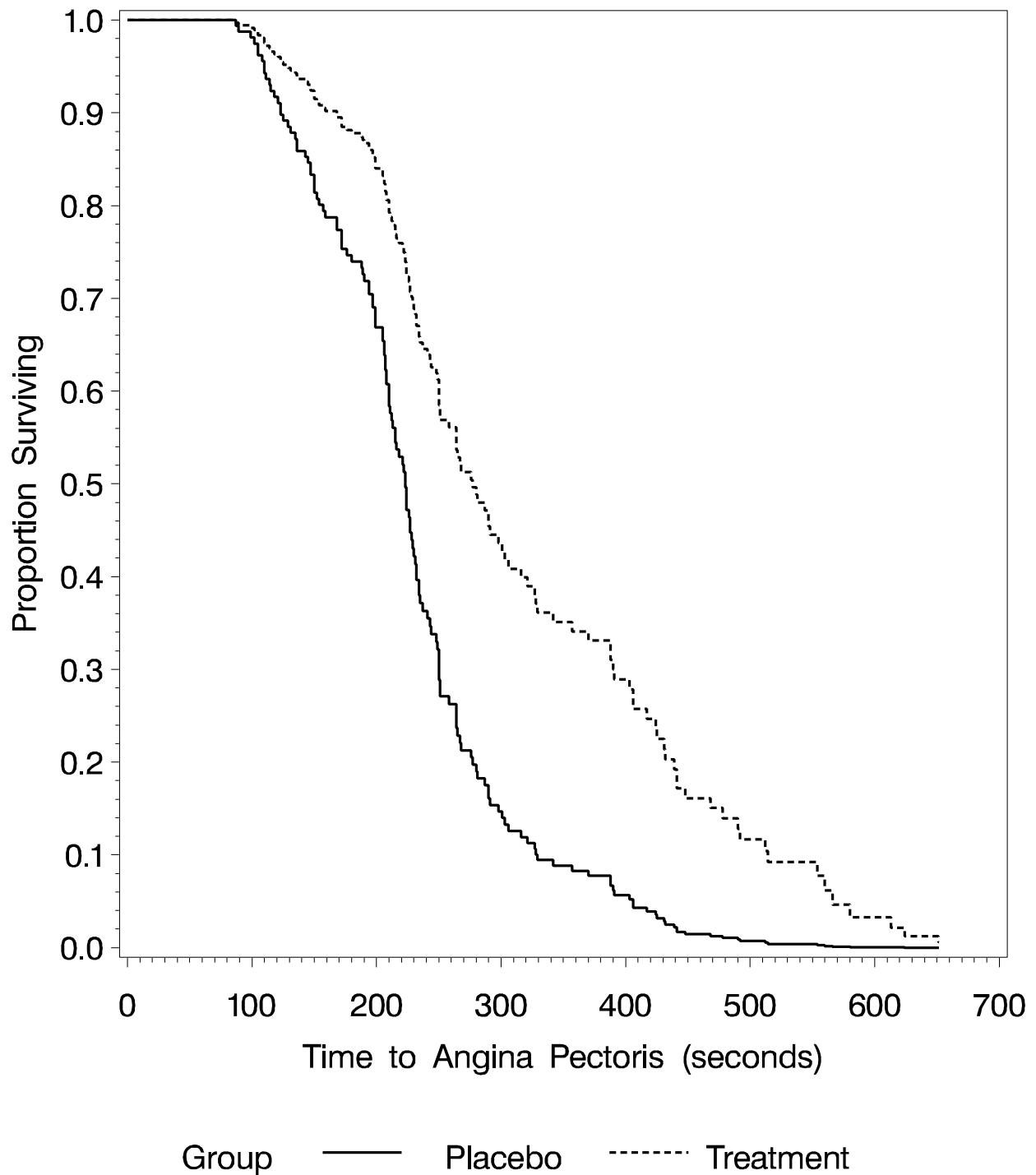
$y_{ij}$  = Exercise time to angina pectoris (in seconds)

**Question:**     **Does treatment delay the time to angina pectoris, after adjusting for time since drug administration and previous conditions?**

# Exercise Time To Angina Pectoris

## Predicted Survival Functions

Adjusted for Time Since Drug Administration and Previous Conditions




## Examples: Data Set 2

---

### Proportional Hazards Model Results

Estimated Regression Coefficient: Treatment vs. Placebo	Estimated Hazards Ratio	Standard Error of Beta		Variance Ratio
		Cluster	Independent	
-0.8395	0.43	<b>0.1474</b>	<b>0.1724</b>	<b>0.73</b> (27% reduction)

  
**SUDAAN**      **Standard Packages:**  
 Too Large

- True variance (via SUDAAN) **smaller** than under independence (e.g., via SAS)
- May fail to detect a treatment effect

## Examples: Data Set 3

---

### Cross-Over Clinical Trial (Ezzett and Whitehead, 1991)

- Two-treatment, 2-period cross-over design
- Comparing two Inhaler Devices in Asthma patients:  
New inhaler vs. a standard (delivering salbutamol).
- Patients randomized to either:
  - Group 1: Device A for 1 week, B the next
  - Group 2: Device B the first week, A the next

No wash-out period

- Outcome of interest: Clarity of leaflet instructions
- Ordinal Scale:

$$y_{ij} = \begin{cases} 1, & \text{Easy} \\ 2, & \text{Only clear after rereading} \\ 3, & \text{Not very clear} \\ 4, & \text{Confusing} \end{cases}$$

**Question:** Is there a difference between the 2 inhaler devices with respect to clarity of leaflet instructions?

## Examples: Data Set 3

---

### Frequency Distribution of Leaflet Clarity in the Cross-Over Clinical Trial

Inhaler Device	Total	Clarity of Leaflet Instructions			
		Easy	Requires Rereading	Not Clear	Confusing
A	286	211	71	2	2
B	286	147	118	15	6

Note: There are 286 patients (clusters) in the study


Source: Ezzet and Whitehead (1991), *Statistics in Medicine* **10**, 901-907.

## Examples: Data Set 3

---

### Proportional Odds Model Results

Estimated Regression Coefficient: Inhaler A vs. B	Estimated Odds Ratio	Standard Error of Beta		Variance Ratio
		Cluster	Independent	
1.0137	2.76	<b>0.1566</b>	<b>0.1733</b>	<b>0.78</b> <b>(22% reduction)</b>

  
**SUDAAN**      **Standard Packages:**  
                          **Too Large**

- True variance (via SUDAAN) **smaller** than under independence (e.g., via SAS)
- May fail to detect a treatment effect

## Design Effects

---

- Measures the impact of clustering on variance estimation and statistical inference for a sample statistic (mean, proportion, regression coefficient)
- Proportional to intraclass correlation and cluster size

### Means, Proportions:

$$\text{DEFF} = \frac{\text{Predicted}}{\text{Observed}} = 1 + \rho_y (m - 1) = \frac{V_{CLUSTER}}{V_{INDEPENDENCE}}$$

$\rho_y$	=	response intraclass correlation
$m$	=	cluster size
$V_{cluster}$	=	variance of the statistic under cluster sampling
$V_{independence}$	=	variance of the statistic under independence

Kish and Frankel (1974)

DEFF	=	1:	No inflation in variance
	>	1:	Variance inflation (over-dispersion)
			$\rho_y > 0$ , almost always

## Design Effects

---

### Regression Coefficients: Continuous and Binary Responses\*

$$Ex: E(y) = X\beta$$

$$DEFF = 1 + \rho_y \rho_x (m - 1)$$

$\rho_x$  = covariate intracluster correlation

$$= \begin{cases} \mathbf{1}, & \text{cluster-level covariates} \\ < \mathbf{0}, & \text{within-cluster covariates (same patterns)} \\ > \mathbf{0}, & \text{within-cluster covariates (patterns differ)} \end{cases}$$

Assumes exchangeable correlation structure

Neuhaus and Segal (1993)

Scott and Holt (1982)

Zeger (1988)

\* Proportional hazards survival models:  
This pattern has been observed in simulation studies, but the exact formula has not been developed



## Analysis Implications



---

DEFF &gt; 1

 $\rho_y > 0$  $\rho_x > 0$ 

DEFF &lt; 1

 $\rho_y > 0$  $\rho_x < 0$ 

	Type I Error	Power
Standard	Inflated	Reduced
SUDAAN	Correct	Increased
		
	Treatment: <b>Cluster-level</b> covariate	Treatment: <b>Within-cluster</b> covariate
	e.g., Teratology	e.g., Cross-Over Trial

## How Much Correlation in Typical Studies?

---

Study	Ave. Cluster Size, $\bar{m}$	$\rho_y$	$\rho_x$	Design Effect
Ex 1: Teratology	13	.26	1	4.0
Ex 2: Angina Exercise Times Cross-Over Study	8	.65	$< 0^*$	0.73
Ex. 3: Leaflet Clarity Cross-Over Study	2	Matrix of positive and negative values	$-1^*$	0.78

$$^* \rho_x = \frac{-1}{\bar{m} - 1} \quad \text{via Kappa Statistic}$$

$$DEFF = 1 + \rho_y \rho_x (\bar{m} - 1)$$

## Design Effects

---

### Relationship to Effective Sample Size

Suppose sample consists of 20 clusters, with 10 observations per cluster:

$n = 20$  clusters

$m = 10$  observations / cluster

$N = 200$  observations in sample ( $n*m$ )

$\rho_y =$ Intracluster Correlation	Design Effect	Effective Sample Size
0 (independence)	1	200 (= $N$ )
1 (perfect positive correlation)	10 (= $m$ )	20 (= $n$ )

$$\text{Design Effect (DEFF)} = 1 + \rho_y (m - 1) = V_{\text{Cluster}} / V_{\text{Indep}}$$

$$\text{Effective Sample Size (EFF)} = N / \text{DEFF}$$

= amount of independent information available for computing variance of sample statistics

Usually:  $0 < \rho_y < 1$  (closer to 0)

$1 < \text{DEFF} < 10$  (closer to 1)

$200 > \text{EFF} > 20$  (closer to 200, number of obs)

## Why SUDAAN?

### A Proportional Hazards Model Simulation Experiment

---

#### Survival Function

$$S(t) = \text{Prob}(T > t)$$

#### Hazard Function

$$h(t) = \frac{f(t)}{S(t)}$$

#### Proportional Hazards Model

$$h(t|\mathbf{x}) = h_o(t) \exp(\mathbf{x}'\boldsymbol{\beta})$$

$$S(t|\mathbf{x}) = [S_o(t)]^{\exp(\mathbf{x}'\boldsymbol{\beta})}$$

#### Variance Estimation

Linearization, combined with a between-cluster variance estimator:

Binder (1983, *International Stat. Review*)

Binder (1992, *Biometrika*)

Lin and Wei (1989, *JASA*)

Shah, et al (1996). *SUDAAN User's Manual, Rel 7.0*

Lee, Wei, Amato (1992)

Lin (1994, *Statistics in Medicine*)

Williams and Bieler (1993, *ASA Biopharm Proceedings*)

## Why SUDAAN?

### A Proportional Hazards Model Simulation Experiment

---

#### Simulation of a Repeated Measures Study

- Based on the design of a Neurotoxicology Study measuring the effect of exposure on learning and memory in laboratory rats
- 4 dose groups
- 25 animals (clusters) per group
- 2 cluster sizes
  - 5 measurements per cluster
  - 20 measurements per cluster
- 2 dose-response relationships

$$h(t|d) = h_0(t) \exp(\beta d) \quad d=0,1,2,3$$

$\beta = 0$                       No dose response

$\beta = .0951$                 33% increased hazard in high dose vs. control

- Also analogous to recurrent events observed on same subject

## Proportional Hazards Simulation Results: Empirical Probabilities of Rejection

$$H_0: \beta = 0$$

Intracluster Correlation	Type I Error		Power (Cluster)
	Independence	Cluster	
<b>Cluster Size = 20, Total Sample Size = 2,000</b>			
All .00	4.4	4.5	96.8
All .05	12.3	5.0	79.6
All .10	20.1	4.8	64.9
All .20	29.3	5.0	42.0
<b>Cluster Size = 5, Total Sample Size = 500</b>			
All .00	5.3	5.9	56.2
All .05	7.0	5.7	46.3
All .10	9.0	5.6	40.5
All .20	11.3	5.1	33.1

Independence: Standard Statistical Packages (e.g., SAS)  
Cluster: SUDAAN

Type I Error: Two-Tail Test  
Power: Upper-tail test with true  $\beta = 0.0951$

## **Summary of the Proportional Hazards Simulation**

- Assuming independence yields inflated Type I error rates
  - Leads to false positive results
  - Increases with cluster size
  - Increases with intracluster correlation
  
- SUDAAN methods maintain 5% Type I error for all cluster sizes and levels of intracluster correlation
  
- SUDAAN gets the variance right
  - Average variance estimates for SUDAAN matched the empirical values

## Why SUDAAN?

### A Logistic Regression Simulation Experiment

---

**Context:** Teratology Screening Study

**Source:** Bieler and Williams (*Biometrics*, 1995)

#### Two Analysis Methods:

- SUDAAN: GEE-independent  
(ordinary logistic regression with a variance correction)
- Independence (as in most standard software packages)

#### Test

$$H_0: \beta_1 = 0 \quad \text{vs.} \quad H_1: \beta_1 > 0$$

#### Linear Logistic Model

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \text{DOSE}$$

where  $p = \Pr(Y_{ij} = 1 \mid \text{DOSE}, \beta)$  = probability of malformation in the  $j$ -th fetus

#### Evaluate

- Empirical Type I Error (nominal  $\alpha = 0.05$ )
- Power (doubling of malformation rate in high dose)



## **Why SUDAAN?**

### **A Logistic Regression Simulation Experiment**

---

#### **Design**

- Kupper, et al (1986)
- Carr (1993)
  
- 4 dose groups:  $\ln(\text{dosage}) = 1, 2, 3, 4$  (C, L, M, H)
- 25 litters / group
- Mean litter size = 11.7 pups

#### **Intralitter Correlations (Beta-binomial model)**

- 3 homogeneous

#### **Model**

- Linear logistic model
- Null Case (background rate = 5%)
- Alternative Case (10% in high dose)

#### **Tests**

- 1,400 replications
- Wald statistic:  $Z = \beta / \text{se}(\beta)$
- Nominal  $\alpha = .05$

## Teratology Simulation Results

### Observed Upper-Tail Probabilities of Rejection for the Null Hypothesis that the Logistic Slope Parameter = 0

Nominal Alpha = 5%  
Background Rate = 5%

Operating Characteristics	High Dose Rate	Intralitter Correlations	Test Statistic	
			Standard $Z_{indep}$	SUDAAN $Z_{cluster}$
<b>TYPE I ERROR</b>	5%	0 0 0 0	4.0	4.9
	5%	.1 .1 .1 .1	13.1	4.5
	5%	.2 .2 .2 .2	17.8	6.9
<b>POWER</b>	10%	0 0 0 0	81.6	84.9
	10%	.1 .1 .1 .1	69.5	52.0
	10%	.2 .2 .2 .2	68.9	41.1

$Z_{INDEP}$  is the standard Wald test statistic estimated from a linear logistic model assuming independence and a binomial likelihood

$Z_{CLUSTER}$  is a Wald test statistic estimated from a linear logistic model under a binomial likelihood, with between-cluster variance estimates based on Taylor linearization (or: GEE with independent working correlations and a robust variance estimator).

## **Why SUDAAN?**

### **A Logistic Regression Simulation Experiment**

---

#### **Results:**

- Assuming independence yields inflated Type I error rates
  - Leads to false positive results
  - Increases with intracluster correlation
  
- SUDAAN maintained nominal Type I error rates (5%)
  
- SUDAAN gets the variance right:
  - Average variance estimates for SUDAAN matched the empirical values

## Analysis of Clustered Data

---

### Basic Concept

- 1) Use consistent estimators of the parameters  
  
*e.g.*, Means, Proportions, Percentages, Odds Ratios, Regression Coefficients
  - *Without* imposing strict distributional assumptions about the response of interest
  - Intracluster correlation treated as a nuisance parameter
  
- 2) Robust variance estimators ensure consistent variance estimates and valid inferences

### Two Variance Estimation Methods in SUDAAN:

- Taylor linearization
- Jackknife resampling (new in Release 7.5)

## Assumptions: Independence Vs. Clustered Data

### Independence

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} \quad V(\mathbf{Y}) = \sigma^2 \mathbf{I}_N = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

Observations independent, constant variance

### Clustered Data (SUDAAN):

$$\mathbf{Y} = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1m_1} \\ \vdots \\ y_{n1} \\ \vdots \\ y_{nm_n} \end{bmatrix} \quad \begin{array}{l} n \text{ clusters of } m_i \text{ observations } \quad (N = \sum_{i=1}^n m_i) \\ \text{Unequal observations per cluster} = m_i \\ \text{Example: } n \text{ litters with } m_i \text{ pups per litter} \end{array}$$

## Assumptions: Independence Vs. Clustered Data

### Clustered Data (SUDAAN):

$$V(Y) = \begin{bmatrix} V_1 & 0 & 0 & \dots & 0 \\ 0 & V_2 & 0 & \dots & 0 \\ 0 & 0 & V_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & V_n \end{bmatrix}$$

Cluster-Correlated Data  
Block-Diagonal by Cluster  
 $V_i$  is an  $m_i \times m_i$  matrix

$$V_i = \begin{bmatrix} \sigma_{(i)1}^2 & \sigma_{(i)12} & \sigma_{(i)13} & \dots & \sigma_{(i)1m} \\ \sigma_{(i)21} & \sigma_{(i)2}^2 & \sigma_{(i)23} & \dots & \sigma_{(i)2m} \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \sigma_{(i)m1} & \sigma_{(i)m2} & \sigma_{(i)m3} & \dots & \sigma_{(i)m}^2 \end{bmatrix}$$

- $V_i$  is an  $m_i \times m_i$  variance covariance matrix of observations in the  $i$ -th cluster
- No assumptions on structure of  $V_i$
- Observations independent between clusters, completely arbitrary correlation structure within clusters

## Independence Vs. Clustered Data: Estimation of a Mean

---

### Independence

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

$$\hat{V}(\bar{y}) = \frac{\hat{\sigma}^2}{N} = \frac{1}{N} \left[ \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N-1} \right]$$

$$= \frac{N}{N-1} \sum_{i=1}^N \left( \frac{y_i - \bar{y}}{N} \right)^2$$

$N$  = number of *observations*

Between-*Unit* Variance Estimator

### Clustered Data (How is SUDAAN different?)

$$Z_{ij} = \frac{y_{ij} - \bar{y}}{N} \quad \text{Linearized Value}$$

$$Z_i = \sum_{j=1}^{m_i} Z_{ij} \quad \text{Cluster Totals}$$

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i \quad \text{Mean of Cluster Totals}$$

$$\hat{V}(\bar{y}) = n S^2 = \frac{n}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2$$

$n$  = number of *clusters*

*Robust* or Between-*Cluster* Variance Estimator

## Why Just the Variance of Cluster Totals?

---

### Simple Example (Teratology)

Let  $y_{ij} = \mu + \gamma_i + \epsilon_{ij}$

$$i = 1, \dots, n \quad (n \text{ litters})$$

$$j = 1, \dots, m \quad (m \text{ pups per litter})$$

$$\gamma_i \sim N(0, \sigma_\gamma^2) \quad \text{litter effect (cluster)}$$

$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \quad \text{pup effect}$$

$\gamma, \epsilon$  independent

Then,

$$\bar{y} = \frac{1}{n} \frac{1}{m} \sum_{i=1}^n \sum_{j=1}^m y_{ij}$$

$$\text{Var}(\bar{y}) = \frac{\sigma_\gamma^2}{n} + \frac{\sigma_\epsilon^2}{nm} \quad \text{True Variance}$$

$$E(n S^2) = \frac{\sigma_\gamma^2}{n} + \frac{\sigma_\epsilon^2}{nm} \quad \begin{array}{l} \text{Expected Value of} \\ \text{Between-Cluster Variance Estimator} \end{array}$$

This says that the expected value of our between-cluster variance estimator is asymptotically equal to the true variance of the estimated parameter, ***no matter how many stages of clustering are present.***

***It also holds true for unequal sample sizes and more general distributional assumptions.***



## Taylor Linearization Approach: Descriptive Statistics

---

*e.g.*, Means, Proportions, Odds ratios

### Two-Step Procedure for Variance Estimation:

- 1) **Apply Taylor series linearization to approximate functions of linear statistics (*e.g.*, ratios of random variables)**

*Example: Teratology*

Proportion of malformed fetuses in a teratology experiment

$$\hat{p} = \frac{\sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij}}{\sum_{i=1}^n m_i} = \frac{\text{Number Malformed Fetuses}}{\text{Total Number of Fetuses}}$$

Find linear approximation to this nonlinear statistic (Kendall and Stuart, 1973);

Between-cluster variance formulas available for *linear* statistics.

Woodruff (1971):

- Equivalent computational procedure using Taylor series linearized values
- Each observational unit gets a linearized value ( $Z_{ij}$ ) for a particular statistic.

- 2) **Apply between-cluster variance estimator to the linearized values  $Z_{ij}$**

## Taylor Series Expansion

---

$X, Y$  are two linear statistics

Then,

$$\begin{aligned} F(X, Y) &= F(\mu_X, \mu_Y) \\ &+ \partial F_X|_{\mu_X} (X - \mu_X) \\ &+ \partial F_Y|_{\mu_Y} (Y - \mu_Y) \\ &+ \textit{higher order terms} \end{aligned}$$

$$\begin{aligned} \text{Var}[F(X, Y)] &\doteq E[F(X, Y) - F(\mu_X, \mu_Y)]^2 \\ &= E\left[\partial F_X|_{\mu_X} (X - \mu_X) + \partial F_Y|_{\mu_Y} (Y - \mu_Y)\right]^2 \end{aligned}$$

## Other Examples of Taylor Series Linearization

### Odds Ratio

	Disease	Not Diseased
Exposed	$a$	$b$
Not Exposed	$c$	$d$

$$OR = \frac{a \cdot d}{b \cdot c}$$

### *Taylor Series Variance*

$$\left( \frac{ad}{bc} \right)^2 \left[ \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \right]$$

### Kaplan-Meier Survival Function

$$S(t_j) = \prod_{i=1}^j \left( 1 - \frac{d_i}{r_i} \right)$$

### *Greenwood's Variance Formula*

$$S(t_j)^2 \sum_{i=1}^j \frac{d_i}{r_i(r_i - d_i)}$$

*Where  $i$  denotes death times*

### Categorical Data Analysis

Grizzle-Starmer-Koch  
variance formula for  
weighted least squares

## Assumptions and Validity for Taylor Approach

---

- Assume:  $n$  clusters selected independently from a hypothetical infinite population of clusters
- No strict distributional assumptions for the response of interest ("model-free")
- SUDAAN variance estimator yields consistent estimates of the variance as the number of clusters tends to infinity

Method is valid for any underlying intracluster correlation structure, as long as:

- 1) Clusters are statistically independent
- 2) Cluster totals  $Z_i$  can be unbiasedly estimated

Therefore, specification of an explicit correlation structure is unnecessary.

- Also valid in presence of additional sources of correlation within each clustermate (*e.g.*, multiple levels of nesting)

## Independence Vs. Clustered Data: Fitting Linear Regression Models

---

### Standard Situation: Linear Regression

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} \quad \begin{array}{l} E(\mathbf{Y}) = \mathbf{X} \boldsymbol{\beta} \\ V(\mathbf{Y}) = \sigma^2 \mathbf{I}_N \\ \text{Independent obs, constant variance} \end{array}$$

### Standard Solution to Normal Equations:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

$$\text{Var}(\mathbf{b}) = \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} \quad \hat{\sigma}^2 = \text{Mean Square Error}$$

This variance formula only holds when  $V(\mathbf{Y}) = \sigma^2 \mathbf{I}_N$

## Independence Vs. Clustered Data: Fitting Linear Regression Models

---

**How is SUDAAN different?**

$$V(\mathbf{Y}) = V_{\mathbf{Y}} = \begin{bmatrix} V_1 & 0 & 0 & \dots & 0 \\ 0 & V_2 & 0 & \dots & 0 \\ 0 & 0 & V_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & V_n \end{bmatrix}$$

Cluster-Correlated Data  
Block-Diagonal by Cluster  
 $V_i$  is an  $m_i \times m_i$  matrix

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

Form Taylor series linearized values,  $Z_{ij}$ , for the linear regression coefficients above. Then use between-cluster variance formula to estimate:

$$\text{Var}(\mathbf{b}) = V_{\mathbf{b}} \quad \text{Estimates each element separately}$$

**KEY POINT:**

$$V_{\mathbf{b}} \neq \hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} \quad \text{due to cluster-correlated data}$$

## Independence Vs. Clustered Data Fitting Linear Regression Models

---

### Null Hypothesis:

$$H_0: C\beta = \mathbf{0}$$

$C$  is a contrast matrix of rank  $r$

### General Form for Test Statistic:

$$Q = (Cb)' [C \text{Var}(b) C']^{-1} (Cb)$$

### Standard Situation

$$Q = (Cb)' [\hat{\sigma}^2 C (X'X)^{-1} C']^{-1} (Cb)$$

$$= \frac{MS_{H_0}}{MS_{error}} \sim r F_{r, N-r}$$

Standard computing formula used by most software packages

### SUDAAN Test Statistic:

$$Q = (Cb)' [CV_b C']^{-1} (Cb)$$

Does not reduce to any simple computing formula

## Jackknife Variance Estimation

---

### Coming in Release 7.5:

### ***Resampling Methods for Correlated Data***

Quenouille (1956): Reducing bias in estimation

Tukey (1958): Approximate confidence intervals

### **Start With Given Point Estimator:**

Descriptive statistics (*e.g.*, means, proportions)

Regression parameter vectors

Use consistent estimators of location parameters

Naively treat the correlated responses as independent

### **Covariance Estimates for Descriptive Statistics**

Binomial proportions (Gladden, 1979 *JASA*):

Proportion of fetuses that are malformed in a teratology study

$$\hat{p} = \frac{\sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij}}{\sum_{i=1}^n m_i}$$



## Jackknife Variance Estimation

---

An estimate based on all clusters **except the  $k$ -th** is as follows:

$$\hat{P}_{(k)} = \frac{\sum_{i \neq k}^n \sum_{j=1}^{m_i} y_{ij}}{\sum_{i \neq k}^n m_i}$$

**Jackknife Variance Estimate for  $\hat{p}$ :**

$$\hat{\sigma}_{JK}^2 = \frac{n-1}{n} \sum_{k=1}^n [\hat{P}_{(k)} - \hat{P}_{(.)}]^2$$

where  $\hat{P}_{(.)}$  is the average of the Jackknife estimates:

$$\hat{P}_{(.)} = \frac{\sum_{k=1}^n \hat{P}_{(k)}}{n} .$$

Assuming

$$\begin{aligned} E(y_i | m_i) &= m_i p \\ V(y_i | m_i) &= h(m_i) , \end{aligned}$$

then

$$\frac{\hat{p} - p}{\hat{\sigma}_{JK}} \rightarrow Z \sim N(0,1)$$

## Jackknife Variance Estimation

---

### Covariance of Regression Parameters

- Logistic regression parameters obtained under a binomial likelihood (Carr and Portier, 1993 *Biometrics*)
- Cox model parameters obtained under a partial likelihood (Lipsitz and Parzen, 1996 *Biometrics*; Lipsitz, Dear, and Zhao, 1994 *Biometrics*)

### Start With Given Point Estimator $\hat{\beta}$ :

Estimated parameter vector obtained by naively assuming the observations within a cluster are independent

Solution to any score estimating equation of the form

$$\mu(\hat{\beta}) = \sum_{i=1}^n \mu_i(\hat{\beta}) = 0$$

where  $\mu_i(\hat{\beta})$  is the contribution to the “score” vector from the  $i$ -th cluster.

### Example

Logistic score equations under binomial likelihood

$$U(\beta) = \frac{\partial \text{Log } L(\beta)}{\partial \beta} = \sum_i \sum_j x'_{ij} y_{ij} - \sum_i \sum_j x'_{ij} p_{ij}(\beta)$$

Solve via iteration:  $U(\beta) = \mathbf{0}$

## Jackknife Variance Estimation

---

### Regression Parameters (continued)

As long as the model for the marginal mean is correctly specified, the MLE  $\hat{\beta}$  is asymptotically consistent and normally distributed

### Jackknife Variance Estimator For $\hat{\beta}$

$$Var_{JK}(\hat{\beta}) = \left( \frac{n-p}{n} \right) \sum_{i=1}^n (\hat{\beta}_{-i} - \hat{\beta})(\hat{\beta}_{-i} - \hat{\beta})'$$

where

- $p$  = number of parameters in the model, and
- $\hat{\beta}_{-i}$  = estimate of  $\beta$  obtained by deleting the  $m_i$  observations in cluster  $i$  and solving the estimating equations via the Newton-Raphson algorithm.

Clusters are removed sequentially and with-replacement

JK variance estimator is consistent for estimating the asymptotic variance of  $\hat{\beta}$

## Jackknife Variance Estimation

---

### Regression Parameters (continued)

### Simulation in Small Sample Situations

Evaluating treatment effect in logistic regression models (Carr and Portier, *Biometrics*, 1993)

Jackknife Method:

- Controlled Type I error
- Estimated location parameters without bias
- Estimated variance of parameter estimates without bias
- Similar to Zeger/Liang GEE in terms of performance

## **Assumptions and Validity for Jackknife**

---

- Assume:  $n$  clusters selected independently from a hypothetical infinite population of clusters
- No strict distributional assumptions for the response of interest
- Jackknife variance estimator yields consistent estimates of the variance as the number of clusters tends to infinity
- Method is valid for any underlying intracluster correlation structure, as long as clusters are statistically independent
- Also valid in presence of additional sources of correlation within each clustermate (*e.g.*, multiple levels of nesting)

## Efficient Parameter Estimation

---

### Efficiently Weight the Data to Estimate Regression Coefficients ( $\beta$ )

**GEE Approach** (Zeger and Liang, 1986):

- 1) Assume a Covariance Structure for  $V_i$ 
  - Mean / Variance Relationship
  - Correlation Model
- 2) Estimate Covariance Parameters
- 3) Weight Data Inversely Proportional to  $V_i$  to Estimate  $\beta$

$y_i = (y_{i1}, \dots, y_{im_i})$       Vector of responses

$\mu_i = E(y_i) = \mu_i(\beta)$       Vector of marginal means  
 $= (\mu_{i1}, \dots, \mu_{im_i})$

$V_i(\alpha) = Cov(y_i; \mu_i, \alpha)$       Working Covariance matrix

$$U(\beta) = \sum_{i=1}^n \frac{\partial \mu_i'}{\partial \beta} V_i(\alpha)^{-1} (y_i - \mu_i) = \mathbf{0}$$

## Covariance Structure for $V_i$

---

$$V_i(\alpha) = A_i^{1/2} R_i(\alpha) A_i^{1/2} \cdot \phi \quad V \text{ is Block diagonal}$$

$A_i$  = diagonal matrix with diagonal elements  
 $g(\mu_{i1}), \dots, g(\mu_{im_i})$

$$= \begin{bmatrix} g(\mu_{i1}) & 0 & 0 & 0 \\ 0 & g(\mu_{i2}) & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & g(\mu_{im_i}) \end{bmatrix}$$

### Relationship Between Variance of $y_{ij}$ and its mean

$$\text{Var}(y_{ij}) = g(\mu_{ij}) \cdot \phi$$

$g$  is a variance function,  $\phi$  is an unknown scale parameter

### Binary Responses

Marginal distribution of  $y_{ij}$  is Bernoulli

$$\text{Var}(y_{ij}) = \mu_{ij}(1 - \mu_{ij}) \text{ and } \phi = 1.$$

$R_i(\alpha)$  is the “Working” Correlation Matrix for  $y_i$

$$\alpha_{jk} = \text{corr}(y_{ij}, y_{ik})$$

## Choices for Working Correlation Matrices

---

### 1) Independent Working Correlation Matrix (Identity matrix)

$$\text{Form: } R_i(\alpha) = I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Estimating equations reduce to familiar forms:
  - Normal equations for linear regression
  - Score equations for logistic regression
- Leads to standard regression coefficient estimates
- Consistent and asymptotically normal, even under cluster sampling
- This approach is offered in SUDAAN, and it is perfectly valid for estimating the regression parameters.



## Choices for Working Correlation Matrices

---

### 2) Exchangeable (equal pairwise correlations)

$$\text{Form: } \mathbf{R}_i(\boldsymbol{\alpha}) = \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

- SUDAAN offers this form as well
- Can improve *efficiency* of parameter estimates over the independence working assumption.

## Robust Variance Estimate for GEE

---

$$\text{Var}(\hat{\boldsymbol{\beta}}) = M_0^{-1} M_1 M_0^{-1}$$

where

$$M_0 = \sum_{i=1}^n \frac{\partial \boldsymbol{\mu}'_i}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1} \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$$

$$M_1 = \sum_{i=1}^n \frac{\partial \boldsymbol{\mu}'_i}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1} \text{Var}(\mathbf{y}_i) \mathbf{V}_i^{-1} \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$$

- $M_0^{-1}$  (outside term) is called the *naive* or *model-based* variance (inverse of information matrix, appropriate when covariance structure correctly specified)
- $M_1$  (middle term) serves as a *variance correction* when the covariance model is misspecified
- Robust variance is consistent even when  $\text{var}(y_{ij}) \neq g(\mu_{ij}) \cdot \phi$  or  $\mathbf{R}_i(\boldsymbol{\alpha})$  is not the true correlation matrix of  $\mathbf{Y}_i$
- $\text{Var}(\mathbf{y}_i)$  empirically estimated by  $(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_i - \hat{\boldsymbol{\mu}}_i)'$
- SUDAAN offers the *robust* and soon (Release 7.5) the *model-based* variance estimates

## Robust Variance Estimate for GEE

---

- Also referred to as Sandwich Estimator or Variance Correction
- Properly accounts for intracluster correlation
- Yields *consistent variance estimates*, even if correlation structure is misspecified (*e.g.*, by specifying “working” independence when the correlations are in fact exchangeable)

Huber (1967)

Royall (1986)

Binder (1983, 1992)

## What Does SUDAAN Model?

---

### Marginal Models (Population-Averaged)

- "Marginal mean" of the multivariate outcomes as a function of the covariates:

$$E(y_{ij} | \mathbf{x}_{ij})$$

- Focus on how X causes Y, while acknowledging the dependence within clusters (as opposed to how one Y causes another)
- Describes relationship between covariates and response *across* clusters
- Intracluster correlation treated as nuisance parameter

### References:

Zeger and Liang (1986)  
Liang and Zeger (1986)  
Zeger, Liang, and Albert (1988)  
Binder (1983, 1992)

## **SUDAAN Software Package**

---

### **Software for Statistical Analysis of Correlated Data**

- Single program, written in the *C* language, consisting of a family of statistical procedures
  
- As easy to use as SAS!
  - Uses a SAS-like interface
  - Accepts SAS data sets or ASCII files as input
  
- SPSS Users: Release 7.5 reads SPSS files
  
- Two Modes of Operation:
  - 1) SAS-Callable  
(VAX/VMS, IBM/MVS, SUN/Solaris, and soon for Win 95)
  
  - 2) Stand-Alone  
(many platforms, including Windows)
  
- 100% compatible across platforms

# SUDAAN Procedures

---

## DESCRIPTIVE PROCEDURES

### CROSSTAB

Computes frequencies, percentage distributions, odds ratios, relative risks, and their standard errors (or confidence intervals) for user-specified cross-tabulations, as well as chi-square tests of independence and the Cochran-Mantel-Haenszel chi-square test for stratified two-way tables.

### DESCRIPT

Computes estimates of means, totals, proportions, percentages, geometric means, quantiles, and their standard errors; also computes standardized estimates and tests of single degree-of-freedom contrasts among levels of a categorical variable.

### RATIO

Computes estimates and standard errors of generalized ratios of the form  $\Sigma y / \Sigma x$ , where  $x$  and  $y$  are observed variables; also computes standardized estimates and tests single-degree-of-freedom contrasts among levels of a categorical variable.

## REGRESSION PROCEDURES

### REGRESS

Fits linear regression models and performs hypothesis tests concerning the model parameters. Uses GEE to efficiently estimate regression parameters, with robust and model-based variance estimation.

### LOGISTIC

Fits logistic regression models to binary data and computes hypothesis tests for model parameters; also estimates odds ratios and their 95% confidence intervals for each model parameter.

### MULTILOG

Fits multinomial logistic regression models to ordinal and nominal categorical data and computes hypothesis tests for model parameters; estimates odds ratios and their 95% confidence intervals for each model parameter; uses GEE to efficiently estimate regression parameters, with robust and model-based variance estimation.

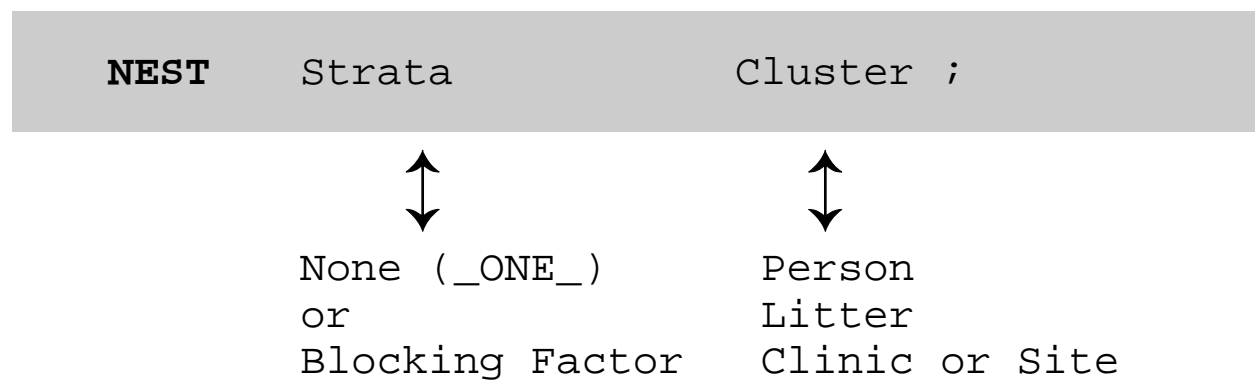
### SURVIVAL

Fits discrete and continuous proportional hazards models to failure time data; also estimates hazard ratios and their 95% confidence intervals for each model parameter.

## Elements of a SUDAAN Procedure

---

<b>PROC</b>	CROSSTAB DESCRIPT MULTILOG	<b>DATA</b> = name	<b>DESIGN</b> = WR JK BRR
	LOGISTIC SURVIVAL		



***For Regression Modelling:***

```
MODEL dependent = independent ;

      Y = DOSE ;
```

***For Descriptive Statistics:***

```
VAR response_variables ;

TABLE categorical effects (e.g.,DOSE) ;
```

## The MULTILOG Procedure

---

### Multinomial Logistic Regression

- **Generalized Logit Models**
  - Nominal Outcomes  
*e.g.*, Type of health plan (A, B, C, D)
  
- **Cumulative Logit Models**
  - Ordinal Outcomes  
*e.g.*, Pain Relief:  
none, mild, moderate, complete relief
  - "Proportional Odds Models"
  
- Binary Logistic is a special case of each
  
- Model-fitting Approach
  - Fits *marginal* or *population-averaged* models
  - Uses GEE to model the intracluster correlations and efficiently estimate regression coefficients



## **SUDAAN Documentation**

---

Shah, Barnwell, and Bieler (1996). ***SUDAAN User's Manual, Release 7.0***, Research Triangle Institute.

## **Enhancements to SUDAAN Release 7.5**

---

### **Resampling Methods for Variance Estimation**

- Jackknife
- Balanced Repeated Replication (BRR)

### **Enhancements of GEE Capabilities**

- Exchangeable correlations in linear regression (as already in logistic and multinomial logistic in Release 7.0)
- Choice of robust or model-based variances in GEE applications

### **Other Regression Enhancements**

- REFLEVEL statement to change the reference level for categorical covariates
- User-friendly contrast statement (the EFFECTS statement) for testing simultaneous regression effects, simple effects in interaction models, and more
- R-square (Cox and Snell, 1989) in logistic regression
- Least Squares Means (LSMEANS) statement in linear regression

### **SAS-Callable Platforms**

- Windows
- SUN/Solaris

**Now reads SPSS files** (in addition to SAS and ASCII)

## References

### Applications of SUDAAN and Related Techniques

- Bieler, G. and Williams, R. (1995). Cluster sampling techniques in quantal response teratology and developmental toxicity studies. *Biometrics* **51**, 764-776.
- Davies, G.M. (1994). Applications of sample survey methodology to repeated measures data structures in dentistry. *Institute of Statistics Mimeo Series* No. 2128T. University of North Carolina: Chapel Hill.
- Donner, A. (1982). An empirical study of cluster randomization. *American Journal of Epidemiology* **11**, 283-286.
- Ennett, S.T., Rosenbaum, D.P., Flewelling, R.L., Bieler, G.S., Ringwalt, C.L., and Bailey, S.L. (1994). Long-term evaluation of drug abuse resistance education. *Addictive Behaviors*, **19**, 113-125.
- Fung, K.Y., Krewski, D., and Scott, A.J. (1994). Tests for trend in developmental toxicity experiments with correlated binary data. *Risk Analysis* **14**, 639-648.
- Gansky, SA, Koch, GG, and Wilson, J. (1994). Statistical evaluation of relationships between analgesic dose and ordered ratings of pain relief over an eight-hour period. *Journal of Biopharmaceutical Statistics* **4**, 233-265.
- Graubard, B.I. and Korn, E.L. (1994). Regression analysis with clustered data. *Statistics in Medicine* **13**, 509-522.
- LaVange, L.M. and Koch, G.G. (1994). Analysis of repeated measures studies with multiple regression methods for sample survey data. Presented at the 1994 Drug Information Association Meetings, and submitted for publication.
- LaVange, L.M., Keyes, L.L., Koch, G.G., and Margolis, P.A. (1994). Application of sample survey methods for modelling ratios to incidence densities. *Statistics in Medicine* **13**, 343-355.
- Norton, E., Bieler, G., Zarkin, G., and Ennett, S. (1996). Analysis of prevention program effectiveness with clustered data using generalized estimating equations. *Journal of Consulting and Clinical Psychology* **64**, 919-926.
- Rao, J. and Colin, D. (1991). Fitting dose-response models and hypothesis testing in teratological studies. In: Statistics in Toxicology, Krewski and Franklin, eds. NY: Gordon and Breach.
- Rao, J. and Scott, A.J. (1992). A simple method for the analysis of clustered binary data. *Biometrics* **48**, 577-585.
- Schmid, JE, Koch, GG, and LaVange, LE (1991). An overview of statistical issues and methods of meta-analysis. *Journal of Biopharmaceutical Statistics*, **1**, 103-120.
- Williams, R.L. (1995). Product-limit survival functions with correlated survival times. *Lifetime Data Analysis* **1**, 171-186.
- Williams, R.L. and Bieler, G.S. (1993). Estimation of proportional hazards models for survival times with nested errors. ASA Proceedings of the Biopharmaceutical Section, 115-120.

## Survey Sampling

- Binder, D. (1992). Fitting Cox's proportional hazards models from survey data. *Biometrika* **79**, 139-147.
- Binder, D. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review* **51**, 279-292.
- Cochran, W.G. (1977). Sampling Techniques. Wiley, New York.
- Folsom, R.E. (1974). National Assessment Approach to Sampling Error Estimation, Sampling Error Monograph. Prepared for the National Assessment of Educational Progress, Denver, CO.
- Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). Sample Survey Methods and Theory, Volume I: Methods and Applications. NY: Wiley.
- Kendall, M.G. and Stuart, A. (1973). The Advanced Theory of Statistics. NY: Hafner Publishing Co.
- Kish, L. (1965). Survey Sampling. NY: Wiley.
- Kish, L. and Frankel, M. (1974). Inference from complex samples (with discussion). *Journal of the Royal Statistical Society, Series B*, **36**, 1-37.
- Koch, G.G., Freeman, D.H. and Freeman, J.L. (1975). Strategies in the multivariate analysis of data from complex surveys. *International Statistical Review* **43**, 59-78.
- LaVange, L.M., Iannacchione, V.G., and Garfinkel, S. (1986). An application of logistic regression methods to survey data: predicting high cost users of medical care. American Statistical Association, Proceedings of the Section on Survey Research Methods, Washington, DC.
- Rao, J. And Scott, A. (1987). On simple adjustments to chi-squared tests with sample survey data. Annals of Statistics, **15**, 385-397.
- Sarndal, C., Swensson, B., and Wretman, J. (1992). Model-Assisted Survey Sampling. Springer-Verlag, NY.
- Scott, A.J. and Holt, D. (1982). The effect of 2-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association* **77**, 848-854.
- Shah, B.V., Barnwell, B.G., and Bieler, G.S. (1996). SUDAAN User's Manual, Release 7, First Edition. Research Triangle Institute, RTP, NC.
- Shah, B.V. and LaVange, L.M. (1994). Mixed models for survey data. American Statistical Association, Proceedings of the Section on Survey Research Methods.
- Shah, B.V., Holt, M.M. and Folsom, R.E. (1977). Inference about regression models from sample survey data. Bulletin of the International Statistical Institute XLVII, **3**, 43-57.
- Thomas, D.R. and Rao, J.N.K. (1987). Small-sample comparisons of level and power for simple goodness-of-fit statistics under cluster sampling. *JASA* **82**, 630-636.
- Wolter, K.M. (1985). Introduction to Variance Estimation. NY: Springer-Verlag.
- Woodruff, R. (1971). A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association* **66**, 411-414.

**GEEs and Generalized Linear Models**

- Carr, G. and Portier, C. (1993). An evaluation of some methods for fitting dose-response models to quantal response data. *Biometrics* **49**, 779-791.
- Diggle, P., Liang, K.Y., and Zeger, S.L. (1994). Analysis of Longitudinal Data. NY: Oxford University Press.
- Dunlop, D.D. (1994). Regression for longitudinal data: a bridge from least squares regression. *American Statistician* **48**, 299-303.
- Huber, P.J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability **1**, 221-233.
- Karim, M.R. and Zeger, S.L. (1989). GEE: A SAS macro for longitudinal data analysis. Technical Report #674 from the Department of Biostatistics, The Johns Hopkins University.
- Liang, K. and Zeger, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.
- Lipsitz, S.R., Kim, K. and Zhao, L. (1994a). Analysis of repeated categorical data using generalized estimating equations. *Statistics in Medicine* **13**, 1149-1163.
- Lipsitz, S.R., Fitzmaurice, G.M., Orav, E.J. and Laird, N.M. (1994b). Performance of generalized estimating equations in practical situations. *Biometrics* **50**, 270-278.
- McCullagh, P. and Nelder, J.A. (1989). Generalized Linear Models. NY: Chapman and Hall.
- Neuhaus, J.M. and Segal, M.R. (1993). Design effects for binary regression models fitted to dependent data. *Statistics in Medicine* **12**, 1259-1268.
- Prentice, R.L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics* **44**, 1033-1048.
- Rotnitzky, A. and Jewell, N.P. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika* **77**, 485-497.
- Royall, R.M. (1986). Model robust confidence intervals using maximum likelihood estimators. *International Statistical Review* **54**, 221-226.
- Zeger, S. (1988). Commentary. *Statistics in Medicine* **7**, 161-168.
- Zeger, S. and Liang, K. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121-130.

### Random Effects Models

- Bryk, A.S. and Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, **101**, 147-158.
- Gibbons, R.D. and Hedeker, D. (1994). Application of random effects probit regression models. *Journal of Consulting and Clinical Psychology*, **62**, 285-296.
- Goldstein, H. (1987). Multilevel Models in Educational and Social Research. New York, NY: Oxford University Press.
- Hedeker, D.R., Gibbons, R.D., and Flay, B.R. (1994). Random-effects regression models for clustered data with an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*, **62**, 757-765.
- Hedeker, D.R., Gibbons, R.D., and Davis, J.M. (1991). Random regression models for multicenter clinical trials data. *Psychopharmacology Bulletin*, **27**, 73-77.
- Hedeker, D.R. and Gibbons, R.D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics* **50**, 933-944.
- Laird, N.M., Donnelly, C., and Ware, J.H. (1992). Longitudinal studies with continuous responses. *Statistical Methods in Medical Research* **1**, 225-247.
- Laird, N.M. and Ware, J.H. (1982). Random effects models for longitudinal data. *Biometrics* **38**, 963-974.
- Stiratelli, R., Laird, N., and Ware, J.H. (1984). Random-effects model for serial observations with binary responses. *Biometrics* **40**, 961-971.

### Survival Methods

- Greenwood, M. (1926). The natural duration of cancer. *Reports on Public Health and Medical Subjects*, **33**, 1-26.
- Lee, E.W., Wei, L.J., and Amato, D.A. (1992). "Cox-type regression analysis for large numbers of small groups of correlated failure time observations". In Klein, J.P. and Goel, P.K. (Eds.), Survival Analysis: State of the Art, Kluwer Academic Publishers, Dordrecht, 237-247.
- Lin, D.Y. (1994). Cox regression analysis of multivariate failure time data: the marginal approach. *Statistics in Medicine* **13**, 2233-2246.
- Lin, D.Y. and Wei, L.J. (1989). The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, **84**, 1074-1078.

### Comparisons of GEE and Random Effects Models

- Neuhaus, J.M. (1993). Estimation efficiency and tests of covariate effects with clustered binary data. *Biometrics* **49**, 989-996.
- Neuhaus, J.M., Kalbfleisch, J.D., and Hauck, W.W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review* **59**, 25-35.

Park, T. (1993). A comparison of the generalized estimating equation approach with the maximum likelihood approach for repeated measurements. *Statistics in Medicine* **12**, 1723-1732.

Zeger, S., Liang, KY, and Albert, P.S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics* **44**, 1049-1060.

### **Jackknife Variance Estimators**

Carr, G. And Portier, C. (1993). An evaluation of some methods for fitting dose-response models to quantal-response developmental toxicity data. *Biometrics* **49**, 779-791.

Gladen, B. (1979). The use of the jackknife to estimate proportions from toxicological data in the presence of litter effects. *Journal of the American Statistical Association* **74**, 278-283.

Lipsitz, S. And Parzen, M. (1996). A jackknife estimator of variance for Cox regression for correlated survival data. *Biometrics* **52**, 291-298.

Lipsitz, S., Dear, K., and Zhao, L. (1994). Jackknife estimators of variance for parameter estimates from estimating equations with applications to clustered survival data. *Biometrics* **50**, 842-846.

Quenouille, M. (1956). Notes on bias in estimation. *Biometrika* **43**, 353-360.

Tukey, J.W. (1958). Bias and confidence in not quite large samples. *Annals of Mathematical Statistics*, **29**, 614.

# **APPENDIX I**

## **Abstracts of Related Papers**





## Statistical Evaluation of Relationships Between Analgesic Dose and Ordered Ratings of Pain Relief Over an Eight-Hour Period

SA Gansky, GG Koch, and J. Wilson

**KEY WORDS:** Longitudinal design; relative potency; weighted least squares; ordinal response; multivariate analysis; marginal models.

1994, *Journal of Biopharmaceutical Statistics*, **4(2)**, 233-265.

Statistical considerations are discussed for the application of alternative methods to a clinical trial involving repeated ordinal ratings and multiple dosage levels of active drugs. Analyses included summary measures traditionally employed in studies of acute pain: sum of pain intensity differences from baseline, total pain relief, and total pain half gone. Estimators and confidence intervals of relative potency are developed for univariate and multivariate situations, using weighted least squares analysis with mean response and variances from Taylor series linearizations. The estimates from these methods are compared to those from traditional methods, such as ordinary least squares regression and Fieller's method for confidence intervals, as well as those from more recent developments, such as generalized estimating equations and sample survey data regression. A double-blind, two-center, randomized clinical trial of acute pain relief comparing placebo with two analgesics, each at two dosage levels, over an 8-hour period serves as an illustrative example for these techniques and comparisons.

## An Overview of Statistical Issues and Methods of Meta-Analysis

JE Schmid, GG Koch, and LM LaVange

**KEY WORDS:** Meta-analysis; random effects model; survey data regression; combination of studies

1991, *Journal of Biopharmaceutical Statistics*, **1(1)**, 103-120.

A meta-analysis is a statistical analysis of the data from some collection of studies in order to synthesize the results. In this paper we discuss issues that frequently arise in meta-analysis and give an overview of the methods used, with particular attention to the use of fixed- and random-effects approaches. The methods are then applied to two sample datasets.

## CLUSTER SAMPLING TECHNIQUES IN QUANTAL RESPONSE TERATOLOGY AND DEVELOPMENTAL TOXICITY STUDIES

Gayle S. Bieler and Rick L. Williams

**Contact: Gayle S. Bieler, Research Triangle Institute,  
P.O. Box 12194, Research Triangle Park, NC 27709**

**KEY WORDS:** clustered binary data; Taylor series variance, dose response; toxicology

This paper presents a model-free approach for evaluating teratology and developmental toxicity data involving clustered binary responses. In teratology studies, a major statistical problem arises from the effect of intralitter correlation, or the potential for littermates to respond similarly. Some statistical methods impose strict distributional assumptions to account for extra-binomial variation, while others rely on nonparametric resampling and empirical variance estimation techniques. Quasi-likelihood methods and GEE's, which model the marginal mean/variance relationship, also avoid strict distributional assumptions. The proposed approach, often used to analyze complex sample survey data, is based on a first-order Taylor series approximation and a between-cluster variance estimation procedure, yielding consistent variance estimates for binomial-based proportions and regression coefficients from dose-response models. The cluster sample technique, presented here in the context of a logistic dose-response model, incorporates many of the advantages of quasi-likelihood methods, are valid for any underlying correlation structure, and are adaptable to a variety of analytical settings. The results of a simulation study show the cluster sample technique to be a viable competitor to other methods currently receiving attention.

1995, *Biometrics* **51**, 764-776.

## ESTIMATION OF PROPORTIONAL HAZARDS MODELS FOR SURVIVAL TIMES WITH NESTED ERRORS

Rick L. Williams and Gayle S. Bieler

**KEY WORDS:** Correlated failure times; Taylor series variance approximation; Cox regression

A simple variance estimation method is demonstrated for proportional hazards models of survival times with a nested error structure. For example, this method is appropriate when repeated measurements are taken on a single animal in neurobehavioral toxicology experiments, when analyzing littermates in teratology studies, or whenever survival times are clustered, as in families or clinics. The method derives from Binder's (1992) work for complex sample surveys using a Taylor series approach which yields consistent estimates of the model parameters and their variance-covariance matrix. In a simulation study of the level and power of significance tests of model parameters, we find that the method maintains the specified Type 1 error rate. We further demonstrate the method using repeated measurements of time to avoidance in a neurobehavioral toxicology experiment. This approach is implemented in the SUDAAN software package.

1993, American Statistical Association, *Proceedings of the Biopharmaceutical Section*.

## REGRESSION ANALYSIS WITH CLUSTERED DATA: APPLICATION OF SURVEY METHODOLOGY

B. I. Graubard and E. L. Korn

**Contact: B. I. Graubard, National Cancer Institute,  
6130 Executive Blvd., Bethesda, MD 20892**

**KEY WORDS:** Population average, cluster specific, logistic regression

Clustered data are found in many different types of studies such as repeated measures, inter-rater agreement, household survey, cross-over and community randomized studies. Analyses based on population average (PA) and cluster specific (CS) models are two commonly used approaches for estimating treatment/exposure effects with clustered data. This paper gives conditions involving marginal balancing of the covariates and the treatment/exposure variable under which the PA and CS analyses will agree. A weighted PA analysis is proposed which will estimate CS treatment effects when marginal balancing does not hold. Methods for variance estimation which draw upon survey methodology are proposed for the weighted PA analysis.

1994, *Statistics in Medicine* **13**, 509-522.

## APPLICATION OF SAMPLE SURVEY METHODS FOR MODELING RATIOS TO INCIDENCE DATA

LM LaVange, GG Koch, LL Keyes, and PA Margolis

**KEY WORDS:** Incidence density analysis, weighted least squares, adverse event associations

We describe ratio estimation methods for analyzing incidence data from follow-up studies. Commonly used in survey data analysis, these ratio methods require minimal distributional assumptions and accurately account for random variability in the at-risk periods and correlations among repeated events. The methods are easy to understand, readily available via commercial software, and provide flexibility for a variety of analytical settings. We suggest that ratio methods may be useful for epidemiological and clinical studies in which quantities such as incidence of illness events or side effects of drug treatment are the focus. The basic strategy consists of a two-step process in which we first estimate subgroup specific incidence densities and their covariance matrix via a first order Taylor series approximation. We then fit log-linear models to the estimated ratios in order to assess covariate effects. The ability to produce direct estimates of adjusted incidence density ratios is an important advantage of this approach. We provide illustrative analyses of incidence data using ratio methods as well as survey logistic regression methods and two applications of generalized estimating equation methodology, repeated logistic and Poisson regression models, for comparison.

1994, *Statistics in Medicine* **13**, 509-522.

## **APPLICATIONS OF SURVEY SAMPLING METHODOLOGY TO ANALYSIS OF REPEATED MEASURES DATA STRUCTURES IN DENTISTRY**

**GM Davies and GG Koch**

**Key Phrases:** Intra-patient correlation, Survey data regression, Stratified cluster sample, Dental data

In the dental sciences, researchers often encounter data which are observed at the site or tooth level within each patient. Consequently, statistical methods for repeated measures or clustered outcomes are needed to analyze dental data with this structure. Survey sampling researchers have developed regression methods to analyze continuous or categorical data from stratified multistage cluster samples. Statistical packages like SUDAAN (Survey Data Analysis) are available to implement such analyses. This paper discusses applications of SUDAAN to a cross-sectional study of school aged children and a longitudinal study of elderly patients over 65. Continuous responses are analyzed with SUDAAN procedures for comparing ratio means and for multiple linear regression; dichotomous responses are analyzed with SUDAAN procedures for logistic regression. For each application, patients are managed as the primary sampling unit through which adjustment for the correlated data structure within patient is achieved.

1993, Presented at the *Biometrics Society Spring Meetings*.

## **ALTERNATIVE METHODS FOR ANALYZING CLUSTERED BINARY DATA IN OPHTHALMOLOGICAL STUDIES**

**AR Localio, JR Landis, SL Weaver, TJ Sharp  
Center for Biostatistics and Epidemiology, Penn State University**

This paper describes modeling alternatives for binary data in ophthalmology, where clustering (correlation) of eyes within subject is a well-recognized issue, to determine how long after fatal accident donor corneas can be transplanted safely. Rabbits were sacrificed, and corneas of one eye were examined microscopically to determine the rate of cell death (under 7 per 1,000). Rabbits (and remaining eyes) were either refrigerated or not, and cell death rate was determined at 6 and 24 hours. Modeling the effect of refrigeration and time on cell viability was performed using: (1) logistic regression (LR) (SAS CATMOD) with no consideration of clustering of cells within eye, or the correlation of eyes within rabbit, (2) exact LR using LogXact, (3) LR using SUDAAN and considering the clustering of cells within eye, (4) LR using SUDAAN and considering two stages of clustering, cell within eye and eye within rabbit, (5) GEE with a logistic link and binomial error that accounts for clustering of cells within eye, and (6) cells within rabbit, (7) GEE with log link and Poisson error term that conditions on eye and accounts for correlation of eyes within rabbit, (8) conditional LR that conditions on rabbit (STATA), and (9) exact conditional LR (LogXact). Methods are compared both in terms of interpretation of parameters and standard errors, and computing requirements for samples of 3,000 observations per eye and a total of 130,000 data points.

1993, Presented at the *Biometrics Society Spring Meetings*.

## ANALYSIS OF REPEATED MEASURES STUDIES WITH MULTIPLE REGRESSION METHODS FOR SAMPLE SURVEY DATA

**LM LaVange and GG Koch**

This presentation discusses how recently developed statistical procedures for fitting multiple regression models to sample survey data enables more effective analysis for repeated measures studies with complicated data structures. Situations where such methods are of interest include dermatology studies where treatment is applied to two or more sites of each patient, multi-visit studies where responses are observed at two or more points for each patient, dental studies where two or more teeth or dental areas of each patient receive treatment or are monitored over time for outcomes such as caries or progression of periodontal disease, multi-period crossover studies, and epidemiologic studies for repeated occurrences of adverse events or illnesses. For these situations, one can specify a primary sampling unit within which repeated measures have intraclass correlation. This intraclass correlation is taken into account by sample survey regression methods through robust estimates of the standard errors of the regression coefficients. Regression estimates are obtained from model fitting estimation equations which ignore the correlation structure of the data (i.e., computing procedures which assume that all observational units are independent or are from simple random samples). The analytic approach is straightforward to apply with logistic models for dichotomous data, proportional odds models for ordinal data, and linear models for continuously scaled data, and results are interpretable in terms of population average parameters. Several examples are presented to illustrate the capabilities of the methodology.

1994, Presented at the *Drug Information Association Annual Meeting*.

## MIXED MODELS FOR SURVEY DATA

**BV Shah and LM LaVange, Research Triangle Institute**

**Key Words:** Mixed models, Maximum Likelihood, Survey Data, Approximate F-test

The classic definition of the likelihood function is limited to simple random samples selected with equal probability. We propose a generalization of the likelihood function that allows for samples selected with unequal probabilities. With this approach, the problem of analyzing sample survey data with linear models is reduced to estimating fixed effects and their variances from a mixed model. In fact, standard methods for fitting linear models to survey data can be viewed as a MINQUE0 estimation procedure, given a set of assumptions regarding the model for the parent population. This approach has the advantage of making explicit the assumptions underlying the analysis methods used in current sample survey practice. We also explore adjustments to the degrees of freedom for the approximate F-test often used in survey data analysis. This approximation can be applied to mixed models in general. We present some simulation results to compare our proposed adjustment to previously recommended approximations. The adjusted F statistic allows for the analysis of cases previously thought to be intractable, and also allows for the analysis of survey data under different sets of assumptions, without ignoring the survey design. Of course, the paper raises many more questions regarding potential models and appropriate analysis techniques.

1994, Presented at the *Joint Statistical Meetings*.

## **Product-Limit Survival Functions with Correlated Survival Times**

**Rick L. Williams**

**Research Triangle Institute**

**Key Words:** Kaplan-Meier estimates, life tables, robust variance, Taylor series linearization, intracluster correlation

A simple variance estimator for product-limit survival functions is demonstrated for survival times with nested errors. Such data arise whenever survival times are observed within clusters of related observations. Greenwood's formula, which assumes independent observations, is not appropriate in this situation. A robust variance estimator is developed using Taylor series linearized values and the between-cluster variance estimator commonly used in multistage sample surveys. A simulation study shows that the between-cluster variance estimator is approximately unbiased and yields confidence intervals that maintain the nominal level for several patterns of correlated survival times. The simulation study also shows that Greenwood's formula underestimates the variance when the survival times are positively correlated within a cluster and yields confidence intervals that are too narrow. Extension to life table methods is also discussed.

1995, *Lifetime Data Analysis* **1**, 171-186.

## **Analysis of Prevention Program Effectiveness with Clustered Data Using Generalized Estimating Equations**

**E.C. Norton, G.S. Bieler, S.T. Ennett, and G.A. Zarkin**

Experimental studies of prevention programs often randomize clusters of individuals rather than individuals to treatment conditions. When the correlation among individuals within clusters is not accounted for in statistical analysis, the standard errors are biased, potentially resulting in misleading conclusions about the significance of treatment effects. This study demonstrates the Generalized Estimating Equation (GEE) method, focusing specifically on the GEE-independent method, to control for within-cluster correlation in regression models with either continuous or binary outcomes. The GEE-independent method yields consistent and robust variance estimates. Data from Project DARE, a youth drug use prevention program, are used for illustration.

1996, *Journal of Consulting and Clinical Psychology* **74**, 278-283.

# **APPENDIX II**

## **Examples**





## Teratology Experiment: Clustered Binary Data

### Evaluation of the Compound DEHP on Fetal Death

This example demonstrates the cluster sample or GEE model-fitting techniques (Zeger and Liang, 1986; Liang and Zeger, 1986) and the Jackknife in the context of a typical teratology experiment. For comparison, we include results based on a strictly binomial model (independence).

The typical teratology screening experiment involves administration of a compound to pregnant dams of a given animal species, followed by evaluation of the fetuses just prior to the end of gestation for various types of malformations. The experimental groups consist of a control group and anywhere from 2 to 4 exposed groups, representing increasing dosages of the compound under test. The data for this example have been taken from Butler (1988) and represent fetal death in CD-1 mice after administration of the compound DEHP at dosages of 0, 250, 500, 1000, or 1500 ppm during gestation. Sample sizes ranged from 24 to 30 litters per group. As reported by Butler, the average litter sizes were slightly larger in the control (13.2) vs. all other dose groups (11.5 to 12.3), but a dose-related trend was not evident for these data.

In this example, the observations on fetuses are clustered within litters, and the variance estimation techniques in SUDAAN are directly applicable for accounting for the intralitter correlation. The SUDAAN program produces dose-specific descriptive statistics (via PROC DESCRIPT) and fits a logistic dose-response model (via PROC LOGISTIC) based on the teratology experiment. For demonstration purposes, we fit two logistic models, one with a single regressor (dose level) and another with indicator variables corresponding to each treatment group.

The sample design option *WR* (shorthand notation for "with-replacement sampling") on the LOGISTIC and DESCRIPT procedure statements invokes the robust variance estimator that is appropriate for these experimental data. The *NEST* statement in SUDAAN indicates that litters (represented by DAM) represent the clusters. The requested test statistics *WALDCHI* and *SATADJCHI* refer to the usual Wald chi-squared test and the Satterthwaite-adjusted chi-squared test (Rao and Scott, 1987), respectively. The latter test is a modification of the usual Wald statistic and has been shown to have superior operating characteristics for multiple-degree-of-freedom hypotheses in small samples (Thomas and Rao, 1987).

The estimated dose group percentages and their standard errors under the cluster sample vs. strictly binomial models are contained in Figure 1. The incidence of fetal death was lowest in the control, 250 ppm, and 500 ppm groups (17%, 10%, and 13%, respectively) and highest in the 1000 ppm and 1500 ppm groups (50% and 84%, respectively).

Figure 1 also contains design effects for the binomial-based percentages. The design effect measures the inflation (or deflation) in variance of a sample statistic due to intracluster correlation beyond that expected if the data were independent. It is estimated as the ratio of the cluster sample variance obtained through Taylor linearization ( $V_{Cluster}$ ) vs. independence ( $V_{Indep}$ ). The predicted design effect for a mean or proportion is directly proportional to the size of the intracluster correlation and the cluster size (Kish and Frankel, 1974):

$$DEFF = 1 + \rho(m - 1) ,$$

where  $m$  is the constant cluster size and  $\rho$  is the intracluster correlation. Neuhaus and Segal (1993) showed that this relationship also provides accurate design effect approximations for coefficients from

binary response regression models with exchangeable correlations, a single cluster-level covariate, and variable cluster sizes. For the case of unequal cluster sizes, it has been recommended that  $m$  be replaced by a weighted analogue:

$$\tilde{m} = \frac{\sum_i \sum_j m_{ij}^2}{\sum_i \sum_j m_{ij}},$$

where  $m_{ij}$  is the cluster size for the  $j$ -th litter in dose group  $i$ .

Observed design effects ( $V_{Cluster}/V_{Indep}$ ) for the dose-specific percentages ranged from 0.85 to 6.32 for these data (see Figure 1). The 250 and 500 ppm groups had design effects just under 1.0 (when  $V_{Cluster} = V_{Indep}$ ), indicating small but slightly negative intralitter correlations. Using the Pearson correlation coefficient, Butler reported intracluster correlations of -0.01 in each of these two groups. The control and higher dose groups had correlations closer to 0.3 and 0.4, and we detected substantial design effects near 5.0 and above in these groups, indicating greater than a 5-fold increase in the strictly binomial variance due to intralitter correlation. The observed design effects closely corresponded to the predicted values (1) in each group, with predictions based on the dose-specific weighted litter sizes and correlations estimated by Butler.

To implement the cluster sample methods (via SUDAAN), we estimated the model parameters under a standard binomial likelihood and computed a robust variance estimate. This is also known as **ordinary logistic regression with a variance correction** and is equivalent to a GEE logistic model with independent “working” correlations (which we refer to as **GEE-independent**). The Wald chi-square test was used to evaluate the null hypothesis of no dose-related effect.

For comparison, the same logistic models were also fit using:

- 1) GEE logistic regression models under exchangeable intralitter correlations (**GEE-exchangeable**),
- 2) ordinary logistic regression with **Jackknife variance estimation**, and
- 3) **ordinary logistic regression with no variance correction**.

Results for the GEE and Jackknife approaches were essentially the same. For testing that the slope parameter from a linear logistic model is equal to zero (Figure 3), the GEE-exchangeable approach yielded a Z-statistic of 9.17, compared to a GEE-independent Z-statistic of 8.63 and a Jackknife Z-statistic of 8.41. The estimated slope parameter was slightly larger using the GEE approach with exchangeable correlations ( $\beta = 0.00256$  vs. 0.00249 for GEE-independent and Jackknife), but this had no substantial impact on test statistics. Estimated standard errors for the GEE-exchangeable and GEE-independent approaches were equivalent (0.00029), and for Jackknife the estimated standard error was 0.00030. The observed design effect for the logistic model slope parameter was over 5.0 for these data, reflecting substantial intralitter correlations. The impact of this design effect is manifested in an inflated Z-statistic of 19.76 obtained from ordinary logistic regression with **no** variance correction.

## Structure of the Fetal Death Data


Dose Group 1 = Control 2 = High Dose	Litter ID	Fetus ID	Y = fetal death 0 = alive 1 = dead
1	1	1	0
1	1	2	1
1	1	3	0
1	2	1	0
1	2	2	0
2	10	1	0
2	10	2	1
2	20	1	1
2	20	2	1
2	30	1	1

$N = 1,619$  records on the file  
 (1,619 fetuses clustered within 131 litters)

Figure 1

## Descriptive Statistics for Fetal Death in the DEHP Data

Dose Group	Number Litters	Number Fetuses	Total Dead	Percentage Dead	Standard Error		Design Effect	
					Cluster	Indep.	Obs.	Predicted
Control	30	396	66	16.67	4.11	1.87	4.82	4.79
250 ppm	26	320	32	10.00	1.53	1.68	0.83	0.88
500 ppm	26	319	42	13.17	1.84	1.89	0.95	0.88
1000 ppm	24	276	139	50.36	7.44	3.01	6.10	6.10
1500 ppm	25	308	258	83.77	4.65	2.10	4.89	4.93

  
**SUDAAN**      **Standard Packages:  
Too Small**

Cluster: SUDAAN (Descript Procedure)

Independence: Standard Statistical Packages (e.g., SAS)

$$\text{Observed DEFF} = \frac{V_{CLUSTER}}{V_{INDEPENDENCE}}$$

$$\text{Predicted DEFF} = 1 + \hat{\rho}_i (\bar{m}_i - 1)$$

$\bar{m}_i$  = dose-specific weighted litter sizes

= (13.62, 12.85, 12.75, 13.14, 12.56)

$\hat{\rho}_i$  = dose-specific intra-cluster correlation (Butler, 1988)

= (0.30, -0.01, -0.01, 0.42, 0.34)

Source: Bieler and Williams (1995), *Biometrics* **51**, 764-776.

Figure 2

## Logistic Regression for the DEHP Data

## Exposed vs. Control Group Contrasts

Contrast	Model-Fitting Method	$\beta_i$	S.E.	Z	P-Value
<b>250 Vs. Control</b>	GEE (indep)	-0.5878	0.3413	-1.72	0.0874
	GEE (exch corr)	-0.5214	0.3307	-1.58	0.1142
	Jackknife	-0.5878	0.3619	-1.62	0.1068
	Independence	-0.5878	0.2300	-2.56	0.0104
<b>500 Vs. Control</b>	GEE (indep)	-0.2769	0.3370	-0.82	0.4128
	GEE (exch corr)	-0.2269	0.3310	-0.69	0.4902
	Jackknife	-0.2769	0.3562	-0.78	0.4383
	Independence	-0.2769	0.2135	-1.30	0.1947
<b>1000 Vs. Control</b>	GEE (indep)	1.6239	0.4197	3.87	0.0002
	GEE (exch corr)	1.6938	0.4004	4.23	0.0000
	Jackknife	1.6239	0.4430	3.67	0.0004
	Independence	1.6239	0.1808	8.98	0.0000
<b>1500 Vs. Control</b>	GEE (indep)	3.2504	0.4523	7.19	0.0000
	GEE (exch corr)	3.3346	0.4470	7.46	0.0000
	Jackknife	3.2504	0.4792	6.78	0.0000
	Independence	3.2504	0.2051	15.85	0.0000

GEE (indep): SUDAAN Logistic Procedure  
 GEE (exch): SUDAAN Multilog Procedure  
 Jackknife: SUDAAN Logistic Procedure  
 Independence: Standard Packages (e.g., SAS Logistic)

Figure 3

## Logistic Regression for the DEHP Data

Test for Dose-Related Trend (  $H_0: \beta = 0$  )

Model-Fitting Method	$\beta$	S.E.	Z	P-Value	Design Effect	
					Observed	Predicted
GEE independent	0.00249	0.00029	8.63	0.0000	4.64	4.11
GEE exchangeable	0.00256	0.00029	9.17	0.0000		
Jackknife	0.00249	0.00030	8.41	0.0000		
Independence	0.00249	0.00013	19.76	0.0000		

GEE independent: SUDAAN Logistic Procedure  
 GEE exchangeable: SUDAAN Multilog Procedure  
 Jackknife: SUDAAN Logistic Procedure  
 Independence: Standard Packages (e.g., SAS Logistic)

Source: Bieler and Williams (1995), *Biometrics* **51**, 764-776.

$$\text{Observed DEFF} = \frac{V_{GEE \text{ Indep.}}}{V_{INDEPENDENCE}}$$

$$\text{Predicted DEFF} = 1 + \hat{\rho}_y (\tilde{n} - 1)$$

$$\tilde{n} = 13.01 \text{ for the DEHP data}$$

$$\hat{\rho}_y = 0.259 \text{ for the DEHP data}$$

## Figure 4

### The LEVEL.DBS File:

#### Contains Value Labels For SUDAAN Data Examples

DEAD	1	Yes
DEAD	0	No
DOSE_5	1	CONTROL
DOSE_5	2	250 ppm
DOSE_5	3	500 ppm
DOSE_5	4	1000 ppm
DOSE_5	5	1500 ppm
HRS	1	1 hr.
HRS	2	3 hrs.
HRS	3	5 hrs.
HRS	4	Pre-Dosing
SUDTRT	1	Treatment
SUDTRT	2	Placebo

#### *Record Layout for the LEVEL.DBS File:*

<u>Columns</u>	<u>Description</u>
1-8	Variable Name
9-10	Level of the Variable
17-66	Text Label For This Level of the Variable

Note: The LEVEL.DBS file can document multiple datasets in the same directory



**Example 1 Results:****Descriptive Statistics**

```

1  PROC DESCRIPT DATA="TERATA" FILETYPE=SAS NOMARG ATLEVEL1=2 DESIGN=WR;
2  NEST _ONE_ DAM;
3  WEIGHT _ONE_;
4  VAR DEAD;
5  CATLEVEL 1;
6  SUBGROUP DOSE_5;
7  LEVELS 5;
8  SETENV LABWIDTH=16 COLWIDTH=10 LINESIZE=78 DECWIDTH=2 PAGESIZE=60;
9  PRINT ATLEV1=" NUMBER LITTERS"
      NSUM= " NUMBER FETUSES"
      TOTAL="TOTAL DEAD"
      PERCENT="PERCENTAGE DEAD"
      SEPERCENT="STANDARD ERROR"
      DEFFPCT="DESIGN EFFECT"/
      STYLE=NCHS ATLEV1FMT=F7.0 NSUMFMT=F7.0 DEFFPCTFMT=F6.2
      SEPERCENTFMT=F8.2 TOTALFMT=F5.0;
10 TITLE "DESCRIPTIVE STATISTICS FOR TERATOLOGY DATA"
      " "
      "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of observations read	:	1619	Weighted count	:	1619
Denominator degrees of freedom	:	130			

**Example 1 Results:**

**Descriptive Statistics**

```

Date: 03-19-97          Research Triangle Institute          Page : 1
Time: 14:53:51         The DESCRIPT Procedure          Table : 1

DESCRIPTIVE STATISTICS FOR TERATOLOGY DATA

FETAL DEATH IN CD-1 MICE

-----
Variable          NUMBER    NUMBER    TOTAL    PERCENTAGE    STANDARD    DESIGN
  Dose Group      LITTERS  FETUSES  DEAD     DEAD          ERROR      EFFECT
-----
DEAD: Yes
  CONTROL          30        396       66        16.67         4.11       4.82
  250 ppm          26        320       32        10.00         1.53       0.83
  500 ppm          26        319       42        13.17         1.84       0.95
  1000 ppm         24        276      139        50.36         7.44       6.10
  1500 ppm         25        308      258        83.77         4.65       4.89
-----

DESCRIPT used
CPU time          : 3.74 seconds
Elapsed time      : 4 seconds
Virtual memory    : 0.84 MB
    
```

**These results are contained in Figure 1. Note the NEST statement specification of DAM as the primary sampling unit (the cluster). With DAM as the cluster and the sample design option WR (with-replacement), the standard errors reported in this table are adjusted for clustering.**

**Example 1 Results:****Descriptive Statistics**

```

11 PROC DESCRIPT DATA="TERATA" FILETYPE=SAS NOMARG DESIGN=WR;
12 NEST _ONE_ DAM;
13 WEIGHT _ONE_;
14 VAR DEAD;
15 CATLEVEL 1;
16 SUBGROUP DOSE_5;
17 LEVELS 5;
18 CONTRAST DOSE_5 = (-1 1 0 0 0) / NAME = "Low Dose Vs. Control";
19 CONTRAST DOSE_5 = (-1 0 1 0 0) / NAME = "500 ppm Vs. Control";
20 CONTRAST DOSE_5 = (-1 0 0 1 0) / NAME = "1500 ppm Vs. Control";
21 CONTRAST DOSE_5 = (-1 0 0 0 1) / NAME = "High Dose Vs. Control";
22 SETENV LABWIDTH=25 COLWIDTH=10 LINESIZE=78 DECWIDTH=2 PAGESIZE=60;
23 PRINT PERCENT="DIFFERENCE"
        SEPERCENT="STANDARD ERROR"
        T_PCT="T-STAT"
        P_PCT="P-VALUE" /
        STYLE=NCHS SEPERCENTFMT=F8.2 T_PCTFMT=F6.2 P_PCTFMT=F7.4;
24 TITLE "DESCRIPTIVE STATISTICS FOR TERATOLOGY DATA"
        " "
        "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of observations read	:	1619	Weighted count :	1619
Denominator degrees of freedom	:	130		

Here we construct *contrasts* to compare the percentages of dead pups across dose groups. We used the CATLEVEL statement to estimate percentages instead of proportions (the response DEAD is a 0-1 variable). The design option and NEST statements are equivalent to the previous run. There are 1,619 pups on the file and 130 denominator DF (#litters - 1) available for computing variance estimates.

**Example 1 Results:****Descriptive Statistics**

Date: 03-19-97                      Research Triangle Institute                      Page : 1  
 Time: 14:53:51                      The DESCRIPT Procedure                      Table : 1

**DESCRIPTIVE STATISTICS FOR TERATOLOGY DATA**

FETAL DEATH IN CD-1 MICE

Variable = DEAD: Yes.

```

-----
Contrast                                STANDARD
                                DIFFERENCE   ERROR   T-STAT   P-VALUE
-----
Low Dose Vs. Control                 -6.67      4.39    -1.52    0.1310
500 ppm Vs. Control                 -3.50      4.51    -0.78    0.4386
1500 ppm Vs. Control                33.70      8.50     3.96    0.0001
High Dose Vs. Control                67.10      6.21    10.81    0.0000
-----

```

DESCRIPT used

CPU time : 4.17 seconds  
 Elapsed time : 5 seconds  
 Virtual memory : 0.92 MB

Here we see that the 1,000 and 1,500 ppm groups have significantly higher fetal death rates than the control group.

**Example 1 Results: GEE-Independent Logistic Regression Model**

```

25  PROC LOGISTIC DATA="TERATA" FILETYPE=SAS DESIGN=WR;
26  NEST _ONE_ DAM;
27  WEIGHT _ONE_;
28  SUBGROUP DOSE_5;
29  LEVELS 5;
30  REFLEVEL DOSE_5 = 1;
31  MODEL DEAD = DOSE_5;
32  EFFECTS DOSE_5 = (-1 0 0 0 1) / NAME = "Control vs. High Dose";
33  TEST SATADJCHI WALDCHI;
34  SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
35  PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" OR LOWOR UPOR
        DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)" SATADCHI=" CHI-SQ (SAT.)"
        WALDCHP=" P-VALUE (WALD)" SATADCHP=" P-VALUE (SAT.)"
        /T_BETAfmt=F8.2 DEFTfmt=F6.2 SEBETAfmt=F8.6
        ORfmt=F5.2 LOWORfmt=F6.2 UPORfmt=F6.2
        DFFmt=F7.0 SATADJDFfmt=F8.2 WALDCHIFmt=F8.2 SATADCHIFmt=F8.2;
36  TITLE "TESTING DOSE GROUP HETEROGENEITY"
        " "
        "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of zero responses : 1082

Number of non-zero responses : 537

Parameters have converged in 4 iterations

Number of observations read : 1619 Weighted count: 1619

Observations used in the analysis : 1619 Weighted count: 1619

Observations with missing values : 0 Weighted count: 0

Denominator degrees of freedom : 130

Maximum number of estimable parameters for the model is 5

R-Square for dependent variable DEAD (Cox & Snell, 1989): 0.304579

Here we fit a *GEE logistic regression model with independent working correlations*. Dose group is modelled as a 5-level categorical covariate so we can compare each group to control. The REFLEVEL statement is used to select dose group level 1 (control) to be the reference level for DOSE\_5 in the model. The R-square statistic is based on Cox and Snell (1989) as the proportion of the log-likelihood that is explained by the model. The EFFECTS statement requests a single degree-of-freedom contrast comparing the high dose to control.

**Example 1 Results:**

**GEE-Independent Logistic Regression Model**

Date: 03-19-97                      Research Triangle Institute                      Page : 1  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD  
**TESTING DOSE GROUP HETEROGENEITY**

FETAL DEATH IN CD-1 MICE

-----

Independent Variables  
 and Effects

	BETA	S.E.	DESIGN EFFECT	T:BETA=0	P-VALUE
Intercept	-1.6094	0.296054	4.82	-5.44	0.0000
DOSE GROUP					
CONTROL	0.0000	0.000000	.	.	.
250 ppm	-0.5878	0.341270	2.20	-1.72	0.0874
500 ppm	-0.2769	0.337047	2.49	-0.82	0.4128
1000 ppm	1.6239	0.419743	5.39	3.87	0.0002
1500 ppm	3.2504	0.452258	4.86	7.19	0.0000

-----

Date: 03-19-97                      Research Triangle Institute                      Page : 2  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD  
**TESTING DOSE GROUP HETEROGENEITY**

FETAL DEATH IN CD-1 MICE

-----

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT.)	P-VALUE (WALD)	P-VALUE (SAT.)
OVERALL MODEL	5	3.60	357.23	107.13	0.0000	0.0000
MODEL MINUS INTERCEPT	4	3.01	132.94	94.87	0.0000	0.0000
INTERCEPT	.	.	.	.	.	.
DOSE_5	4	3.01	132.94	94.87	0.0000	0.0000
<b>Control vs. High Dose</b>	1	1.00	51.65	51.65	0.0000	0.0000

-----

**Example 1 Results:****GEE Independent Logistic Regression Model**

```

Date: 03-19-97          Research Triangle Institute          Page : 3
Time: 14:53:51         The LOGISTIC Procedure              Table : 1

Response variable DEAD: DEAD
TESTING DOSE GROUP HETEROGENEITY

FETAL DEATH IN CD-1 MICE

-----
Independent Variables
and Effects                Odds    Lower    Upper
                          Ratio    95%     95%
                          Ratio    Limit   Limit
-----
Intercept                  0.20    0.11    0.36
DOSE GROUP
  CONTROL                  1.00    1.00    1.00
  250 ppm                  0.56    0.28    1.09
  500 ppm                  0.76    0.39    1.48
  1000 ppm                 5.07    2.21   11.63
  1500 ppm                25.80   10.55   63.10
-----

LOGISTIC used
CPU time      : 7.75 seconds
Elapsed time  : 8 seconds
Virtual memory : 1.31 MB

```

These results indicate that the two highest dose groups have a significantly higher fetal death risk than the control group (odds ratios are 5.07 and 25.80, respectively). The treatment effect is statistically significant ( $p=0.0000$ ).

**Example 1 Results:****GEE-Independent Logistic Regression Model**

```

37 PROC LOGISTIC DATA="TERATA" FILETYPE=SAS DESIGN=WR;
38 NEST _ONE_ DAM;
39 WEIGHT _ONE_;
40 MODEL DEAD = DOSE;
41 TEST WALDCHI SATADJCHI;
42 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
43 PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF="DF" SATADJDF="ADJ DF"
      WALDCHI=" CHI-SQ (WALD)" SATADCHI=" CHI-SQ (SAT.)"
      WALDCHP=" P-VALUE (WALD)" SATADCHP=" P-VALUE (SAT.)"
      /SEBETAfmt=F8.6 DFFMT=F7.0 T_BETAfmt=F8.2 DEFTfmt=F6.2
      SATADJDFfmt=F8.2 WALDCHIFMT=F8.2 SATADCHIFMT=F8.2;
44 TITLE "TESTING DOSE-RELATED TREND"
      " "
      "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of zero responses : 1082

Number of non-zero responses : 537

Parameters have converged in 4 iterations

Number of observations read : 1619 Weighted count: 1619

Observations used in the analysis : 1619 Weighted count: 1619

Observations with missing values : 0 Weighted count: 0

Denominator degrees of freedom : 130

Maximum number of estimable parameters for the model is 2

R-Square for dependent variable DEAD (Cox & Snell, 1989): 0.277411

Now we model the treatment effect as a continuous covariate, using the actual dosage levels as the covariate values. For this reason, we do not use a SUBGROUP statement here.



**Example 1 Results:****GEE Independent Logistic Regression Model**

Date: 03-19-97                      Research Triangle Institute                      Page : 1  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD

**TESTING DOSE-RELATED TREND**

FETAL DEATH IN CD-1 MICE

-----  
 Independent Variables  
 and Effects

	BETA	S.E.	DESIGN EFFECT	T:BETA=0	P-VALUE
Intercept	-2.4300	0.255035	5.01	-9.53	0.0000
DOSAGE	0.0025	0.000289	5.27	8.63	0.0000

-----

Date: 03-19-97                      Research Triangle Institute                      Page : 2  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD

**TESTING DOSE-RELATED TREND**

FETAL DEATH IN CD-1 MICE

-----  
 Contrast

	DF	ADJ	DF	CHI-SQ (WALD)	CHI-SQ (SAT.)	P-VALUE (WALD)	P-VALUE (SAT.)
OVERALL MODEL	2	1.98		91.65	97.21	0.0000	0.0000
MODEL MINUS INTERCEPT	1	1.00		74.53	74.53	0.0000	0.0000
INTERCEPT	1	1.00		90.78	90.78	0.0000	0.0000
DOSE	1	1.00		74.53	74.53	0.0000	0.0000

-----

LOGISTIC used

CPU time : 6.92 seconds  
 Elapsed time : 7 seconds  
 Virtual memory : 1.24 MB

These results indicate there is a significant dose-related trend on the fetal death rate ( $p=0.0000$ ).

## Example 1 Results

### Jackknife Variance Estimation

Below are the results obtained using Jackknife variance estimation. The option *DESIGN=Jackknife* is added to the PROC statement. All other programming statements are the same as previous. We begin with dose group modelled as a categorical covariate.

```

45  PROC LOGISTIC DATA="TERATA" FILETYPE=SAS DESIGN=JACKKNIFE;
46  NEST _ONE_ DAM;
47  WEIGHT _ONE_;
48  SUBGROUP DOSE_5;
49  LEVELS 5;
50  REFLEVEL DOSE_5=1;
51  MODEL DEAD = DOSE_5;
52  EFFECTS DOSE_5 = (-1 0 0 0 1) / NAME = "Control vs. High Dose";
53  TEST SATADJCHI WALDCHI;
54  SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
55  PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" OR LOWOR UPOR
        DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)" SATADCHI=" CHI-SQ (SAT.)"
        WALDCHP=" P-VALUE (WALD)" SATADCHP=" P-VALUE (SAT.)"
        /T_BETAFMT=F8.2 DEFTFMT=F6.2 SEBETAFMT=F8.6
        ORFMT=F5.2 LOWORFMT=F6.2 UPORFMT=F6.2
        DFFMT=F7.0 SATADJDFMT=F8.2 WALDCHIFMT=F8.2 SATADCHIFMT=F8.2;
56  TITLE " " "TESTING DOSE GROUP HETEROGENEITY VIA JACKKNIFE"
        " " "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of observations read	:	1619	Weighted count:	1619
Observations used in the analysis	:	1619	Weighted count:	1619
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	130		

Maximum number of estimable parameters for the model is 5

Number of zero responses	:	1082
Number of non-zero responses	:	537

Parameters have converged in 4 iterations

R-Square for dependent variable DEAD (Cox & Snell, 1989): 0.304579

## Example 1 Results

## Jackknife Variance Estimation

Date: 03-19-97                      Research Triangle Institute                      Page : 1  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD

**TESTING DOSE GROUP HETEROGENEITY VIA JACKKNIFE**

FETAL DEATH IN CD-1 MICE

Independent Variables and Effects	DESIGN				
	BETA	S.E.	EFFECT	T:BETA=0	P-VALUE
Intercept	-1.6094	0.314927	5.45	-5.11	0.0000
DOSE GROUP					
CONTROL	0.0000	0.000000	.	.	.
250 ppm	-0.5878	0.361909	2.48	-1.62	0.1068
500 ppm	-0.2769	0.356192	2.78	-0.78	0.4383
1000 ppm	1.6239	0.443029	6.01	3.67	0.0004
1500 ppm	3.2504	0.479198	5.46	6.78	0.0000

Date: 03-19-97                      Research Triangle Institute                      Page : 2  
 Time: 14:53:51                      The LOGISTIC Procedure                      Table : 1

Response variable DEAD: DEAD

**TESTING DOSE GROUP HETEROGENEITY VIA JACKKNIFE**

FETAL DEATH IN CD-1 MICE

Contrast	DF		CHI-SQ	CHI-SQ	P-VALUE	P-VALUE
	DF	ADJ	(WALD)	(SAT.)	(WALD)	(SAT.)
OVERALL MODEL	5	3.59	327.07	96.31	0.0000	0.0000
MODEL MINUS INTERCEPT	4	3.00	119.97	85.19	0.0000	0.0000
INTERCEPT	.	.	.	.	.	.
DOSE_5	4	3.00	119.97	85.19	0.0000	0.0000
Control vs. High Dose	1	1.00	46.01	46.01	0.0000	0.0000

Here we see that the *estimated regression coefficients* for the Jackknife are identical to those used for GEE-independent, but the estimated standard errors are just slightly larger. Nevertheless, the *p*-values from the two approaches are still quite similar, and both approaches have been shown to be valid for adjusting for intraclass correlation.

## Example 1 Results

### Jackknife Variance Estimation

Date: 03-19-97	Research Triangle Institute	Page : 3	
Time: 14:53:51	The LOGISTIC Procedure	Table : 1	
Response variable DEAD: DEAD			
<b>TESTING DOSE GROUP HETEROGENEITY VIA JACKKNIFE</b>			
FETAL DEATH IN CD-1 MICE			
-----			
Independent Variables and Effects	Odds Ratio	Lower 95% Limit	Upper 95% Limit
-----			
Intercept	0.20	0.11	0.37
DOSE GROUP			
CONTROL	1.00	1.00	1.00
250 ppm	0.56	0.27	1.14
500 ppm	0.76	0.37	1.53
1000 ppm	5.07	2.11	12.18
1500 ppm	25.80	10.00	66.56
-----			
LOGISTIC used			
CPU time	: 19.99 seconds		
Elapsed time	: 20 seconds		
Virtual memory	: 1.25 MB		

Since the estimated standard errors are slightly larger for the Jackknife vs. GEE-independent approaches using these data, the **95% confidence bands** around the **estimated odds ratios** are also slightly wider using the Jackknife. Note that the odds ratios themselves are identical because the same regression coefficients are used for both approaches.

## Example 1 Results

### Jackknife Variance Estimation

```

57 PROC LOGISTIC DATA="TERATA" FILETYPE=SAS DESIGN=JACKKNIFE;
58 NEST _ONE_ DAM;
59 WEIGHT _ONE_;
60 MODEL DEAD = DOSE;
61 TEST WALDCHI SATADJCHI;
62 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
63 PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF="DF" SATADJDF="ADJ DF"
      WALDCHI=" CHI-SQ (WALD)" SATADCHI=" CHI-SQ (SAT.)"
      WALDCHP=" P-VALUE (WALD)" SATADCHP=" P-VALUE (SAT.)"
      /SEBETAFMT=F8.6 DFFMT=F7.0 T_BETAFMT=F8.2 DEFTFMT=F6.2
      SATADJDFMT=F8.2 WALDCHIFMT=F8.2 SATADCHIFMT=F8.2;
64 TITLE " " "TESTING DOSE-RELATED TREND VIA JACKKNIFE"
      " " "FETAL DEATH IN CD-1 MICE";

```

Opened SAS data file C:\TERA\EXAMPLES\TERATA.SSD for reading.

Number of observations read	:	1619	Weighted count:	1619
Observations used in the analysis	:	1619	Weighted count:	1619
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	130		

Maximum number of estimable parameters for the model is 2

Number of zero responses	:	1082
Number of non-zero responses	:	537

Parameters have converged in 4 iterations

R-Square for dependent variable DEAD (Cox & Snell, 1989): 0.277411

Here are the Jackknife results with dosage modelled as a continuous covariate.

## Example 1 Results

### Jackknife Variance Estimation

```

Date: 03-19-97          Research Triangle Institute          Page : 1
Time: 14:53:51         The LOGISTIC Procedure          Table : 1

Response variable DEAD: DEAD

TESTING DOSE-RELATED TREND VIA JACKKNIFE

FETAL DEATH IN CD-1 MICE

-----
Independent Variables
and Effects
                BETA      S.E.  DESIGN  T:BETA=0  P-VALUE
-----
Intercept      -2.4300 0.262856   5.32    -9.24    0.0000
DOSAGE         0.0025 0.000297   5.56     8.41    0.0000
-----

```

```

Date: 03-19-97          Research Triangle Institute          Page : 2
Time: 14:53:51         The LOGISTIC Procedure          Table : 1

Response variable DEAD: DEAD

TESTING DOSE-RELATED TREND VIA JACKKNIFE

FETAL DEATH IN CD-1 MICE

-----
Contrast
                DF      ADJ DF      CHI-SQ   CHI-SQ   P-VALUE   P-VALUE
                (WALD)  (SAT.)  (WALD)   (SAT.)  (WALD)   (SAT.)
-----
OVERALL MODEL      2      1.98      86.29    92.58    0.0000    0.0000
MODEL MINUS INTERCEPT  1      1.00      70.66    70.66    0.0000    0.0000
INTERCEPT       1      1.00      85.46    85.46    0.0000    0.0000
DOSE              1      1.00      70.66    70.66    0.0000    0.0000
-----

LOGISTIC used
CPU time      : 16.92 seconds
Elapsed time  : 17 seconds
Virtual memory : 1.24 MB

```

These Jackknife results are almost identical to the GEE-independent results shown earlier.



## Cross-Over Clinical Trial with Multivariate Failure Time Data:

### Evaluation of a Coronary Heart Disease Drug on Repeated Exercise Times to Angina Pectoris

This example demonstrates SUDAAN's correlated data techniques in the context of a clinical trial. The data for this example represent repeated exercise times (in seconds) to angina pectoris in patients with coronary heart disease. We analyzed the data reported by Crouchley and Pickles (1993), in which 21 subjects were each tested four times on one day and a further four times two days later. On each day exercise time measurements were taken just before and at 1 hour, 3 hours, and 5 hours following drug administration. On one day the drug was an active treatment (an oral dose of isosorbide dinitrate) and on the other placebo. Although undertaken as a double-blind randomized cross-over design, the published data do not indicate the order of treatment, preventing any testing for carry-over effects.

The Cox proportional hazards model was used to evaluate the regression effect of treatment (or test day), after adjusting for several covariates: time since drug administration (4-level factor), and indicators for previous myocardial infarction (MI), previous coronary artery bypass surgery (CAB), and previous propranolol treatment (PP). Note that treatment day and time since drug administration are *within-cluster* covariates, while MI, PP, and CAB represent *cluster-level* covariates. For comparison, we include results based on assuming complete independence among the 8 failure times per subject.

The SUDAAN program contains code to fit the Cox proportional hazards model to the observed event times. The default sample design option *DESIGN=WR* (shorthand notation for "with-replacement sampling") invokes the robust variance estimator that is appropriate for the study. The *NEST* statement in SUDAAN indicates that the patient (PATIENT) represents the cluster or primary sampling unit, with the keyword *\_ONE\_* indicating there is a single design stratum. Additional sources of intracluster correlation, such as time within each study day, need not be specified. The requested test statistics *WALDCHI* and *SATADJCHI* refer to the usual Wald chi-square test and the Satterthwaite-adjusted chi-square test (Rao and Scott, 1987), respectively. The latter test is a modification of the usual Wald statistic and has been shown to have superior operating characteristics for multiple-degree-of-freedom hypotheses in small samples (Thomas and Rao, 1987).

Three sets of proportional hazards models were fit:

- 1) Model 1 was the *main effects model*, and it included the main effects of treatment (or study day), time since drug administration (modelled as a 4-level categorical variable corresponding to pre-dosing, 1-hour, 3-hours, and 5-hours post-dosing), and the three continuous covariates MI, CAB, and PP.
- 2) Model 2 was the *interaction model*, containing the main effects in Model 1 and the interaction effects between treatment and time since drug administration.
- 3) Finally, in Model 3 we evaluated the *simple effects of treatment* at each of the four times since drug administration. Model 3 required four separate runs of the proportional hazards model containing the treatment effect and the three continuous covariates. The four runs corresponded to each of the four times since drug administration.

SUDAAN results from fitting Models 1-3 are contained in the SUDAAN output, and results from the



main effects model are contained in Figure 1.

To implement the cluster sample methods using SUDAAN, we estimated the model parameters under a standard partial likelihood and applied a robust variance estimator (labelled *Robust* in Figure 1). The Wald chi-square test was used to evaluate the null hypothesis of no treatment effect. For comparison, the same proportional hazards model was also fit assuming complete independence of the response times (labelled *Naive* in Figure 1).

Figure 1 contains results for the main effects model. Note that for parameters which represent *cluster-level covariates*, the cluster sample method results in a substantial *increase* in standard errors. However, for *within-cluster covariates* (e.g., the treatment and time effects), the cluster variance estimates are substantially *smaller* than the independence estimates. Using the design effect results of Neuhaus and Segal (1993) and proceeding by analogy to failure time data, the large observed design effects for the cluster-level covariates (e.g., previous bypass surgery) indicate large response intracluster correlations. In this situation, the variance of the regression coefficients for such covariates is increased. However, the observed design effects for within-cluster covariates whose patterns do not vary from cluster to cluster (time since drug administration and treatment day) were much less than 1 (as low as 0.30), which would be expected when the response intracluster correlation is positive and the covariate intracluster correlation is negative. In this case, variance estimates for the regression coefficients would be smaller than that expected under independence, corresponding to a gain in efficiency.

As seen in Figure 1, tests for treatment effects and time since drug administration were statistically significant under the cluster sample and independence approaches, but were slightly more significant under the cluster sample approach. Using cluster sample techniques, SUDAAN reports the estimated hazard ratio for treatment vs. control in the main effects only model to be 0.43, with a 95% confidence interval of (0.32 - 0.59). A hazard ratio less than 1.0 indicates longer exercise times in the treatment group (a protective effect against angina pectoris), and this can be seen in the predicted survival (Kaplan-Meier) functions (computed at pre-dosing, and 1-, 3-, and 5-hours post-dosing). The Kaplan-Meier functions suggest that the treatment differences are largest at 1 and 3-hours post-dosing, and in fact, SUDAAN reports a significant interaction effect between treatment day and time since drug administration ( $p=0.0204$ , Wald chi-square test). The estimated hazard ratios at 1 and 3-hours post-dosing are 0.28 and 0.34, respectively; and the hazard ratios at pre-dosing and 5-hours post-dosing are 0.56 and 0.48, respectively.

Tests for the cluster-level covariates (previous MI, bypass surgery, and propranolol treatment) became less significant under the cluster sample approach, and only previous myocardial infarction remained statistically significant in each of the three models (interaction, main effects, and time-specific treatment effects models) due to the large design effects. A user-defined general linear contrast for testing the joint effects of the three covariates is demonstrated for the main effects model (via the EFFECTS statement).

## Cross-Over Clinical Trial

---

Repeated Exercise Times to Angina Pectoris  
(Crouchley and Pickles, *Biometrics*, 1993)

- Double-blind randomized cross-over design  
(not enough info to test carry-over effects)
- 21 male patients (clusters) with coronary heart disease
- Tested 4 times on each of two consecutive days  
(Cluster size = 8)

Just before drug administration

1 hr post

3 hrs post

5 hrs post

- One day: Active treatment (isosorbide dinitrate)  
Other day: Placebo
- Outcome at each of 8 time points:

$y$  = exercise time to angina pectoris (in seconds)

**Question:** Does treatment delay the time to angina pectoris, after adjusting for time since drug administration and previous conditions?

## Structure of the Angina Data

Patient ID	Treatment Day	Time Since Drug Admin (Hours)	Y = Exercise Time (seconds)	MI
1	1 = Placebo Day	1 = Pre	150	1
1	1	2 = 1 hr	172	1
1	1	3 = 3 hrs	118	1
1	1	4 = 5 hrs	143	1
1	2 = Treatment Day	1	136	1
1	2	2	445	1
1	2	3	393	1
1	2	4	226	1
2	1 = Placebo Day	1	205	0
2	1	2	287	0
2	1	3	211	0
2	1	4	207	0
2	2 = Treatment Day	1	250	0
2	2	2	306	0
2	2	3	206	0
2	2	4	224	0

**N** = 168 records (21 patients, 8 records per patient)

## Exercise Time to Angina Pectoris

---

### Proportional Hazards Model Results

Estimated Regression Coefficient: Treatment vs. Placebo	Estimated Hazards Ratio	Standard Error of Beta		Variance Ratio
		Cluster	Independent	
-0.8395	0.43	<b>0.1474</b>	<b>0.1724</b>	<b>0.73</b> (27% reduction)

  
**SUDAAN**                      **Standard Packages:**  
*Too Large*

- True variance ***smaller*** than under independence
- May fail to detect a treatment effect

Figure 1

## Proportional Hazards Regression for Exercise Time Data

## Main Effects Model

Covariate	Model-Fitting Method	$\beta_i$	S.E.	Design Effect <sup>1</sup>	Z	P-Value
<b>Treatment Day (Treatment vs. Placebo)</b>	Robust	-0.8395	0.1474	0.73	-5.70	.0000
	Naive	-0.8395	0.1724	1.00	-4.87	.0000
<b>Time Since Drug Administration</b>						
<b>1 hour</b>	Robust	-0.9295	0.2085	0.74	-4.46	.0001
	Naive	-0.9295	0.2417	1.00	-3.85	.0001
<b>3 hours</b>	Robust	-0.6040	0.1294	0.31	-4.67	.0001
	Naive	-0.6040	0.2311	1.00	-2.61	.0090
<b>5 hours</b>	Robust	-0.1827	0.1216	0.30	-1.50	.1487
	Naive	-0.1827	0.2232	1.00	-0.82	.4130
<b>Previous MI</b>	Robust	-1.2263	0.3636	3.29	-3.37	.0030
	Naive	-1.2263	0.2004	1.00	-6.12	.0000
<b>Previous Bypass Surgery</b>	Robust	0.7525	0.4025	4.17	1.87	.0762
	Naive	0.7525	0.1970	1.00	3.82	.0000
<b>Previous Propranolol Treatment</b>	Robust	-0.6282	0.4737	4.71	-1.33	.1998
	Naive	-0.6282	0.2182	1.00	-2.88	.0040

Number Clusters = 21; Cluster Size = 2 days X 4 times each day = 8

Estimated Hazard Ratio = 0.4319 (over 50% reduction in hazard, treatment vs. control)

Notes: Significant treatment-by-time interaction effect (via SUDAAN,  $p < 0.05$ )  
Largest effects occur at 1 and 3 hours post-dosing.

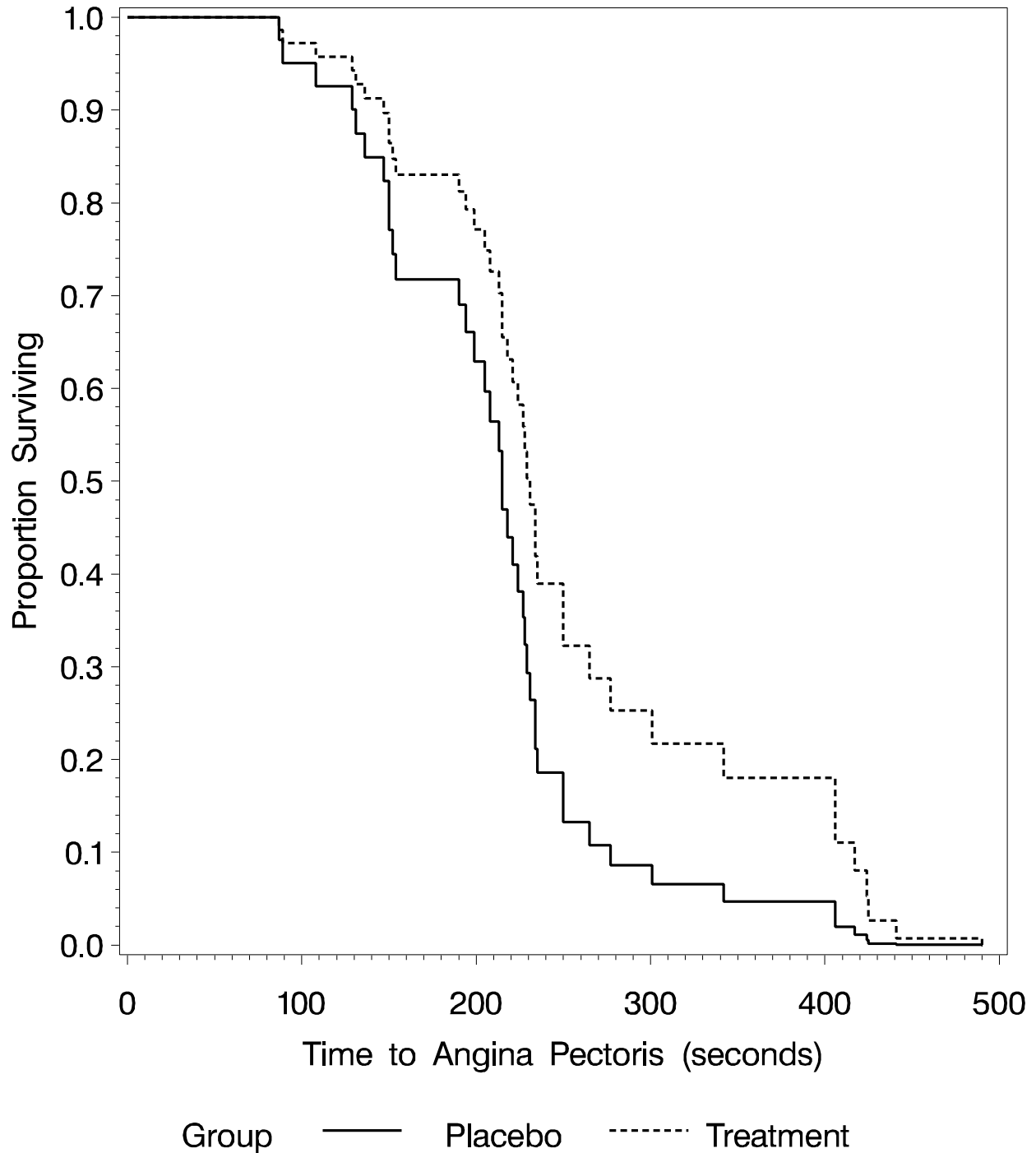
$$^1 \text{ Design Effect} = \left( \frac{SE_{Robust}}{SE_{Naive}} \right)^2$$

# Exercise Time To Angina Pectoris

Time-Specific Predicted Survival Functions

Adjusted for Previous Conditions

Time = Pre-Dosing

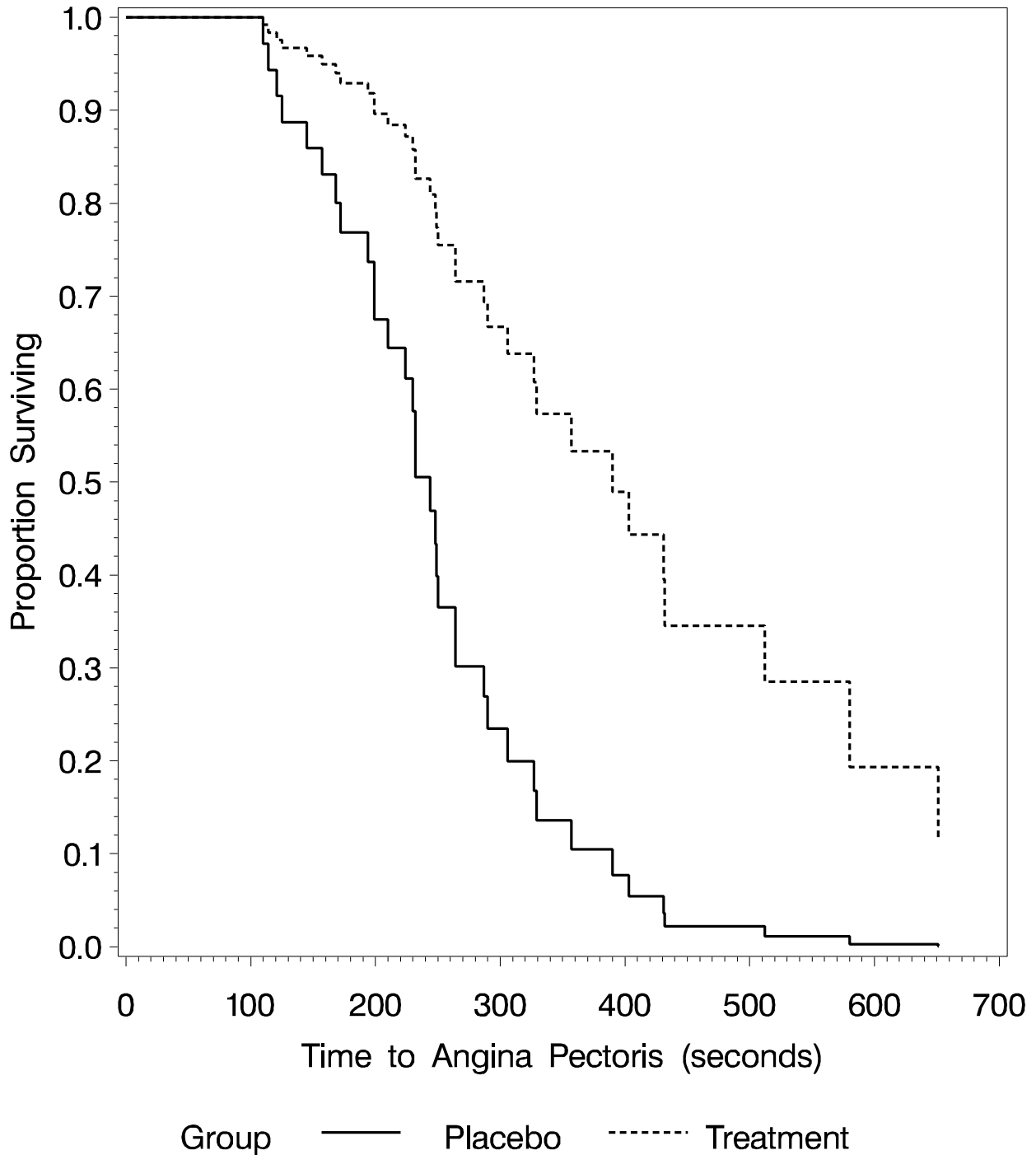


# Exercise Time To Angina Pectoris

Time-Specific Predicted Survival Functions

Adjusted for Previous Conditions

Time=1 hr. Post

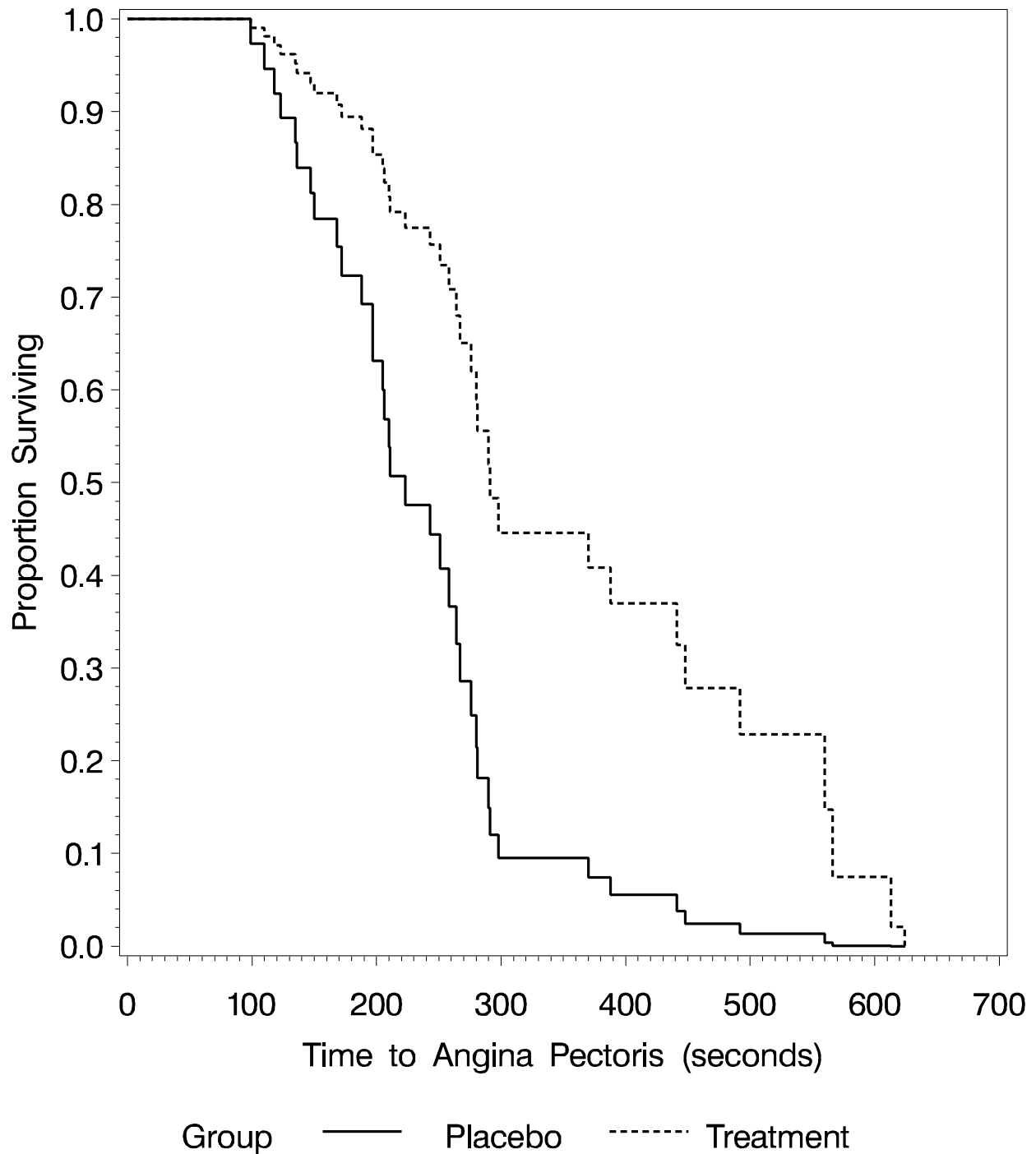


# Exercise Time To Angina Pectoris

Time-Specific Predicted Survival Functions

Adjusted for Previous Conditions

Time = 3 hrs. Post



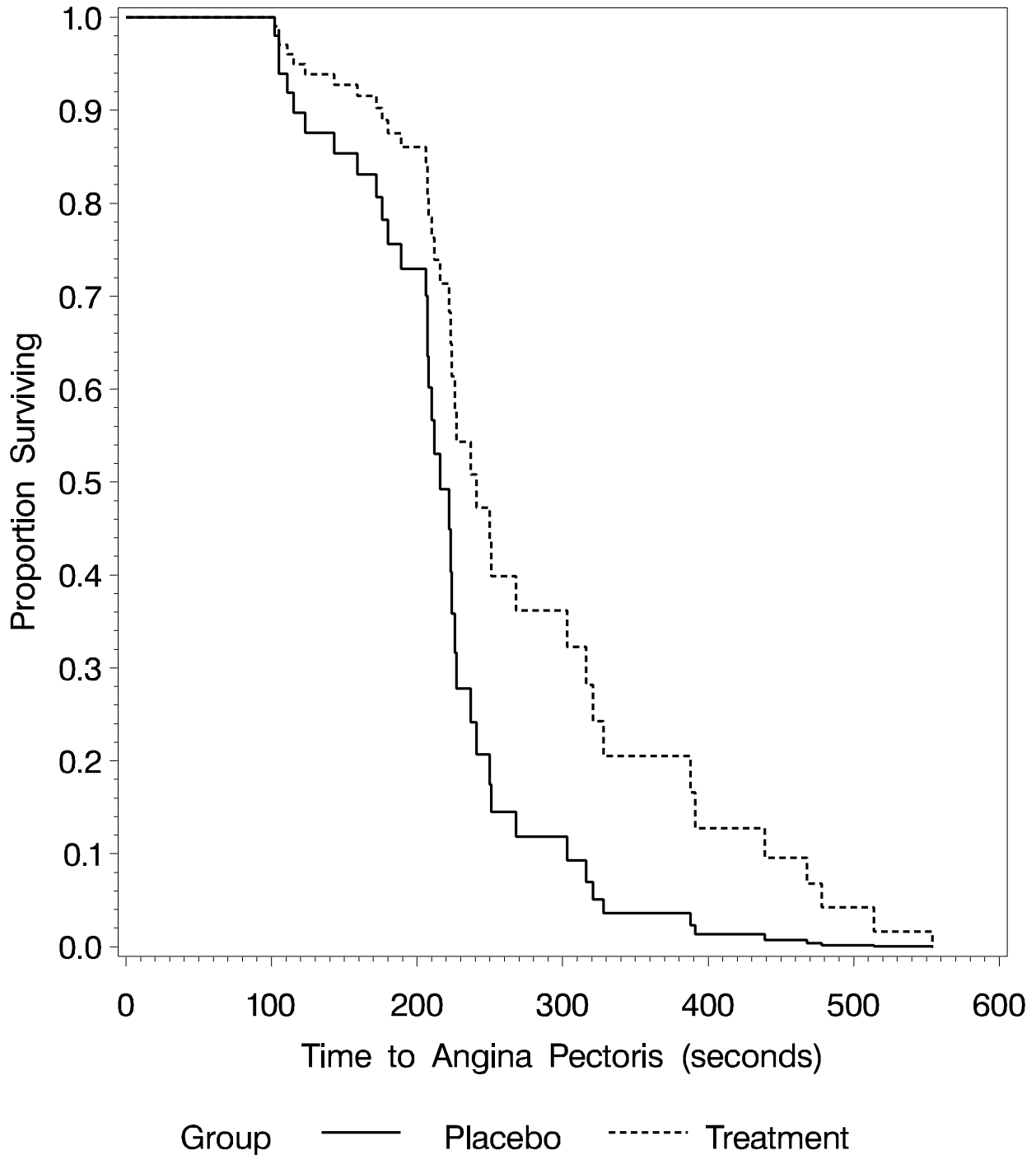


# Exercise Time To Angina Pectoris

Time-Specific Predicted Survival Functions

Adjusted for Previous Conditions

Time = 5 hrs. Post



## Example 2 Results: Testing Interaction

```

14 PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
15 NEST _ONE_ PATIENT;
16 WEIGHT _ONE_;
17 SUBGROUP HRS SUDTRT;
18 LEVELS 4 2;
19 EVENT COMPLETE;
20 MODEL EXTIME = SUDTRT HRS SUDTRT*HRS MI CAB PP;
21 TEST WALDCHI SATADJCHI;
22 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
23 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
      P_BETA="P-VALUE"
      DF="DF" SATADJDF="ADJ DF"
      WALDCHI=" CHI-SQ (WALD)"
      SATADCHI=" CHI-SQ (SAT)"
      WALDCHP=" P-VALUE (WALDC)"
      SATADCHP=" P-VALUE (SAT)"
      /DFFMT=F7.0 BETAFMT=F10.6 SEBETAFMT=F10.6 T_BETAFMT=F8.2 WALDCHPFMT=F8.4
      P_BETAFMT=F8.4 SATADCHPFMT=F8.4 DEFTFMT=F6.2;
24 TITLE " "
      "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
      " "
      "PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR"
      " " "Interaction Model";
25 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

```

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read	:	168	Weighted count:	168
Observations used in the analysis	:	168	Weighted count:	168
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	20		

Maximum number of estimable parameters for the model is 10

Number of non-censored events:	155
Number of censored events	: 13

SURVIVAL has converged to a solution in 5 iterations.

**Example 2 Results: Testing Interaction**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR

**Interaction Model**

-----  
 Independent Variables  
 and Effects

	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
-----					
Day					
Treatment	-0.405588	0.133014	0.18	-3.05	0.0063
Placebo	0.000000	0.000000	.	.	.
Hours Since Drug Admin					
1 hr.	-0.463372	0.201299	0.42	-2.30	0.0322
3 hrs.	-0.339857	0.132493	0.18	-2.57	0.0185
5 hrs.	-0.087686	0.113670	0.13	-0.77	0.4495
Pre-Dosing	0.000000	0.000000	.	.	.
Day, Hours Since Drug Admin					
Treatment, 1 hr.	-1.107631	0.413010	0.72	-2.68	0.0143
Treatment, 3 hrs.	-0.639324	0.251528	0.30	-2.54	0.0194
Treatment, 5 hrs.	-0.228561	0.195745	0.19	-1.17	0.2567
Treatment, Pre-Dosing	0.000000	0.000000	.	.	.
Placebo, 1 hr.	0.000000	0.000000	.	.	.
Placebo, 3 hrs.	0.000000	0.000000	.	.	.
Placebo, 5 hrs.	0.000000	0.000000	.	.	.
Placebo, Pre-Dosing	0.000000	0.000000	.	.	.
Previous MI	-1.239716	0.370078	3.38	-3.35	0.0032
Previous Bypass Surgery	0.736154	0.403746	4.18	1.82	0.0832
Previous Propranolol Trt	-0.615225	0.484650	4.91	-1.27	0.2189

-----  
 Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: Testing Interaction**

Date: 03-24-97                      Research Triangle Institute                      Page : 2  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR

**Interaction Model**

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
OVERALL MODEL	10	3.98	44.84	20.81	0.0000	0.0004
SUDTRT	.	.	.	.	.	.
HRS	.	.	.	.	.	.
<b>SUDTRT * HRS</b>	<b>3</b>	<b>1.80</b>	<b>9.80</b>	<b>10.35</b>	<b>0.0204</b>	<b>0.0046</b>
MI	1	1.00	11.22	11.22	0.0008	0.0008
CAB	1	1.00	3.32	3.32	0.0683	0.0686
PP	1	1.00	1.61	1.61	0.2043	0.2046

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used  
 CPU time : 3.29 seconds  
 Elapsed time : 4 seconds  
 Virtual memory : 1.08 MB

**Example 2 Results: Testing Main Effects**

```

1  PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
2  NEST _ONE_ PATIENT;
3  WEIGHT _ONE_;
4  SUBGROUP HRS SUDTRT;
5  LEVELS 4 2;
6  EVENT COMPLETE;
7  MODEL EXTIME = SUDTRT HRS MI CAB PP;
8  EFFECTS MI CAB PP / NAME = "Combined Effect: MI,CAB,PP";
9  TEST WALDCHI SATADJCHI;
10 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
11 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" HR LOWHR UPHR DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)" SATADCHI=" CHI-SQ (SAT)"
        WALDCHP=" P-VALUE (WALDC)" SATADCHP=" P-VALUE (SAT)"
        /DFMT=F7.0 BETAFMT=F10.6 SEBETAFMT=F10.6 T_BETAFMT=F8.2 WALDCHPFMT=F8.4
        P_BETAFMT=F8.4 SATADCHPFMT=F8.4 DEFTFMT=F6.2
        HRFMT=F7.2 LOWHRFMT=F6.2 UPHRFMT=F6.2;
12 TITLE " "
        "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
        " "
        "PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR:"
        " " "Main Effects Model";
13 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

```

NOTE: Terms in the MODEL statement have been rearranged  
to follow subgroup order.

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read	:	168	Weighted count:	168
Observations used in the analysis	:	168	Weighted count:	168
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	20		

Maximum number of estimable parameters for the model is 7

Number of non-censored events: 155  
Number of censored events : 13  
SURVIVAL has converged to a solution in 5 iterations.

**Example 2 Results: Testing Main Effects**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR:

**Main Effects Model**

```
-----
```

Independent Variables and Effects	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
-----					
Hours Since Drug Admin					
1 hr.	-0.929513	0.208504	0.74	-4.46	0.0002
3 hrs.	-0.603992	0.129440	0.31	-4.67	0.0001
5 hrs.	-0.182658	0.121615	0.30	-1.50	0.1487
Pre-Dosing	0.000000	0.000000	.	.	.
Day					
Treatment	-0.839508	0.147408	0.73	-5.70	0.0000
Placebo	0.000000	0.000000	.	.	.
Previous MI	-1.226269	0.363640	3.29	-3.37	0.0030
Previous Bypass Surgery	0.752530	0.402488	4.17	1.87	0.0762
Previous Propranolol Trt	-0.628185	0.473715	4.71	-1.33	0.1998

```
-----
```

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: Testing Main Effects (continued)**

Date: 03-24-97

Research Triangle Institute

Page : 2

Time: 08:50:19

The SURVIVAL Procedure

Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR:

**Main Effects Model**

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
OVERALL MODEL	7	3.57	49.58	20.57	0.0000	0.0003
HRS	3	2.29	31.22	30.73	0.0000	0.0000
SUDTRT	1	1.00	32.43	32.43	0.0000	0.0000
MI	1	1.00	11.37	11.37	0.0007	0.0008
CAB	1	1.00	3.50	3.50	0.0615	0.0618
PP	1	1.00	1.76	1.76	0.1848	0.1851
<b>Combined Effect: MI,CAB,PP</b>	<b>3</b>	<b>2.86</b>	<b>15.43</b>	<b>13.17</b>	<b>0.0015</b>	<b>0.0039</b>

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: Testing Main Effects (continued)**

Date: 03-24-97                      Research Triangle Institute                      Page : 3  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING ROBUST VARIANCE ESTIMATOR:

**Main Effects Model**

Independent Variables and Effects	Hazards Ratio	Lower 95% Limit	Upper 95% Limit
Hours Since Drug Admin			
1 hr.	0.39	0.26	0.61
3 hrs.	0.55	0.42	0.72
5 hrs.	0.83	0.65	1.07
Pre-Dosing	1.00	1.00	1.00
Day			
Treatment	0.43	0.32	0.59
Placebo	1.00	1.00	1.00
Previous MI	0.29	0.14	0.63
Previous Bypass Surgery	2.12	0.92	4.91
Previous Propranolol Trt	0.53	0.20	1.43

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used

CPU time : 3.0 seconds  
 Elapsed time : 3 seconds  
 Virtual memory : 1.08 MB



**Example 2 Results: Pre-Dosing Treatment Effect**

```

26 PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
27 NEST _ONE_ PATIENT;
28 WEIGHT _ONE_;
29 SUBPOPN HOURS=1 / NAME = "PRE-DOSING TREATMENT EFFECT";
30 SUBGROUP SUDTRT;
31 LEVELS 2;
32 EVENT COMPLETE;
33 MODEL EXTIME = SUDTRT MI CAB PP;
34 TEST WALDCHI SATADJCHI;
35 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
36 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" HR LOWHR UPHR
      DF="DF" SATADJDF="ADJ DF"
      WALDCHI=" CHI-SQ (WALD)"
      SATADCHI=" CHI-SQ (SAT)"
      WALDCHP=" P-VALUE (WALDC)"
      SATADCHP=" P-VALUE (SAT)"
      /DFFMT=F7.0 BETAFMT=F10.6 SEBETAFMT=F10.6 T_BETAFMT=F8.2 WALDCHPFMT=F8.4
      P_BETAFMT=F8.4 SATADCHPFMT=F8.4 DEFTFMT=F6.2
      HRFMT=F7.2 LOWHRFMT=F7.2 UPHRFMT=F7.2;
37 TITLE " "
      "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
      " "
      "PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE" " ";
38 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

```

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read	:	168	Weighted count:	168
Observations in subpopulation	:	42	Weighted count:	42
Observations used in the analysis	:	42	Weighted count:	42
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	20		

Maximum number of estimable parameters for the model is 4

Number of non-censored events: 42  
Number of censored events : 0  
WARNING: All values of the censoring variable COMPLETE are 1.  
SURVIVAL has converged to a solution in 5 iterations.

**Example 2 Results: Pre-Dosing Treatment Effect**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: PRE-DOSING TREATMENT EFFECT**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----

Independent Variables  
and Effects

	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
-----					
Day					
Treatment	-0.580044	0.190196	0.32	-3.05	0.0063
Placebo	0.000000	0.000000	.	.	.
Previous MI	-1.456612	0.502090	1.32	-2.90	0.0088
Previous Bypass Surgery	0.201999	0.503483	1.58	0.40	0.6925
Previous Propranolol Trt	-0.516548	0.575401	1.76	-0.90	0.3800

-----

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Date: 03-24-97                      Research Triangle Institute                      Page : 2  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: PRE-DOSING TREATMENT EFFECT**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
-----						
OVERALL MODEL	4	3.18	13.89	9.33	0.0077	0.0298
SUDTRT	1	1.00	9.30	9.30	0.0023	0.0024
MI	1	1.00	8.42	8.42	0.0037	0.0038
CAB	1	1.00	0.16	0.16	0.6883	0.6884
PP	1	1.00	0.81	0.81	0.3693	0.3696

-----

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: Pre-Dosing Treatment Effect**

Date: 03-27-97

Research Triangle Institute

Page : 3

Time: 08:50:19

The SURVIVAL Procedure

Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

**For Subpopulation: PRE-DOSING TREATMENT EFFECT**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

```

-----
Independent Variables          Lower   Upper
and Effects                   Hazards 95%   95%
                               Ratio   Limit Limit
-----

```

Day

```

Treatment          0.56   0.38   0.83
Placebo            1.00   1.00   1.00
Previous MI        0.23   0.08   0.66
Previous Bypass Surgery 1.22   0.43   3.50
Previous Propranolol Trt 0.60   0.18   1.98
-----

```

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used

CPU time : 2.8 seconds

Elapsed time : 3 seconds

Virtual memory : 1.06 MB

**Example 2 Results: 1-Hour Post-Dosing Treatment Effect**

```

39 PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
40 NEST _ONE_ PATIENT;
41 WEIGHT _ONE_;
42 SUBPOPN HOURS = 2 / NAME = "TREATMENT EFFECT @ 1 HR. POST-DOSING";
43 SUBGROUP SUDTRT;
44 LEVELS 2;
45 EVENT COMPLETE;
46 MODEL EXTIME = SUDTRT MI CAB PP;
47 TEST WALDCHI SATADJCHI;
48 SETENV COLSPACE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
49 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" HR LOWHR UPHR
        DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)"
        SATADCHI=" CHI-SQ (SAT)"
        WALDCHP=" P-VALUE (WALDC)"
        SATADCHP=" P-VALUE (SAT)"
        /DFMT=F7.0 BETAfmt=F10.6 SEBETAfmt=F10.6 T_BETAfmt=F8.2 WALDCHPfmt=F8.4
        P_BETAfmt=F8.4 SATADCHPfmt=F8.4 DEFTfmt=F6.2
        HRFMT=F7.2 LOWHRfmt=F7.2 UPHRfmt=F7.2;
50 TITLE " "
        "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
        " "
        "PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE" " ";
51 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read      :    168    Weighted count:    168
Observations in subpopulation   :     42    Weighted count:     42
Observations used in the analysis :     42    Weighted count:     42
Observations with missing values :      0    Weighted count:      0
Denominator degrees of freedom  :     20

Maximum number of estimable parameters for the model is 4

Number of non-censored events:   35
Number of censored events      :    7

SURVIVAL has converged to a solution in 5 iterations.

```

**Example 2 Results: 1-Hour Post-Dosing Treatment Effect**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

**For Subpopulation: TREATMENT EFFECT @ 1 HR. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----  
 Independent Variables  
 and Effects

	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
Day					
Treatment	-1.276868	0.290823	0.57	-4.39	0.0003
Placebo	0.000000	0.000000	.	.	.
Previous MI	-0.955064	0.437032	1.19	-2.19	0.0409
Previous Bypass Surgery	1.160058	0.443047	1.06	2.62	0.0165
Previous Propranolol Trt	-0.415035	0.436418	0.87	-0.95	0.3530

-----  
 Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Date: 03-24-97                      Research Triangle Institute                      Page : 2  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris

**For Subpopulation: TREATMENT EFFECT @ 1 HR. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
OVERALL MODEL	4	3.13	28.50	17.68	0.0000	0.0006
SUDTRT	1	1.00	19.28	19.28	0.0000	0.0000
MI	1	1.00	4.78	4.78	0.0289	0.0291
CAB	1	1.00	6.86	6.86	0.0088	0.0090
PP	1	1.00	0.90	0.90	0.3416	0.3418

-----  
 Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: 1-Hour Post-Dosing Treatment Effect**

Date: 03-24-97                      Research Triangle Institute                      Page : 3  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 1 HR. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

```

-----
Independent Variables          Lower   Upper
and Effects                    Hazards 95%  95%
                               Ratio   Limit Limit
-----
Day
Treatment                      0.28   0.15   0.51
Placebo                        1.00   1.00   1.00
Previous MI                     0.38   0.15   0.96
Previous Bypass Surgery        3.19   1.27   8.04
Previous Propranolol Trt      0.66   0.27   1.64
-----
    
```

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used  
 CPU time : 2.85 seconds  
 Elapsed time : 3 seconds  
 Virtual memory : 1.06 MB

## Example 2 Results: 3-Hours Post-Dosing Treatment Effect

```

52 PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
53 NEST _ONE_ PATIENT;
54 WEIGHT _ONE_;
55 SUBPOPN HOURS = 3 / NAME = "TREATMENT EFFECT @ 3 HRS. POST-DOSING";
56 SUBGROUP SUDTRT;
57 LEVELS 2;
58 EVENT COMPLETE;
59 MODEL EXTIME = SUDTRT MI CAB PP;
60 TEST WALDCHI SATADJCHI;
61 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
62 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" HR LOWHR UPHR
        DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)"
        SATADCHI=" CHI-SQ (SAT)"
        WALDCHP=" P-VALUE (WALDC)"
        SATADCHP=" P-VALUE (SAT)"
        /DFFMT=F7.0 BETAFMT=F10.6 SEBETAFMT=F10.6 T_BETAFMT=F8.2 WALDCHPFMT=F8.4
        P_BETAFMT=F8.4 SATADCHPFMT=F8.4 DEFTFMT=F6.2 HRFMT=F7.2 LOWHRFMT=F7.2
        UPHRFMT=F7.2;
63 TITLE " "
        "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
        " "
        "PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE" " ";
64 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

```

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read	:	168	Weighted count:	168
Observations in subpopulation	:	42	Weighted count:	42
Observations used in the analysis	:	42	Weighted count:	42
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	20		

Maximum number of estimable parameters for the model is 4

Number of non-censored events:	38
Number of censored events	: 4

SURVIVAL has converged to a solution in 5 iterations.

**Example 2 Results: 3-Hours Post-Dosing Treatment Effect**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 3 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----

Independent Variables and Effects	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
Day					
Treatment	-1.068183	0.250056	0.46	-4.27	0.0004
Placebo	0.000000	0.000000	.	.	.
Previous MI	-1.246678	0.400267	0.94	-3.11	0.0055
Previous Bypass Surgery	0.994977	0.441095	1.16	2.26	0.0354
Previous Propranolol Trt	-0.626157	0.555806	1.52	-1.13	0.2733

-----

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Date: 03-24-97                      Research Triangle Institute                      Page : 2  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 3 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
OVERALL MODEL	4	3.38	27.53	18.53	0.0000	0.0006
SUDTRT	1	1.00	18.25	18.25	0.0000	0.0000
MI	1	1.00	9.70	9.70	0.0018	0.0019
CAB	1	1.00	5.09	5.09	0.0241	0.0243
PP	1	1.00	1.27	1.27	0.2599	0.2602

-----

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)



**Example 2 Results: 3-Hours Post-Dosing Treatment Effect**

Date: 03-24-97  
Time: 08:50:19

Research Triangle Institute  
The SURVIVAL Procedure

Page : 3  
Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 3 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

```
-----
Independent Variables          Lower   Upper
and Effects                   Hazards 95%   95%
                               Ratio   Limit  Limit
-----
```

```
Day
  Treatment          0.34   0.20   0.58
  Placebo            1.00   1.00   1.00
  Previous MI        0.29   0.12   0.66
  Previous Bypass Surgery 2.70   1.08   6.79
  Previous Propranolol Trt 0.53   0.17   1.70
-----
```

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used

CPU time : 2.80 seconds  
Elapsed time : 3 seconds  
Virtual memory : 1.06 MB

## Example 2 Results: 5-Hours Post-Dosing Treatment Effect

```

65 PROC SURVIVAL DATA="EXERCISE" FILETYPE=SAS;
66 NEST _ONE_ PATIENT;
67 WEIGHT _ONE_;
68 SUBPOPN HOURS = 4 / NAME = "TREATMENT EFFECT @ 5 HRS. POST-DOSING";
69 SUBGROUP SUDTRT;
70 LEVELS 2;
71 EVENT COMPLETE;
72 MODEL EXTIME = SUDTRT MI CAB PP;
73 TEST WALDCHI SATADJCHI;
74 SETENV COLSPCE=1 LABWIDTH=22 COLWIDTH=8 LINESIZE=78 PAGESIZE=60;
75 PRINT BETA="BETA" SEBETA="STDERR" DEFT="DEFF" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" HR LOWHR UPHR
        DF="DF" SATADJDF="ADJ DF"
        WALDCHI=" CHI-SQ (WALD)"
        SATADCHI=" CHI-SQ (SAT)"
        WALDCHP=" P-VALUE (WALDC)"
        SATADCHP=" P-VALUE (SAT)"
        /DFFMT=F7.0 BETAFMT=F10.6 SEBETAFMT=F10.6 T_BETAFMT=F8.2 WALDCHPFMT=F8.4
        P_BETAFMT=F8.4 SATADCHPFMT=F8.4 DEFTFMT=F6.2
        HRFMT=F7.2 LOWHRFMT=F7.2 UPHRFMT=F7.2;
76 TITLE " "
        "EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT"
        " "
        "PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE"
        " ";
77 FOOTNOTE "Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

```

Opened SAS data file C:\TERA\EXAMPLES\EXERCISE.SSD for reading.

Number of observations read	:	168	Weighted count:	168
Observations in subpopulation	:	42	Weighted count:	42
Observations used in the analysis	:	42	Weighted count:	42
Observations with missing values	:	0	Weighted count:	0

Denominator degrees of freedom : 20

Maximum number of estimable parameters for the model is 4

Number of non-censored events:	40
Number of censored events	: 2

SURVIVAL has converged to a solution in 5 iterations.

**Example 2 Results: 5-Hours Post-Dosing Treatment Effect**

Date: 03-24-97                      Research Triangle Institute                      Page : 1  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 5 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

-----  
 Independent Variables  
 and Effects

	BETA	STDERR	DEFF	T:BETA=0	P-VALUE
Day					
Treatment	-0.742211	0.222916	0.43	-3.33	0.0033
Placebo	0.000000	0.000000	.	.	.
Previous MI	-1.952236	0.504155	1.23	-3.87	0.0009
Previous Bypass Surgery	0.872620	0.409208	1.12	2.13	0.0456
Previous Propranolol Trt	-0.765749	0.636049	2.10	-1.20	0.2427

-----  
 Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Date: 03-24-97                      Research Triangle Institute                      Page : 2  
 Time: 08:50:19                      The SURVIVAL Procedure                      Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 5 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

Contrast	DF	ADJ DF	CHI-SQ (WALD)	CHI-SQ (SAT)	P-VALUE (WALDC)	P-VALUE (SAT)
OVERALL MODEL	4	3.00	20.05	14.62	0.0005	0.0023
SUDTRT	1	1.00	11.09	11.09	0.0009	0.0009
MI	1	1.00	14.99	14.99	0.0001	0.0001
CAB	1	1.00	4.55	4.55	0.0330	0.0332
PP	1	1.00	1.45	1.45	0.2286	0.2289

-----  
 Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

**Example 2 Results: 5-Hours Post-Dosing Treatment Effect**

Date: 03-24-97  
 Time: 08:50:19

Research Triangle Institute  
 The SURVIVAL Procedure

Page : 3  
 Table : 1

For response variable EXTIME: Exercise Time to Angina Pectoris  
**For Subpopulation: TREATMENT EFFECT @ 5 HRS. POST-DOSING**

EXERCISE TIME TO ANGINA PECTORIS (SECONDS): PLACEBO VS. TREATMENT

PROPORTIONAL HAZARDS REGRESSION USING CLUSTER SAMPLE TECHNIQUE

```

-----
Independent Variables      Lower      Upper
and Effects                Hazards    95%      95%
                           Ratio      Limit    Limit
-----
Day
  Treatment                 0.48      0.30     0.76
  Placebo                   1.00      1.00     1.00
Previous MI                 0.14      0.05     0.41
Previous Bypass Surgery    2.39      1.02     5.62
Previous Propranolol Trt   0.46      0.12     1.75
-----
    
```

Source: Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

SURVIVAL used  
 CPU time : 2.86 seconds  
 Elapsed time : 3 seconds  
 Virtual memory : 1.06 MB



## Cross-Over Clinical Trial With Ordinal Outcomes:

### Evaluation of a New Inhaler Device via a Cross-Over Clinical Trial

Qualitative responses in a cross-over clinical trial are often ordinal. Such responses might be, for example, relief, slight relief, or no relief in studies of painkiller effectiveness. Due to the nature of cross-over studies, repeated measurements on the same subject are likely to be correlated. The intra-subject correlation must be taken into account in order to make valid inferences about the treatment effect.

Data for this example are from a two-treatment two-period crossover study conducted by 3M Health Care Ltd (Ezzet and Whitehead, 1991) to compare the suitability of two inhalation devices (A and B) in patients who are currently using a standard inhaler device delivering salbutamol. The first sequence of patients were randomized to Device A for one week (period 1) followed by Device B for another week (period 2). The second sequence of patients received the treatments in the opposite order (Device B in period 1, Device A in period 2). Patients gave their assessment on clarity of leaflet instructions accompanying the devices, recorded on an ordinal scale of: 1 = *easy*, 2 = *clear only after re-reading*, 3 = *not very clear*, and 4 = *confusing*.

Variables in the regression models included:

TREATMENT: A or B

PERIOD: 1 or 2.

The accompanying output contains results from the following SUDAAN procedures:

- 1) PROC RECORDS - contents of the data set
- 2) PROC CROSSTAB - descriptive statistics: distribution of the 4-level ordinal outcome across treatment group
- 3) PROC MULTLOG - proportional odds and multinomial logit regression of treatment and period effects on leaflet clarity

## Structure of the Clarity Data

Patient	Period	Treatment	Y = Clarity
1	1	1 = New	1 = Easy
1	2	2 = Standard	1 = Easy
2	1	1	1 = Easy
2	2	2	2 = Rereading
3	1	2	3 = Not Clear
3	2	1	2 = Rereading
4	1	2	4 = Confusing
4	2	1	1 = Easy

**N** = 572 records on the file

(286 clusters, 2 records per cluster)

**Example 3 Results:****File Contents**

```
1  PROC RECORDS DATA="C:\\TERA\\GEEORD\\CROSS" FILETYPE=SAS
      CONTENTS COUNTREC NOPRINT;
```

SAS Record File C:\\TERA\\GEEORD\\CROSS.SSD

Variables

Name	Type	Format	Description
PERSON	Numeric	F15.3	PERSON
TREAT	Numeric	F15.3	TREAT
SEQUENCE	Numeric	F15.3	SEQUENCE
PERIOD	Numeric	F15.3	PERIOD
CLARITY	Numeric	F15.3	CLARITY

Codes and Labels for Variable TREAT:

Code	Label
1	Inhaler A
2	Inhaler B

Codes and Labels for Variable PERIOD:

Code	Label
1	1=AB
2	2=BA

Codes and Labels for Variable CLARITY:

Code	Label
1	Easy
2	Rereading
3	Not Clear
4	Confusing

Number of records on file : 572

RECORDS used

CPU time : 0.55 seconds  
 Elapsed time : 1 second  
 Virtual memory : 0.75 MB

There are 572 records (one record for each person and treatment occasion) on the SAS data set. The outcome of interest is CLARITY of leaflet instructions, coded 1=*easy*, 2=*rereading required*, 3=*not clear*, and 4=*confusing*. SUDAAN picks up the labels for dependent and independent variables from the user-defined LEVEL.DBS file.

In the proportional odds model, we will model the probability of increasing clarity across treatment group and period (1 vs. 2). In the multinomial logit model, we will model the probability of being in each of the first 3 levels of CLARITY vs. the last.



**The LEVEL.DBS file for Example 3:**

**Value labels for categorical variables:**

CLARITY	1	Easy
CLARITY	2	Rereading
CLARITY	3	Not Clear
CLARITY	4	Confusing
TREAT	1	Inhaler A
TREAT	2	Inhaler B
SEQUENCE	1	1=AB
SEQUENCE	2	2=BA

**Example 3 Results:**

```

2  PROC CROSSTAB DATA="C:\\TERA\\NCHS\\CROSS" FILETYPE=SAS;
3  NEST _ONE_ PERSON;
4  WEIGHT _ONE_;
5  SUBGROUP TREAT CLARITY;
6  LEVELS 2 4;
7  TABLES TREAT*CLARITY;
8  SETENV DECWIDTH=0 COLWIDTH=10 LABWIDTH=15 COLSPCE=2;
9  PRINT NSUM/STYLE=NCHS;
10 TITLE " "
        "FREQUENCY DISTRIBUTION FOR INHALER DEVICE CROSS-OVER STUDY"
        " " "Ezzett and Whitehead, 1991" " ";

Number of observations read      :    572    Weighted count :    572
Number of observations skipped  :         0
(WEIGHT variable nonpositive)
Denominator degrees of freedom :    285
    
```

```

Date: 03-18-97                Research Triangle Institute                Page : 1
Time: 11:18:22                The CROSSTAB Procedure                    Table : 1
    
```

FREQUENCY DISTRIBUTION FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

Sample Size

TREAT	CLARITY				
	Total	Easy	Rereading	Not Clear	Confusing
Total	572	358	189	17	8
Inhaler A	286	211	71	2	2
Inhaler B	286	147	118	15	6

The CROSSTAB procedure was used to obtain the *frequency distribution* of CLARITY across treatment. It appears that the Inhaler B leaflet is less easy to read than that for Inhaler A.

### **MULTILOG Programming Statements and Options**

The first set of MULTILOG programming statements fits the proportional odds model in SUDAAN PROC MULTILOG. The *DATA* option on the *PROC* statement specifies a SAS data set as input. Since there is no *DESIGN* option specified, SUDAAN is using the default *DESIGN=WR* (with-replacement) option for variance estimation.

We will fit the following types of models:

- 1) ***SEMETHOD=ZEGER* and *R=INDEPENDENT***  
Implements the GEE model-fitting technique under an independent “working” assumption and a robust variance estimator.
- 2) ***SEMETHOD=ZEGER* and *R=EXCHANGEABLE***  
Implements the GEE model-fitting technique under exchangeable “working” correlations and a robust variance estimator.
- 3) ***SEMETHOD=MODEL* and *R=EXCHANGEABLE***  
We compare the results using the robust variance estimator (*SEMETHOD=ZEGER*) to the model-based, or naive, variance assumption (*SEMETHOD=MODEL*). When *R=exchangeable* is specified in conjunction with *SEMETHOD=MODEL*, variances are then computed as if the exchangeable “working” correlation assumption were correct.

The *NEST* statement indicates that PERSON is the cluster variable. The *WEIGHT* statement indicates equal sampling weights of 1.0 for each person and measurement occasion.

In MULTILOG, the *SUBGROUP* statement contains the dependent variable and all covariates that are to be modelled as categorical covariates (with level values of 1,2,...,k), where the maximum number of levels (*K*) appears on the *LEVELS* statement.

The *MODEL* statement specifies the categorical dependent variable CLARITY on the left of the "=" sign (with levels 1, 2, 3, and 4), and regressors on the right. The *CUMLOGIT* (cumulative logit) link specifies the proportional odds model (the GENLOGIT link comes later in the output). The CUMLOGIT link will model the log-odds that CLARITY  $\leq k$ , where  $k=1, \dots, K-1$  (or the tendency for CLARITY to be less than confusing). The *GENLOGIT* link will model the log-odds that CLARITY= $k$  vs.  $K$  (or the log-odds that CLARITY is easy, requires re-reading, or not clear vs. confusing). The CUMLOGIT option produces common slopes but separate intercepts for each of the  $K-1 = 3$  cutpoints, while the GENLOGIT option produces a separate logit equation (intercepts and slopes) for each of the 3 cutpoints.

The *TEST* statement specifies that we want the Wald chi-square statistic to be the default for testing main effects, interactions, and user-defined contrasts. This statement is optional. If omitted, the Wald F statistic becomes the default. However, any default statistic can be overridden on the *PRINT* statement.

The *SETENV* and *PRINT* statements are both optional, and control the printing of results (which statistics get printed, as well as their labels, formats, and layout).

## MULTILOG Programming Statements for the Proportional Odds Model: CUMLOGIT Link

### GEE with Independent Working Correlations and Robust Variance Estimates

```

11  PROC MULTILOG DATA="C:\\TERA\\GEEORD\\CROSS" FILETYPE=SAS
      SEMETHOD=ZEGER  R=INDEPENDENT;

12  NEST _ONE_ PERSON;

13  WEIGHT _ONE_;

14  SUBGROUP CLARITY TREAT PERIOD;

15  LEVELS 4 2 2;

16  MODEL CLARITY = TREAT PERIOD / CUMLOGIT;

17  TEST WALDCHI;

18  SETENV LABWIDTH=28 MAXIND=4 LINESIZE=78 PAGESIZE=60 COLSPCE=2;

19  PRINT  BETA="BETA" SEBETA="STDERR" DEFT="DESIGN EFFECT"
      T_BETA="T:BETA=0" P_BETA="P-Value"/
      RISK=ALL TESTS=DEFAULT
      BETAFMT=F7.4 SEBETAFMT=F6.4 T_BETAFMT=F8.2 P_BETAFMT=F7.4
      DEFTFMT=F6.2 WALDCHIFMT=F6.2 WALDCHPFMT=F7.4
      ORFMT=F5.2 LOWORFMT=F6.2 UPORFMT=F6.2 DFFMT=F7.0;

20  TITLE " " "PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY"
      " " "Ezzett and Whitehead, 1991";

```

Opened SAS data file C:\TERA\GEEORD\CROSS.SSD for reading.

Independence parameters have converged in 3 iterations

Number of observations read	:	572	Weighted count:	572
Observations used in the analysis	:	572	Weighted count:	572
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	285		

Maximum number of estimable parameters for the model is 5

File C:\TERA\GEEORD\CROSS.SSD contains 286 Clusters

Maximum cluster size is 2 records

Minimum cluster size is 2 records

Sample and Population Counts for Response Variable CLARITY

Easy	:	Sample Count	358	Population Count	358
Rereading:		Sample Count	189	Population Count	189
Not Clear:		Sample Count	17	Population Count	17
Confusing:		Sample Count	8	Population Count	8

CLARITY is the outcome variable in the model, while TREAT and PERIOD are covariates. There are 572 records on the file, corresponding to 286 clusters, with a minimum and maximum cluster size of 2 (since this is a 2-period crossover design). There are no missing values in the data set and no SUBPOPN statement to subset the analysis, so all observations on the file are used in fitting the model. SUDAAN displays the frequency distribution of the response in the data and the number of iterations needed to estimate the regression coefficients.

**Example 3 Results:**

**Proportional Odds Model: CUMLOGIT Link  
GEE with Independent Working Correlations and Robust Variance Estimates**

```

Date: 03-18-97              Research Triangle Institute          Page : 1
Time: 11:18:22             The MULTLOG Procedure              Table : 1

Variance Estimation Method: Robust (Zeger-Liang, 1986)
Working Correlations: Independent
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

-----
Independent Variables and          DESIGN
Effects                            BETA  STDERR  EFFECT  T:BETA=0  P-Value
-----
CLARITY (cum-logit)
  Intercept 1: Easy                0.1110  0.1383   0.88     0.80     0.4229
  Intercept 2: Rereading           2.7656  0.2357   1.03    11.74     0.0000
  Intercept 3: Not Clear           3.9464  0.3638   0.95    10.85     0.0000
TREAT
  Inhaler A                        1.0137  0.1566   0.78     6.47     0.0000
  Inhaler B                        0.0000  0.0000   .         .         .
PERIOD
  1=AB                             -0.1512  0.1565   0.80    -0.97     0.3347
  2=BA                             0.0000  0.0000   .         .         .
-----

```

The *estimated regression coefficients* for the proportional odds model indicate that Inhaler A is significantly clearer in its leaflet instructions than Inhaler B ( $p=0.0000$ ,  $t$ -test). This is reflected in the positive regression coefficient estimate (1.0137) and in the estimated odds ratio on the next page (2.76). In other words, the odds of being  $\leq$  any response level  $k$  are increased almost 3-fold over Inhaler B. Note the 3 intercept terms in the model are non-decreasing because they are cumulative over the categories of the response (*i.e.*, intercept 1 = *easy*; 2 = *easy or rereading required*; 3 = *easy, rereading, or not clear*). The fitted proportional odds model is as follows:

$$\log \left[ \frac{\text{prob}(Y \leq k)}{\text{prob}(Y > k)} \right] = 0.11_{k=1} + 2.77_{k=2} + 3.95_{k=3} + 1.01 \cdot TREAT - 0.1512 \cdot PERIOD$$

where TREAT and PERIOD are converted to 0-1 indicator variables because of their appearance on the SUBGROUP statement.

Note the *design effect* of 0.78 for the treatment parameter. We expect design effects less than 1.0 for treatment parameters nested within the cluster, as occurs in many repeated measures designs. This indicates that an improvement in precision was obtained because of the cross-over design and that SUDAAN was able to recognize this gain.

**Example 3 Results:****Proportional Odds Model: CUMLOGIT Link  
GEE with Independent Working Correlations and Robust Variance Estimates**

Date: 03-18-97                      Research Triangle Institute                      Page : 2  
Time: 11:18:22                      The MULTLOG Procedure                      Table : 1

Variance Estimation Method: **Robust (Zeger-Liang, 1986)**  
Working Correlations: **Independent**  
Link Function: **Cumulative Logit**  
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	5	272.62	0.0000
MODEL MINUS INTERCEPT	2	42.13	0.0000
TREAT	1	41.88	0.0000
PERIOD	1	0.93	0.3338

**Example 3 Results:**

**Proportional Odds Model: CUMLOGIT Link  
GEE with Independent Working Correlations and Robust Variance Estimates**

```

Date: 03-18-97                      Research Triangle Institute          Page : 3
Time: 11:18:22                      The MULTILOG Procedure              Table : 1

Variance Estimation Method: Robust (Zeger-Liang, 1986)
Working Correlations: Independent
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

-----
Independent Variables and          Lower   Upper
Effects                            Odds    95%    95%
                                   Ratio   Limit  Limit
-----
CLARITY (cum-logit)
  Intercept 1: Easy                 1.12    0.85    1.47
  Intercept 2: Rereading            15.89    9.99   25.26
  Intercept 3: Not Clear            51.75   25.30  105.86
TREAT
  Inhaler A                         2.76    2.03    3.75
  Inhaler B                         1.00    1.00    1.00
PERIOD
  1=AB                               0.86    0.63    1.17
  2=BA                               1.00    1.00    1.00
-----

MULTILOG used
CPU time       : 4.44 seconds
Elapsed time   : 5 seconds
Virtual memory : 1.11 MB

```

This output contains the *main effects tests* for the proportional odds model, in addition to the *estimated odds ratios* and their *95% confidence limits*.



**Example 3 Results:****MULTILOG Programming Statements for the Proportional Odds Model:  
Exchangeable Correlations and Robust Variance Estimates**

```

31  PROC MULTILOG DATA="C:\\TERA\\GEEORD\\CROSS" FILETYPE=SAS
      SEMETHOD=ZEGER R=EXCHANGE;

32  NEST _ONE_ PERSON;

33  WEIGHT _ONE_;

34  SUBGROUP CLARITY TREAT PERIOD;

35  LEVELS  4      2      2;

36  MODEL CLARITY = TREAT PERIOD / CUMLOGIT;

37  TEST WALDCHI;

38  SETENV LABWIDTH=15 LINESIZE=78 PAGESIZE=60;

39  PRINT RHO / RHOFMT=F10.4;

40  SETENV LABWIDTH=28 MAXIND=4 LINESIZE=78 PAGESIZE=60 COLSPCE=2;

41  PRINT  BETA="BETA" SEBETA="STDERR" DEFT="DESIGN EFFECT"
      T_BETA="T:BETA=0" P_BETA="P-Value" /
      RISK=ALL TESTS=DEFAULT
      BETAFMT=F7.4 SEBETAFMT=F6.4 T_BETAFMT=F8.2 P_BETAFMT=F7.4
      DEFTFMT=F6.2 WALDCHIFMT=F6.2 WALDCHPFMT=F7.4
      ORFMT=F5.2 LOWORFMT=F6.2 UPORFMT=F6.2 DFFMT=F7.0;

42  TITLE " " "PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY"
      " " "Ezzett and Whitehead, 1991";

```

*Continued on next page...*

**Example 3 Results:****MULTILOG Programming Statements for the Proportional Odds Model:  
Exchangeable Correlations and Robust Variance Estimates***continued from previous page...*

```

Opened SAS data file C:\TERA\GEEORD\CROSS.SSD for reading.

Independence parameters have converged in 3 iterations

Step 1 parameters have converged in 5 iterations.

Number of observations read      :    572    Weighted count:    572
Observations used in the analysis :    572    Weighted count:    572
Observations with missing values :     0     Weighted count:     0
Denominator degrees of freedom  :    285

Maximum number of estimable parameters for the model is 5

File C:\TERA\GEEORD\CROSS.SSD contains 286 Clusters
Maximum cluster size is 2 records
Minimum cluster size is 2 records

Sample and Population Counts for Response Variable CLARITY
Easy      : Sample Count    358    Population Count    358
Rereading: Sample Count    189    Population Count    189
Not Clear: Sample Count     17    Population Count     17
Confusing: Sample Count     8     Population Count     8

```

In the above programming statements, we request *SEMETHOD=ZEGGER* and *R=exchangeable* to implement GEE under exchangeable working correlations. All other statements remain as previously for the proportional odds model (CUMLOGIT link). The starting parameter estimates, computed in the usual way under the naive assumption of independence, converged to a solution in 4 iterations. The Step 1 GEE estimates, which update the independence estimates with the estimated correlation structure, converged in 5 iterations.

**Example 3 Results:****Proportional Odds Model: CUMLOGIT Link  
GEE with Exchangeable Working Correlations and Robust Variance Estimates**

Date: 03-18-97                      Research Triangle Institute                      Page : 1  
Time: 11:18:22                      The MULTILOG Procedure                      Table : 1

Variance Estimation Method: **Robust (Zeger-Liang, 1986)**  
Working Correlations: **Exchangeable**  
Link Function: **Cumulative Logit**  
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

**Correlation Matrix**

```
-----
CLARITY          CLARITY
                  Easy    Rereading    Not Clear
-----
Easy              0.2156
Rereading         -0.1975         0.2069
Not Clear         -0.0564         -0.0168         0.1427
-----
```

The *estimated correlation structure* is contained in the above table. Note that for a 4-level response variable, a cluster size of 2, and an exchangeable correlation model, there are exactly 6 unique correlation estimates. SUDAAN prints the lower portion of the symmetric 3-by-3 matrix. These estimates indicate that the correlation between the “*Easy to Read*” categories on both treatments ( $Y_{i1s}, Y_{i1t}$ ) was 0.2156, and the correlation between the “*Rereading Required*” categories on both treatments ( $Y_{i2s}, Y_{i2t}$ ) was 0.2069. Therefore, the most frequently occurring pairs are identical outcomes. The smaller negative correlations indicate that crossing response categories from Inhaler A to B is not as likely as remaining in the same response category on each treatment.

**Example 3 Results:**

**Proportional Odds Model: CUMLOGIT Link  
GEE with Exchangeable Working Correlations and Robust Variance Estimates**

Date: 03-18-97	Research Triangle Institute	Page : 2		
Time: 11:18:22	The MULTLOG Procedure	Table : 1		
Variance Estimation Method: <b>Robust (Zeger-Liang, 1986)</b>				
Working Correlations: <b>Exchangeable</b>				
Link Function: <b>Cumulative Logit</b>				
Response variable CLARITY: CLARITY				
PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY				
Ezzett and Whitehead, 1991				
-----				
Independent Variables and Effects	BETA	STDERR	T:BETA=0	P-Value
-----				
CLARITY (cum-logit)				
Intercept 1: Easy	0.1085	0.1379	0.79	0.4320
Intercept 2: Rereading	2.7424	0.2344	11.70	0.0000
Intercept 3: Not Clear	3.9568	0.3639	10.87	0.0000
TREAT				
Inhaler A	1.0140	0.1562	6.49	0.0000
Inhaler B	0.0000	0.0000	.	.
PERIOD				
1=AB	-0.1531	0.1556	-0.98	0.3258
2=BA	0.0000	0.0000	.	.
-----				

This table contains the *regression coefficient estimates* under the exchangeable correlation structure. We see that the regression estimates are slightly larger and the variance estimates are slightly smaller compared to the independence working assumption shown previously. However, the results are qualitatively the same. Inhaler A is significantly clearer in its leaflet instructions than Inhaler B. Both working assumptions are valid no matter what the true correlation structure since SUDAAN is using the *robust variance estimates* (SEMETHOD=ZEGER) for computing variance and testing hypotheses.

**Example 3 Results:****Proportional Odds Model: CUMLOGIT Link  
GEE with Exchangeable Working Correlations and Robust Variance Estimates**

Date: 03-18-97                      Research Triangle Institute                      Page : 3  
Time: 11:18:22                      The MULTILOG Procedure                      Table : 1

Variance Estimation Method: **Robust (Zeger-Liang, 1986)**  
Working Correlations: **Exchangeable**  
Link Function: **Cumulative Logit**  
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	5	272.33	0.0000
MODEL MINUS INTERCEPT	2	42.39	0.0000
TREAT	1	42.16	0.0000
PERIOD	1	0.97	0.3250

This table summarizes the *main effects tests* under the exchangeable correlation “working” assumption. Again, these results are qualitatively similar to the “working” independence model with robust variance estimates.

**Example 3 Results:**

**Proportional Odds Model: CUMLOGIT Link  
GEE with Exchangeable Working Correlations and Robust Variance Estimates**

```

Date: 03-18-97                      Research Triangle Institute          Page : 4
Time: 11:18:22                      The MULTILOG Procedure              Table : 1

Variance Estimation Method: Robust (Zeger-Liang, 1986)
Working Correlations: Exchangeable
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

-----
CLARITY (cum-logit),
Independent Variables and          Lower   Upper
Effects                            Odds    95%    95%
                                   Ratio   Limit  Limit
-----
CLARITY (cum-logit)
Intercept 1: Easy                  1.11    0.85   1.46
Intercept 2: Rereading             15.52    9.79  24.62
Intercept 3: Not Clear             52.29   25.56 106.99
TREAT
Inhaler A                          2.76    2.03   3.75
Inhaler B                          1.00    1.00   1.00
PERIOD
1=AB                                0.86    0.63   1.17
2=BA                                1.00    1.00   1.00
-----

MULTILOG used
CPU time       : 11.91 seconds
Elapsed time   : 12 seconds
Virtual memory : 1.14 MB
    
```

These *odds ratios and 95% confidence limits* for the exchangeable “working” assumption are identical to the independence “working” model. Modelling the correlations under exchangeability did not significantly improve the efficiency of the parameter estimates in this example.

**Example 3 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

Below are results from the exchangeable correlation model using the *model-based* or *naive variance-covariance matrix* of the estimated regression coefficients. The model-based variance is the  $M_0^{-1}$  matrix, or the outside portion of the robust variance estimate:  $M_0^{-1} = [D'V^{-1}D]^{-1}$ , where  $D = \partial\pi_i / \partial\beta$  is the vector of first partial derivatives of the response probabilities  $\pi_i$  with respect to the regression coefficients  $\beta$ . In this case, the naive variance estimate is computed *as if the exchangeable "working" correlation assumption were correct*. Since this is close to truth for litter data, we will see that results are essentially the same as with the robust variance estimator. To obtain the model-based results, we specify **SEMETHOD=MODEL** on the PROC statement.

```

43  PROC MULTILOG DATA="C:\\TERA\\GEEORD\\CROSS" FILETYPE=SAS
      SEMETHOD=MODEL R=EXCHANGE;

44  NEST _ONE_ PERSON;

45  WEIGHT _ONE_;

46  SUBGROUP CLARITY TREAT PERIOD;

47  LEVELS 4 2 2;

48  MODEL CLARITY = TREAT PERIOD / CUMLOGIT;

49  TEST WALDCHI;

50  SETENV LABWIDTH=15 LINESIZE=78 PAGESIZE=60;

51  PRINT RHO / RHOFMT=F10.4;

52  SETENV LABWIDTH=28 MAXIND=4 LINESIZE=78 PAGESIZE=60 COLSPCE=2;

53  PRINT BETA="BETA" SEBETA="STDERR" DEFT="DESIGN EFFECT"
      T_BETA="T:BETA=0" P_BETA="P-Value" /
      RISK=ALL TESTS=DEFAULT
      BETAFMT=F7.4 SEBETAFMT=F6.4 T_BETAFMT=F8.2 P_BETAFMT=F7.4
      DEFTFMT=F6.2 WALDCHIFMT=F6.2 WALDCHPFMT=F7.4
      ORFMT=F5.2 LOWORFMT=F6.2 UPORFMT=F6.2 DFFMT=F7.0;

54  TITLE " " "PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY"
      " " "Ezzett and Whitehead, 1991" "Model-Based Variance Estimation";

```

...continued next page

**Example 3 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation***...continued from previous page*

```
Opened SAS data file C:\TERA\GEEORD\CROSS.SSD for reading.
```

```
Number of observations read      :    572    Weighted count:    572
Observations used in the analysis :    572    Weighted count:    572
Observations with missing values :     0     Weighted count:     0
Denominator degrees of freedom  :    285
```

```
Maximum number of estimable parameters for the model is 5
```

```
File C:\TERA\GEEORD\CROSS.SSD contains 286 Clusters
```

```
Maximum cluster size is 2 records
Minimum cluster size is 2 records
```

```
Independence parameters have converged in 3 iterations
```

```
Step 1 parameters have converged in 5 iterations.
```

```
Sample and Population Counts for Response Variable CLARITY
```

Easy	: Sample Count	358	Population Count	358
Rereading:	Sample Count	189	Population Count	189
Not Clear:	Sample Count	17	Population Count	17
Confusing:	Sample Count	8	Population Count	8



**Example 3 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

```
Date: 03-18-97          Research Triangle Institute          Page : 1
Time: 11:18:22        The MULTILOG Procedure              Table : 1
```

```
Variance Estimation Method: Model-Based (Naive)
Working Correlations: Exchangeable
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY
```

```
PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY
```

```
Ezzett and Whitehead, 1991
Model-Based Variance Estimation
```

**Correlation Matrix**

```
-----
CLARITY          CLARITY
                  Easy    Rereading    Not Clear
-----
Easy              0.2156
Rereading         -0.1975    0.2069
Not Clear         -0.0564    -0.0168    0.1427
-----
```

The *estimated correlation matrix* under exchangeability is unaffected by the choice of robust vs. model-based variance estimation.

**Example 3 Results:**

**GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

```

Date: 03-18-97          Research Triangle Institute          Page : 2
Time: 11:18:22         The MULTLOG Procedure          Table : 1

Variance Estimation Method: Model-Based (Naive)
Working Correlations: Exchangeable
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991
Model-Based Variance Estimation

-----
Independent Variables and
Effects                BETA  STDERR  T:BETA=0  P-Value
-----
CLARITY (cum-logit)
  Intercept 1: Easy          0.1085  0.1415    0.77    0.4437
  Intercept 2: Rereading     2.7424  0.2363   11.61    0.0000
  Intercept 3: Not Clear     3.9568  0.3510   11.27    0.0000
TREAT
  Inhaler A                  1.0140  0.1577    6.43    0.0000
  Inhaler B                   0.0000  0.0000    .        .
PERIOD
  1=AB                       -0.1531  0.1555   -0.98    0.3256
  2=BA                       0.0000  0.0000    .        .
-----

```

Here we have the *estimated regression coefficients* computed under exchangeability and the estimated standard errors as if the exchangeable working assumption were correct. The standard errors are roughly the same as with the robust variance estimator for these data, indicating that the exchangeable correlation assumption is close to truth.

**Example 3 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

Date: 03-18-97                      Research Triangle Institute                      Page : 3  
 Time: 11:18:22                      The MULTILOG Procedure                      Table : 1

Variance Estimation Method: **Model-Based (Naive)**  
 Working Correlations: **Exchangeable**  
 Link Function: **Cumulative Logit**  
 Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991  
 Model-Based Variance Estimation

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	5	271.76	0.0000
MODEL MINUS INTERCEPT	2	42.17	0.0000
TREAT	1	41.33	0.0000
PERIOD	1	0.97	0.3248

Here we have the *main effects tests* computed under exchangeability, using the model-based variance approach. Results are essentially the same as with the robust variance estimator.

**Example 3 Results:**

**GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

```

Date: 03-18-97                Research Triangle Institute                Page : 4
Time: 11:18:22                The MULTILOG Procedure                Table : 1

Variance Estimation Method: Model-Based (Naive)
Working Correlations: Exchangeable
Link Function: Cumulative Logit
Response variable CLARITY: CLARITY

PROPORTIONAL ODDS MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991
Model-Based Variance Estimation

-----
CLARITY (cum-logit),
  Independent Variables and          Lower   Upper
  Effects                            Odds    95%    95%
                                   Ratio   Limit  Limit
-----
CLARITY (cum-logit)
  Intercept 1: Easy                   1.11    0.84   1.47
  Intercept 2: Rereading               15.52    9.75  24.71
  Intercept 3: Not Clear               52.29   26.21 104.31
TREAT
  Inhaler A                           2.76    2.02   3.76
  Inhaler B                           1.00    1.00   1.00
PERIOD
  1=AB                                 0.86    0.63   1.17
  2=BA                                 1.00    1.00   1.00
-----

MULTILOG used
CPU time      : 10.60 seconds
Elapsed time  : 11 seconds
Virtual memory : 1.14 MB
    
```

Here we have the *estimated odds ratios and their 95% confidence limits* computed under exchangeability, using the model-based variance approach. Odds ratios are unaffected by the choice of robust vs. model-based variance estimates, and estimated confidence limits are essentially the same as with the robust variance estimator.

**Example 3 Results:****MULTILOG Programming Statements for the Multinomial Logit Model:  
GENLOGIT Link**

```

55  PROC MULTILOG DATA="C:\\TERA\\GEEORD\\CROSS" FILETYPE=SAS
      SEMETHOD=ZEGER R=INDEPENDENT;

56  NEST _ONE_ PERSON;

57  WEIGHT _ONE_;

58  SUBGROUP CLARITY TREAT PERIOD;

59  LEVELS  4      2      2;

60  MODEL CLARITY = TREAT PERIOD / GENLOGIT;

61  TEST WALDCHI;

62  SETENV LABWIDTH=15 COLWIDTH=10 DECWIDTH=4 MAXIND=4 LINESIZE=78
      PAGESIZE=60;

63  PRINT BETA="BETA" SEBETA="STDERR" DEFT="DESIGN EFFECT"
      T_BETA="T:BETA=0" P_BETA="P-Value"/
      RISK=ALL TESTS=DEFAULT T_BETAFMT=F8.2 WALDCHIFMT=F6.2
      ORFMT=F10.2 LOWORFMT=F10.2 UPORFMT=F10.2 DFFMT=F7.0;

64  TITLE " "
      "GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY"
      " " "Ezzett and Whitehead, 1991";

```

Opened SAS data file C:\TERA\GEEORD\CROSS.SSD for reading.

Independence parameters have converged in 5 iterations

Number of observations read	:	572	Weighted count:	572
Observations used in the analysis	:	572	Weighted count:	572
Observations with missing values	:	0	Weighted count:	0
Denominator degrees of freedom	:	285		

Maximum number of estimable parameters for the model is 9

File C:\TERA\GEEORD\CROSS.SSD contains 286 Clusters

Maximum cluster size is 2 records

Minimum cluster size is 2 records

Sample and Population Counts for Response Variable CLARITY

Easy	:	Sample Count	358	Population Count	358
Rereading:		Sample Count	189	Population Count	189
Not Clear:		Sample Count	17	Population Count	17
Confusing:		Sample Count	8	Population Count	8

The **GENLOGIT** option invokes the multinomial logit model based on the generalized logit link function. All other options remain the same as for the proportional odds model.

**Example 3 Results:**

**Multinomial Logit Model: GENLOGIT Link**

**GEE with Independent Working Correlations and Robust Variance Estimates**

Date: 03-18-97	Research Triangle Institute	Page : 1
Time: 11:18:22	The MULTLOG Procedure	Table : 1
Variance Estimation Method: <b>Robust (Zeger-Liang, 1986)</b>		
Working Correlations: <b>Independent</b>		
Link Function: <b>Generalized Logit</b>		
Response variable CLARITY: CLARITY		
GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY		
Ezzett and Whitehead, 1991		
CLARITY (log-odds)		Independent Variables and Effects
		Intercept   TREAT =   TREAT =
		Inhaler A   Inhaler B
Easy vs Confusing	BETA	3.5099   1.4615   0.0000
	STDERR	0.6858   0.8254   0.0000
	DESIGN EFFECT	1.2232   1.0037   .
	T:BETA=0	5.12   1.77   .
	P-Value	0.0000   0.0777   .
Rereading vs Confusing	BETA	3.2510   0.5919   0.0000
	STDERR	0.6908   0.8311   0.0000
	DESIGN EFFECT	1.2281   1.0015   .
	T:BETA=0	4.71   0.71   .
	P-Value	0.0000   0.4769   .
Not Clear vs Confusing	BETA	1.0089   -0.9159   0.0000
	STDERR	0.7634   1.1557   0.0000
	DESIGN EFFECT	1.1092   1.0830   .
	T:BETA=0	1.32   -0.79   .
	P-Value	0.1874   0.4287   .

-continued-

**Example 3 Results:****Multinomial Logit Model: GENLOGIT Link****GEE with Independent Working Correlations and Robust Variance Estimates**

Date: 03-18-97	Research Triangle Institute	Page : 2
Time: 11:18:22	The MULTLOG Procedure	Table : 1
Variance Estimation Method: <b>Robust (Zeger-Liang, 1986)</b>		
Working Correlations: <b>Independent</b>		
Link Function: <b>Generalized Logit</b>		
Response variable CLARITY: CLARITY		
GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY		
Ezzett and Whitehead, 1991		
-----		
CLARITY (log-odds)		Independent Variables and Effects
		PERIOD = 1=AB      PERIOD = 2=BA
-----		
Easy vs Confusing	BETA	-0.5593      0.0000
	STDERR	0.7401      0.0000
	DESIGN EFFECT	0.9995      .
	T:BETA=0	-0.76      .
	P-Value	0.4505      .
-----		
Rereading vs Confusing	BETA	-0.4805      0.0000
	STDERR	0.7456      0.0000
	DESIGN EFFECT	1.0016      .
	T:BETA=0	-0.64      .
	P-Value	0.5198      .
-----		
Not Clear vs Confusing	BETA	-0.1527      0.0000
	STDERR	0.8992      0.0000
	DESIGN EFFECT	1.0411      .
	T:BETA=0	-0.17      .
	P-Value	0.8653      .
-----		

In this and the previous box we have the *estimated regression coefficient vector* and related statistics. Note that we now have 3 separate logit equations. So, for example, the logit equation for CLARITY = *Easy* vs. CLARITY = *Confusing* is as follows:

$$\log \left[ \frac{\hat{\pi}_{EASY}}{\hat{\pi}_{CONFUSING}} \right] = 3.51 + 1.46 \cdot TREAT - 0.5593 \cdot PERIOD$$

where TREAT and PERIOD are converted to 0-1 indicator variables because of their appearance on the SUBGROUP statement. The treatment effect appears to be largest when comparing the *Easy* vs. *Confusing* categories.

**Example 3 Results:**

**Multinomial Logit Model: GENLOGIT Link**

**GEE with Independent Working Correlations and Robust Variance Estimates**

```

Date: 03-18-97                Research Triangle Institute                Page : 3
Time: 11:18:22                The MULTLOG Procedure                    Table : 1

Variance Estimation Method: Robust (Zeger-Liang, 1986)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable CLARITY: CLARITY

GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

-----
Contrast          Degrees      Wald      P-value
                  of          ChiSq     Wald ChiSq
                  Freedom
-----
OVERALL MODEL          9    233.14    0.0000
MODEL MINUS INTERCEPT 6    45.27    0.0000
INTERCEPT            .      .        .
TREAT                  3    39.88    0.0000
PERIOD                 3     1.44    0.6962
-----

```

The *treatment effect* (now with 3 degrees of freedom in the multinomial logit model) is statistically significant, as in the proportional odds model.



**Example 3 Results:****Multinomial Logit Model: GENLOGIT Link****GEE with Independent Working Correlations and Robust Variance Estimates**

Date: 03-18-97                      Research Triangle Institute                      Page : 4  
 Time: 11:18:22                      The MULTILOG Procedure                      Table : 1

Variance Estimation Method: **Robust (Zeger-Liang, 1986)**  
 Working Correlations: **Independent**  
 Link Function: **Generalized Logit**  
 Response variable CLARITY: CLARITY

GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY

Ezzett and Whitehead, 1991

CLARITY (log-odds)		Independent Variables and Effects		
		Intercept	TREAT = Inhaler A	TREAT = Inhaler B
Easy vs Confusing	Odds Ratio	33.44	4.31	1.00
	Lower 95% Limit	8.68	0.85	1.00
	Upper 95% Limit	128.91	21.87	1.00
Rereading vs Confusing	Odds Ratio	25.82	1.81	1.00
	Lower 95% Limit	6.63	0.35	1.00
	Upper 95% Limit	100.47	9.27	1.00
Not Clear vs Confusing	Odds Ratio	2.74	0.40	1.00
	Lower 95% Limit	0.61	0.04	1.00
	Upper 95% Limit	12.31	3.89	1.00

- continued -

The *estimated odds* of being in the *EASY* vs. *CONFUSING* categories is increased over 4-fold for Inhaler A vs. B.

**Example 3 Results:**

**Multinomial Logit Model: GENLOGIT Link**

**GEE with Independent Working Correlations and Robust Variance Estimates**

Date: 03-18-97	Research Triangle Institute	Page : 5
Time: 11:18:22	The MULTILOG Procedure	Table : 1
Variance Estimation Method: <b>Robust (Zeger-Liang, 1986)</b>		
Working Correlations: <b>Independent</b>		
Link Function: <b>Generalized Logit</b>		
Response variable CLARITY: CLARITY		
GENERALIZED LOGIT MODEL FOR INHALER DEVICE CROSS-OVER STUDY		
Ezzett and Whitehead, 1991		
-----		
CLARITY (log-odds)		Independent Variables and Effects PERIOD = 1=AB      PERIOD = 2=BA
-----		
Easy vs Confusing	Odds Ratio	0.57      1.00
	Lower 95% Limit	0.13      1.00
	Upper 95% Limit	2.45      1.00
-----		
Rereading vs Confusing	Odds Ratio	0.62      1.00
	Lower 95% Limit	0.14      1.00
	Upper 95% Limit	2.68      1.00
-----		
Not Clear vs Confusing	Odds Ratio	0.86      1.00
	Lower 95% Limit	0.15      1.00
	Upper 95% Limit	5.03      1.00
-----		
MULTILOG used		
CPU time	: 5.50 seconds	
Elapsed time	: 6 seconds	
Virtual memory	: 1.18 MB	



## Example 4.

### Teratology Experiment: Clustered Continuous Data

This example demonstrates the GEE (Zeger and Liang, 1986; Liang and Zeger, 1986) and Jackknife model-fitting techniques in the context of a pre-clinical teratology experiment.

The typical teratology screening experiment involves administration of a compound to pregnant dams of a given animal species, followed by evaluation of the fetuses just prior to the end of gestation for fetal body weight, fetal death, and various types of fetal malformations. The experimental groups consist of a control group and anywhere from 2 to 5 exposed groups, representing increasing dosages of the compound under test. The data for this example represent *fetal body weight* in rats after administration of boric acid (0, 0.025, 0.05, 0.075, 0.1, or 0.2% in feed) to the dam daily during gestation. There were a total of 164 litters in the experiment (average of 27 litters per group) and anywhere from 2-14 fetuses per litter (1,302 fetuses total).

In this example, the observations on fetuses are clustered within litters. The design effect measures the inflation (or deflation) in variance of a sample statistic due to intraclass correlation beyond that expected if the data were independent. It is estimated as the ratio of the cluster sample variance obtained through GEE or Jackknife vs. independence. Design effects in this study ranged from 3 to 6, reflecting high intralitter correlations.

To implement the GEE methods in SUDAAN, we first estimated the model parameters via *ordinary least squares (OLS)* with a *robust variance estimate*. This is the GEE linear model with independent “working” correlations (which we refer to as *GEE-independent*). The Wald chi-square test was used to evaluate the null hypothesis of no dose-related effect.

For comparison, the same linear model was also fit using:

- 1) GEE linear regression under *exchangeable intralitter correlations*,
- 2) linear regression with *Jackknife variance estimation*, and
- 3) *ordinary least squares with no variance correction*

Results for GEE-exchangeable, GEE-independent, and the Jackknife approach were essentially the same. For comparing the high dose to control in the linear model, the GEE-exchangeable approach yielded a Z-statistic of -7.40, compared to a GEE-independent Z-statistic of -8.11 and a Jackknife Z-statistic of -7.83. The Z-statistic which ignores clustering altogether was misleadingly high, -14.54. The observed design effect for the high dose vs. control regression parameter was over 3.0 for these data, reflecting substantial intralitter correlations (estimated to be 0.5056).

Naively ignoring the clustering of the design in both parameter and variance estimation yields significant reductions ( $p < 0.05$ ) in body weights in dose groups as low as .05% (2nd lowest dose group) and marginally significant reductions ( $p = 0.06$ ) in the lowest dose group, while all three alternative approaches (GEE-independent, GEE-exchangeable, and Jackknife) only detect significant reductions in the two highest dose groups. Therefore, if we ignore intraclass correlations for cluster-level covariates (dose group in this study), we run the risk of detecting false-positive results.

### Structure of the Fetal Body Weight Data

<b>Dose Group</b> 1 = Control 6 = High Dose	<b>Litter ID</b>	<b>Fetus ID</b>	<b>Y = fetal body weight (gms)</b>
1	1	1	3.56
1	1	2	3.20
1	1	3	4.14
1	2	1	2.99
1	2	2	3.21
6	10	1	2.11
6	10	2	3.43
6	20	1	4.88
6	20	2	3.10
6	30	1	2.67

$N = 1,302$  records on the file  
(1,302 fetuses clustered within 164 litters)

Figure 1

Linear Regression for the Boric Acid Data

Exposed vs. Control Group Contrasts

Contrast	Model-Fitting Method	$\beta$	S.E.	Z	P-Value
<b>0.025% Vs. Control</b>	GEE (indep)	-0.0611	0.0596	-1.03	0.3067
	GEE (exch corr)	-0.0509	0.0622	-0.82	0.4148
	Jackknife	-0.0611	0.0616	-0.99	0.3227
	Independence	-0.0611	0.0332	-1.84	0.0676
<b>0.050% Vs. Control</b>	GEE (indep)	-0.0789	0.0724	-1.09	0.2777
	GEE (exch corr)	-0.0656	0.0738	-0.89	0.3753
	Jackknife	-0.0789	0.0752	-1.05	0.2955
	Independence	-0.0789	0.0330	-2.39	0.0180
<b>0.075% Vs. Control</b>	GEE (indep)	-0.1219	0.0740	-1.65	0.1016
	GEE (exch corr)	-0.1363	0.0790	-1.73	0.0864
	Jackknife	-0.1219	0.0765	-1.59	0.1128
	Independence	-0.1219	0.0327	-3.73	0.0003
<b>0.10% Vs. Control</b>	GEE (indep)	-0.2062	0.0627	-3.29	0.0012
	GEE (exch corr)	-0.2409	0.0680	-3.54	0.0005
	Jackknife	-0.2062	0.0648	-3.18	0.0018
	Independence	-0.2062	0.0323	-6.39	0.0000
<b>0.20% Vs. Control</b>	GEE (indep)	-0.4883	0.0602	-8.11	0.0000
	GEE (exch corr)	-0.4822	0.0651	-7.40	0.0000
	Jackknife	-0.4883	0.0624	-7.83	0.0000
	Independence	-0.4883	0.0336	-14.54	0.0000

GEE (independent): SUDAAN REGRESS Procedure  
 GEE (exchangeable): SUDAAN REGRESS Procedure  
 Jackknife: SUDAAN REGRESS Procedure  
 Independence: Standard Packages (e.g., SAS GLM)

All procedures in this table (except Independence) use a robust variance estimator.

**Example 4 Results:****Descriptive Statistics**

Here we present the average fetal body weight in each dose group, along with their estimated standard errors (using a robust variance estimate to adjust for clustering) and design effects. These design effects were in the range of 3-6, reflecting more than a tripling in the variance of the estimated means under the clustered design. In the program code, the DAMID variable represents the cluster on the NEST statement. Results show that fetal body weight is reduced in the two highest dose groups compared to control.

```
1  PROC DESCRIPT DATA="c:\\tera\\examples\\boric" FILETYPE=SAS NOMARG DESIGN=WR;
2  NEST _ONE_ DAMID;
3  WEIGHT _ONE_;
4  VAR BW;
5  SUBGROUP DOSEGRP;
6  LEVELS 6;
7  SETENV LABWIDTH=30 COLWIDTH=8 LINESIZE=78 DECWIDTH=4 PAGESIZE=60;
8  PRINT NSUM="SAMPLE SIZE"
      MEAN="MEAN"
      SEMEAN="SE"
      DEFFMEAN="DESIGN EFFECT" / STYLE=NCHS NSUMFMT=F6.0;
9  TITLE "Fetal Body Weight in a Teratology Study";

Opened SAS data file c:\\tera\\examples\\boric.SSD for reading.

Number of observations read      :   1302      Weighted count :   1302
Denominator degrees of freedom  :    163
```

## Example 4 Results

### Descriptive Statistics

Date: 03-18-97                      Research Triangle Institute                      Page : 1  
 Time: 12:51:34                      The DESCRIPT Procedure                      Table : 1

Fetal Body Weight in a Teratology Study

```
-----
```

Variable				
DOSEGRP	SAMPLE	MEAN	SE	DESIGN
	SIZE			EFFECT
<b>Fetal Body Weight</b>				
Control	217	3.4979	0.0406	3.2295
0.025%	210	3.4367	0.0436	3.5601
0.05%	215	3.4190	0.0600	5.4348
0.075%	223	3.3760	0.0619	6.0734
0.1%	236	3.2916	0.0477	4.8155
0.2%	201	3.0096	0.0444	4.7528

```
-----
```

DESCRIPT used  
 CPU time : 2.47 seconds  
 Elapsed time : 3 seconds  
 Virtual memory : 0.83 MB



**Example 4 Results:****Regression Modelling**

This first REGRESS procedure fits a *linear regression model* to the fetal body weights, with dose group (6-level categorical variable, from 1=*control* to 6=*high dose*) as the only predictor. We use the *REFLEVEL statement* to change the reference level for DOSEGRP from the default last level (*high dose*) to the first (*control*). Now the regression parameters will be comparing each treatment group to the control. We specify *R=INDEPENDENT* to estimate the model via GEE under independent “working” correlations. The DAMID variable remains as the cluster on the NEST statement. We also request the “Least Squares Means” for DOSEGRP. In this case, the least squares means will be equal to the raw means in each dose group, since DOSEGRP is the only covariate in the model. SUDAAN notifies the user that there are 164 clusters, 1302 fetuses (records on the file), and a min and max cluster size of 2 and 14, respectively.

```

10 PROC REGRESS DATA="c:\\tera\\examples\\boric" FILETYPE=SAS R=INDEPENDENT;

11 NEST _ONE_ DAMID;

12 WEIGHT _ONE_;

13 RELEVEL DOSEGRP = 1;

14 SUBGROUP DOSEGRP;

15 LEVELS 6;

16 MODEL BW = DOSEGRP;

17 LSMEANS DOSEGRP;

18 SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;

19 PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF WALDCHI WALDCHP / LSMEANS=ALL
      T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;

20 TITLE " "
      "Treatment Effect on Fetal Body Weight in a Teratology Experiment";

Opened SAS data file c:\tera\examples\boric.SSD for reading.
Number of observations read      :   1302      Weighted count:   1302
Observations used in the analysis :   1302      Weighted count:   1302
Observations with missing values :     0        Weighted count:     0
Denominator degrees of freedom  :    163

Maximum number of estimable parameters for the model is 6

File c:\tera\examples\boric.SSD contains 164 clusters
Maximum cluster size is 14 records
Minimum cluster size is 2 records
Weighted mean response is 3.341329

Multiple R-Square for the dependent variable BW: 0.171089

```

**Example 4 Results:**

**GEE Under Independent “Working” Correlations**

Below are the estimated regression coefficients under working independence. By default, SUDAAN uses the robust variance estimator, which appropriately corrects for intracluster correlation and yields valid results. For linear regression models, the robust variance estimator of Binder (1983) is equivalent to that of Zeger and Liang (1986)

```

Date: 03-18-97                Research Triangle Institute                Page : 1
Time: 12:51:34                The REGRESS Procedure                    Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Independent Variables and
Effects
                BETA      S.E.  DESIGN  T:BETA=0  P-VALUE
-----
Intercept                3.4979   0.0406   3.04     86.10    0.0000
DOSEGRP
  Control                0.0000   0.0000   .         .         .
  0.025%               -0.0611   0.0596   3.22     -1.03    0.3067
  0.05%                -0.0789   0.0724   4.81     -1.09    0.2777
  0.075%               -0.1219   0.0740   5.12     -1.65    0.1016
  0.1%                 -0.2062   0.0627   3.77     -3.29    0.0012
  0.2%                 -0.4883   0.0602   3.21     -8.11    0.0000
-----

```

Looking at the *estimated regression coefficients*, we see that the two highest dose groups have significantly lower body weights than controls ( $p=0.0012$  and  $0.0000$  for the 0.1% and 0.2% groups vs. controls, respectively). Again, design effects for regression coefficients in the range of 3-4 indicate more than a tripling in the variance of the estimated regression coefficients under the clustered design. The SUDAAN standard errors appropriately reflect this increase.

**Example 4 Results:****GEE Under Independent “Working” Correlations**

Date: 03-18-97                      Research Triangle Institute                      Page : 2  
 Time: 12:51:34                      The REGRESS Procedure                      Table : 1

Variance Estimation Method: **Robust** (Binder, 1983)  
 Working Correlations: **Independent**  
 Link Function: **Identity**  
 Response variable BW: Fetal Body Weight

**Treatment Effect on Fetal Body Weight in a Teratology Experiment**

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	6	29189.67	0.0000
MODEL MINUS INTERCEPT	5	78.38	0.0000
INTERCEPT	.	.	.
DOSEGRP	5	78.38	0.0000

The *main effects tests* indicate that the overall effect of treatment (with 5 degrees of freedom) is statistically significant, *after adjusting for clustering* ( $p=0.0000$ ).

**Example 4 Results:**

**GEE Under Independent “Working” Correlations**

```

Date: 03-18-97                Research Triangle Institute                Page : 3
Time: 12:51:34                The REGRESS Procedure                    Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Least-Square Means
                SE LS      T-Test      P-value
                LS Mean Mean      LSM=0      T-Test
                LSM=0
-----
DOSEGRP
Control          3.4979   0.0406   86.1031   0.0000
0.025%          3.4367   0.0436   78.7976   0.0000
0.05%           3.4190   0.0600   57.0111   0.0000
0.075%          3.3760   0.0619   54.5351   0.0000
0.1%            3.2916   0.0477   68.9535   0.0000
0.2%            3.0096   0.0444   67.7343   0.0000
-----

REGRESS used
CPU time       : 7.91 seconds
Elapsed time   : 8 seconds
Virtual memory : 1.29 MB
    
```

These *Least-Squares Means* are identical to the raw means and standard errors presented earlier, since there is only one covariate in the model. Body weights are reduced in the two highest dose groups vs. controls. The standard errors of the least-squares means are adjusted for clustering.

**Example 4 Results:****GEE Under Independent Working Correlations  
Model-based (Naive) Variance Estimates**

Below are the results obtained under working independence using the *model-based* or *naive variance-covariance matrix* of the estimated regression coefficients. The model-based variance is equal to the outside of the robust variance estimator,  $M_0^{-1}$ , or  $[X'V^{-1}X]^{-1}$ . In this case, the naive variance estimate is computed *as if the independent working correlation assumption were correct*. In other words, these are the results that would be obtained if clustering were ignored altogether. Although it is not recommended for analysis of clustered data, we are showing it to demonstrate the effects of clustering. We use **SEMETHOD=MODEL** on the PROC statement to obtain the model-based results.

```

43 PROC REGRESS DATA="c:\\tera\\examples\\boric" FILETYPE=SAS
      R=INDEPENDENT SEMETHOD=MODEL;

44 NEST _ONE_ DAMID;

45 WEIGHT _ONE_;

46 REFLEVEL DOSEGRP = 1;

47 SUBGROUP DOSEGRP;

48 LEVELS 6;

49 MODEL BW = DOSEGRP;

50 LSMEANS DOSEGRP;

51 SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
52 PRINT BETA="BETA" SEBETA="S.E." T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF WALDCHI WALDCHP / LSMEANS=ALL
      T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;
53 TITLE "Treatment Effect on Fetal Body Weight in a Teratology Experiment";

```

```

Opened SAS data file c:\tera\examples\boric.SSD for reading.
Number of observations read      :   1302   Weighted count:   1302
Observations used in the analysis :   1302   Weighted count:   1302
Observations with missing values :     0     Weighted count:     0
Denominator degrees of freedom  :    163

```

```

Maximum number of estimable parameters for the model is 6
File c:\tera\examples\boric.SSD contains 164 clusters
Maximum cluster size is 14 records
Minimum cluster size is 2 records

```

```

Weighted mean response is 3.341329

```

```

Multiple R-Square for the dependent variable BW: 0.171089

```

**Example 4 Results:**

**GEE Under Independent Working Correlations  
Model-based (Naive) Variance Estimates**

```

Date: 03-18-97                Research Triangle Institute                Page : 1
Time: 12:51:34                The REGRESS Procedure                Table : 1

Variance Estimation Method: Model-Based (Naive)
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Independent Variables and
Effects

                BETA      S.E.   T:BETA=0  P-VALUE
-----
Intercept                3.4979   0.0233   150.15   0.0000
DOSEGRP
  Control                0.0000   0.0000    .        .
  0.025%                -0.0611   0.0332   -1.84    0.0676
  0.05%                 -0.0789   0.0330   -2.39    0.0180
  0.075%                -0.1219   0.0327   -3.73    0.0003
  0.1%                  -0.2062   0.0323   -6.39    0.0000
  0.2%                  -0.4883   0.0336  -14.54    0.0000
-----

```

Here we see the *estimated standard errors* using the model-based approach under independence are much smaller than with the robust variance estimator, with several of the lower dose groups appearing significantly different from control. These estimates are overly optimistic (naive), computed as if the data were truly independent. Therefore, these results are not valid for the data at hand. They merely demonstrate the consequences of ignoring the experimental design.

**Example 4 Results:****GEE Under Independent Working Correlations  
Model-based (Naive) Variance Estimates**

Date: 03-18-97                      Research Triangle Institute                      Page : 2  
 Time: 12:51:34                      The REGRESS Procedure                      Table : 1

Variance Estimation Method: **Model-Based (Naive)**

Working Correlations: **Independent**

Link Function: **Identity**

Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	6	*****	0.0000
MODEL MINUS INTERCEPT	5	267.50	0.0000
INTERCEPT	.	.	.
DOSEGRP	5	267.50	0.0000

This table contains the *main effects tests* assuming the naive assumption of independence were true. The  $p$ -value according to the treatment effect is still significant ( $p=0.0000$ ), but the Wald chi-square is much larger (267.5) than under the robust variance approach.

**Example 4 Results:**

**GEE Under Independent Working Correlations  
Model-based (Naive) Variance Estimates**

```

Date: 03-18-97                Research Triangle Institute                Page : 3
Time: 12:51:34                The REGRESS Procedure                Table : 1

Variance Estimation Method: Model-Based (Naive)
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Least-Square Means                SE LS                T-Test                P-value
                                LS Mean Mean                LSM=0                T-Test
                                LSM=0
-----
DOSEGRP
Control                3.4979    0.0233    150.1543    0.0000
0.025%                3.4367    0.0237    145.1315    0.0000
0.05%                3.4190    0.0234    146.0900    0.0000
0.075%                3.3760    0.0230    146.9116    0.0000
0.1%                3.2916    0.0223    147.3570    0.0000
0.2%                3.0096    0.0242    124.3387    0.0000
-----

REGRESS used
CPU time                : 7.9 seconds
Elapsed time            : 8 seconds
Virtual memory          : 1.29 MB
    
```

Under the model-based approach, the *Least Squares Means* are the same as before, but their estimated standard errors are misleadingly small, computed as if the independence assumption were true.



**Example 4 Results:****GEE under Exchangeable Working Correlations**

Below are the programming statements to estimate the *linear model under exchangeable working correlations*. The only change from the previous statements is the switch from R=INDEPENDENT to *R=EXCHANGE* on the PROC statement. All other statements remain unchanged. By default, SUDAAN will use the GEE 1-step approach for estimating regression parameters, with the independence parameter estimates being updated exactly once with the estimated correlation structure. SUDAAN notifies the user that there are 164 clusters, 1302 fetuses, and a min and max cluster size of 2 and 14, respectively. Below is the first page of the REGRESS procedure output:

```

21  PROC REGRESS DATA="c:\\tera\\examples\\boric" FILETYPE=SAS R=EXCHANGE;
22  NEST _ONE_ DAMID;
23  WEIGHT _ONE_;
24  REFLEVEL DOSEGRP = 1;
25  SUBGROUP DOSEGRP;
26  LEVELS 6;
27  MODEL BW = DOSEGRP;
28  LSMEANS DOSEGRP;
29  SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
30  PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" DF WALDCHI WALDCHP / LSMEANS=ALL RHOS=ALL
        T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;
31  TITLE " "
        "Treatment Effect on Fetal Body Weight in a Teratology Experiment";

```

```

Opened SAS data file c:\tera\examples\boric.SSD for reading.
Number of observations read      :   1302   Weighted count:   1302
Observations used in the analysis :   1302   Weighted count:   1302
Observations with missing values :     0     Weighted count:     0
Denominator degrees of freedom  :    163

```

Maximum number of estimable parameters for the model is 6

```

File c:\tera\examples\boric.SSD contains 164 clusters
Maximum cluster size is 14 records
Minimum cluster size is 2 records
Weighted mean response is 3.341329

```

Multiple R-Square for the dependent variable BW: 0.169118

**Example 4 Results:**

**GEE Under Exchangeable Working Correlations**

```

Date: 03-18-97                Research Triangle Institute                Page : 3
Time: 12:51:34                The REGRESS Procedure                Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Exchangeable
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Variable                Correlation
                        Matrix
-----
Fetal Body Weight                0.5056
-----

```

The *estimated exchangeable correlation parameter* (measure of pairwise dependence within clusters) is 0.5056. The relatively large size of the intracluster correlation is partly responsible for the large design effects (variance inflation) for estimated means and regression parameters seen already. Variance inflation is directly related to the size of the intracluster correlation and the average cluster size (here, number of fetuses per litter).

**Example 4 Results:****GEE Under Exchangeable Working Correlations**

```

Date: 03-18-97           Research Triangle Institute           Page : 1
Time: 12:51:34          The REGRESS Procedure             Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Exchangeable
Link Function: Identity
Response variable BW: Fetal Body Weight

```

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----  
Independent Variables and  
Effects

	BETA	S.E.	T:BETA=0	P-VALUE
-----				
Intercept	3.5125	0.0470	74.67	0.0000
DOSEGRP				
Control	0.0000	0.0000	.	.
0.025%	-0.0509	0.0622	-0.82	0.4148
0.05%	-0.0656	0.0738	-0.89	0.3753
0.075%	-0.1363	0.0790	-1.73	0.0864
0.1%	-0.2409	0.0680	-3.54	0.0005
0.2%	-0.4822	0.0651	-7.40	0.0000
-----				

Above are the *estimated regression coefficients under exchangeability, with a robust variance estimator*. We see that these results are qualitatively the same as “working” independence showed previously. Modelling the within-cluster covariance structure has not improved efficiency in these data.

**Example 4 Results:**

**GEE Under Exchangeable Working Correlations**

Below are the *main effects tests* and *least squares means* under exchangeability with a robust variance estimator. Again, these results are similar to “working” independence shown earlier.

```

Date: 03-18-97          Research Triangle Institute          Page : 2
Time: 12:51:34        The REGRESS Procedure          Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Exchangeable
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Contrast                Degrees      P-value
                        of          Wald
                        Freedom   ChiSq    Wald
                        of          ChiSq    ChiSq
-----
OVERALL MODEL                6 28271.55  0.0000
MODEL MINUS INTERCEPT     5   75.41  0.0000
INTERCEPT                  .         .         .
DOSEGRP                      5   75.41  0.0000
-----

```

```

Date: 03-18-97          Research Triangle Institute          Page : 4
Time: 12:51:34        The REGRESS Procedure          Table : 1

Variance Estimation Method: Robust (Binder, 1983)
Working Correlations: Exchangeable
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Least-Square Means                SE LS      T-Test      P-value
                                   LS Mean    Mean        LSM=0       T-Test
                                   LS Mean    Mean        LSM=0       LSM=0
-----
DOSEGRP
Control                3.5125  0.0470  74.6666  0.0000
0.025%                 3.4616  0.0408  84.9373  0.0000
0.05%                  3.4469  0.0568  60.6350  0.0000
0.075%                 3.3762  0.0635  53.1718  0.0000
0.1%                   3.2716  0.0491  66.6914  0.0000
0.2%                   3.0303  0.0450  67.3092  0.0000
-----

```

**Example 4 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

Below are results from the exchangeable correlation model using the *model-based* or *naive variance-covariance matrix* of the estimated regression coefficients. The model-based variance is the  $M_0^{-1}$  matrix, or the outside portion of the robust variance estimate:  $M_0^{-1} = [X'V^{-1}X]^{-1}$ . In this case, the naive variance estimate is computed *assuming that the exchangeable "working" correlation assumption were correct*. Since that is close to truth for litter data, we will see that results are essentially the same as with the robust variance estimator.

```

32  PROC REGRESS DATA="c:\\tera\\examples\\boric" FILETYPE=SAS
      R=EXCHANGE  SEMETHOD=MODEL;

33  NEST  _ONE_  DAMID;

34  WEIGHT  _ONE_;

35  REFLEVEL DOSEGRP = 1;

36  SUBGROUP DOSEGRP;

37  LEVELS  6;

38  MODEL BW = DOSEGRP;

39  LSMEANS DOSEGRP;

40  SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;

41  PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
      P_BETA="P-VALUE" DF WALDCHI WALDCHP / LSMEANS=ALL RHOS=ALL
      T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;

42  TITLE  " "
      "Treatment Effect on Fetal Body Weight in a Teratology Experiment";

```

```

Opened SAS data file c:\tera\examples\boric.SSD for reading.
Number of observations read      :   1302   Weighted count:   1302
Observations used in the analysis :   1302   Weighted count:   1302
Observations with missing values :     0     Weighted count:     0
Denominator degrees of freedom  :    163

```

```

Maximum number of estimable parameters for the model is 6
File c:\tera\examples\boric.SSD contains 164 clusters
Maximum cluster size is 14 records
Minimum cluster size is 2 records
Weighted mean response is 3.341329

```

```

Multiple R-Square for the dependent variable BW: 0.169118

```

**Example 4 Results:**

**GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

```

Date: 03-18-97                Research Triangle Institute                Page : 1
Time: 12:51:34                The REGRESS Procedure                    Table : 1

Variance Estimation Method: Model-Based (Naive)
Working Correlations: Exchangeable
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Independent Variables and
Effects

                                BETA      S.E.    T:BETA=0  P-VALUE
-----
Intercept                      3.5125   0.0519    67.69    0.0000
DOSEGRP
  Control                       0.0000   0.0000     .         .
  0.025%                       -0.0509  0.0719    -0.71    0.4799
  0.05%                         -0.0656  0.0730    -0.90    0.3701
  0.075%                       -0.1363  0.0723    -1.88    0.0613
  0.1%                          -0.2409  0.0721    -3.34    0.0010
  0.2%                          -0.4822  0.0737    -6.54    0.0000
-----

```

Here we have the *estimated regression coefficients* computed under exchangeability and the standard errors as if the exchangeable working assumption were correct. The standard errors are roughly the same as with the robust variance estimator for these data, indicating that the exchangeable correlation assumption is close to truth.

**Example 4 Results:****GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

Date: 03-18-97                      Research Triangle Institute                      Page : 2  
 Time: 12:51:34                      The REGRESS Procedure                      Table : 1

Variance Estimation Method: **Model-Based (Naive)**

Working Correlations: **Exchangeable**

Link Function: **Identity**

Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	6	26050.93	0.0000
MODEL MINUS INTERCEPT	5	58.46	0.0000
INTERCEPT	.	.	.
DOSEGRP	5	58.46	0.0000

Here we have the *main effects tests* computed under exchangeability, using the model-based variance approach. Results are essentially the same as with the robust variance estimator.

**Example 4 Results:**

**GEE Under Exchangeable Working Correlations  
Model-Based (Naive) Variance Estimation**

Date: 03-18-97	Research Triangle Institute	Page : 4
Time: 12:51:34	The REGRESS Procedure	Table : 1
Variance Estimation Method: <b>Model-Based (Naive)</b>		
Working Correlations: <b>Exchangeable</b>		
Link Function: <b>Identity</b>		
Response variable BW: Fetal Body Weight		
Treatment Effect on Fetal Body Weight in a Teratology Experiment		
-----		
Least-Square Means		P-value
	SE LS	T-Test
	LS Mean Mean	LSM=0
		LSM=0
-----		
DOSEGRP		
Control	3.5125	0.0519 67.6924 0.0000
0.025%	3.4616	0.0497 69.6389 0.0000
0.05%	3.4469	0.0513 67.1676 0.0000
0.075%	3.3762	0.0504 66.9663 0.0000
0.1%	3.2716	0.0501 65.3301 0.0000
0.2%	3.0303	0.0523 57.9232 0.0000
-----		
REGRESS used		
CPU time	: 7.19 seconds	
Elapsed time	: 8 seconds	
Virtual memory	: 1.34 MB	



**Example 4 Results:****Jackknife Variance Estimation**

Below are the modelling results using an alternative approach: Jackknife variance estimation. We obtained these results by specifying *DESIGN=JACKKNIFE* on the PROC statement.

```

54  PROC REGRESS DATA="c:\\tera\\examples\\boric" FILETYPE=SAS  DESIGN=JACKKNIFE;
55  NEST _ONE_ DAMID;
56  WEIGHT _ONE_;
57  REFLEVEL DOSEGRP = 1;
58  SUBGROUP DOSEGRP;
59  LEVELS 6;
60  MODEL BW = DOSEGRP;
61  LSMEANS DOSEGRP;
62  SETENV COLSPCE=1 LABWIDTH=25 COLWIDTH=8 DECWIDTH=4 LINESIZE=78 PAGESIZE=60;
63  PRINT BETA="BETA" SEBETA="S.E." DEFT="DESIGN EFFECT" T_BETA="T:BETA=0"
        P_BETA="P-VALUE" DF WALDCHI WALDCHP / LSMEANS=ALL
        T_BETAfmt=F8.2 DEFTfmt=F6.2 DFFMT=F8.0 WALDCHIFMT=F8.2;
64  TITLE " "
        "Treatment Effect on Fetal Body Weight in a Teratology Experiment";

Opened SAS data file c:\tera\examples\boric.SSD for reading.

Number of observations read      :   1302      Weighted count:   1302
Observations used in the analysis :   1302      Weighted count:   1302
Observations with missing values :         0      Weighted count:         0
Denominator degrees of freedom  :    163

Maximum number of estimable parameters for the model is 6
File c:\tera\examples\boric.SSD contains 164 clusters
Maximum cluster size is 14 records
Minimum cluster size is 2 records
Weighted mean response is 3.341329

Multiple R-Square for the dependent variable BW: 0.171089

```

**Example 4 Results:**

**Jackknife Variance Estimation**

Below are the *estimated regression coefficients and standard errors* under the Jackknife option. Note that the regression coefficients are simply those computed under independence. The estimated standard errors are computed using the Jackknife variance estimator. Note the results for this example are similar to those obtained under GEE-independent and GEE-exchangeable.

```

Date: 03-18-97          Research Triangle Institute          Page : 1
Time: 12:51:34        The REGRESS Procedure          Table : 1

Variance Estimation Method: Jackknife
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Independent Variables and
Effects
          BETA      S.E.  DESIGN  T:BETA=0  P-VALUE
-----
Intercept          3.4979   0.0419   3.23     83.52    0.0000
DOSEGRP
  Control          0.0000   0.0000   .         .         .
  0.025%         -0.0611   0.0616   3.44     -0.99    0.3227
  0.05%          -0.0789   0.0752   5.18     -1.05    0.2955
  0.075%         -0.1219   0.0765   5.46     -1.59    0.1128
  0.1%           -0.2062   0.0648   4.03     -3.18    0.0018
  0.2%           -0.4883   0.0624   3.45     -7.83    0.0000
-----

```

**Example 4 Results:****Jackknife Variance Estimation**

Below are the *main effects tests* using the Jackknife approach. Again, they are similar to those obtained under GEE.

Date: 03-18-97                      Research Triangle Institute                      Page : 2  
 Time: 12:51:34                      The REGRESS Procedure                      Table : 1

Variance Estimation Method: **Jackknife**  
 Working Correlations: **Independent**  
 Link Function: **Identity**  
 Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

Contrast	Degrees of Freedom	Wald ChiSq	P-value Wald ChiSq
OVERALL MODEL	6	27208.92	0.0000
MODEL MINUS INTERCEPT	5	72.72	0.0000
INTERCEPT	.	.	.
DOSEGRP	5	72.72	0.0000

**Example 4 Results:**

**Jackknife Variance Estimation**

Below are the *least squares means* and their standard errors using the Jackknife approach.

```

Date: 03-18-97                Research Triangle Institute                Page : 3
Time: 12:51:34                The REGRESS Procedure                Table : 1

Variance Estimation Method: Jackknife
Working Correlations: Independent
Link Function: Identity
Response variable BW: Fetal Body Weight

Treatment Effect on Fetal Body Weight in a Teratology Experiment

-----
Least-Square Means

                SE LS      T-Test      P-value
                LS Mean Mean      LSM=0      LSM=0
-----
DOSEGRP
Control                3.4979    0.0419    83.5233    0.0000
0.025%                3.4367    0.0452    76.0471    0.0000
0.05%                 3.4190    0.0624    54.7695    0.0000
0.075%                3.3760    0.0640    52.7745    0.0000
0.1%                  3.2916    0.0495    66.5411    0.0000
0.2%                  3.0096    0.0462    65.0926    0.0000
-----

REGRESS used
CPU time      : 7.30 seconds
Elapsed time  : 8 seconds
Virtual memory : 1.22 MB
    
```



**Appendix III:**  
**Comparisons to Other Methods**

## Modelling vs. Accounting for Correlation

---

### EFFICIENCY

#### **Cluster-level or Time-stationary covariates:**

That is, a covariate that takes on the same value for all members of a given cluster (*i.e.*, constant within a cluster)

*e.g.*, a common treatment applied to all members of a given cluster.

*Modelling* the correlation structure and *accounting for* intracluster correlation have similar efficiency and power to detect cluster-level covariate effects.

References:      Bieler and Williams, 1994  
                     Neuhaus, 1993  
                     Lipsitz, et al., 1994a, 1994b  
                     Mancl and Leroux, 1996

## Modelling vs. Accounting for Correlation

---

### EFFICIENCY

#### **Within-Cluster or Time-Varying Covariates:**

A covariate where each member of a given cluster can potentially take on its own value (*i.e.*, not constant within a cluster)

*Modelling* the intracluster correlation:

- Provides more efficient estimation and more powerful tests of within-cluster covariates with different covariate patterns across individuals:

*e.g.*, Smoking Status at various points in time

Gains in efficiency increase with response intracluster correlation

- Provides no increase in efficiency for within-cluster covariates in which the pattern of the within-cluster covariate is the same for all clusters (for example, covariates that can be experimentally manipulated):

*e.g.*, the effect of *time*, where *time* may be recorded as the occasions of measurement (1, 2, 3, 4) for each individual

References:      Neuhaus, 1993  
                      Lipsitz, et al., 1994a, 1994b  
                      Mancl and Leroux, 1996



## Other Comparisons

---

### GEE / SUDAAN: Comparisons to other procedures and packages:

#### Vs. SAS PROC MIXED and BMDP5V:

- Same idea (Population-averaged or "marginal" models), and same estimating equations IF you specify only **fixed effects** and supply an assumed correlation structure for the random errors;
- Mixed models in SAS also allow for **random effects** (which are sometimes called cluster-specific models, because individual cluster deviations from population-averaged intercepts and slopes can be estimated); SUDAAN and the GEE macro of Karim and Zeger (1989) do not allow for random effects; this area is developing!
- Continuous outcomes only; use maximum likelihood (or REML) under a multivariate normality assumption to estimate  $\beta$  and  $V_i$
- SAS produces the *naive* or *model-based* variance estimate, and recently the robust variance estimator. The model-based variance assumes the working correlation structure is correct. Most important if you specify "independent" working correlations!
- SAS does not iterate between correlations and regression coefficients; it estimates  $V_i$  directly

#### Biomedical Research:

Laird and Ware (1982); Laird, et al (1992)

#### Psychology and Behavioral Research:

Bryk and Raudenbush (1987)

Goldstein (1987)

Hedeker, Gibbons, and Davis (1991)

Hedeker, Gibbons, and Flay (1994)

## Other Comparisons

---

### Review of the General Linear Mixed Model

- Extension of the General Linear Model by allowing both correlation and non-identity (heterogeneity of variance), although still assuming normality:

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

- Elements of  $\epsilon$  *not* independent and identically distributed with zero mean and variance  $\sigma^2$  (however, normality still assumed).
- $y$  is an observed data vector  
 $X$  is the design matrix for the unknown fixed effects,  $\beta$   
 $Z$  is a design matrix for the unknown random effects,  $v$   
 $\epsilon$  is a vector of unobserved random errors.

$$E \begin{bmatrix} v \\ \epsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\text{Var} \begin{bmatrix} v \\ \epsilon \end{bmatrix} = \begin{bmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{R} \end{bmatrix}$$

$$V(y) = \mathbf{ZGZ}' + \mathbf{R}$$

## Other Comparisons

---

### Review of the General Linear Mixed Model (continued)

- Model  $V$  by setting up a random-effects design matrix  $Z$  and by specifying covariance structures for  $G$  and  $R$

For example, the General Linear Model is a special case of a mixed model with  $Z = \mathbf{0}$  and  $R = \sigma^2 I$

- Estimates of fixed effects parameters (ML or REML):

$$\hat{\beta} = (X'V^{-1}X)^{-1} X'\hat{V}^{-1}y$$

$$Var(\hat{\beta}) = (X'\hat{R}^{-1}X)^{-1} \quad (\text{Model-based})$$

- Therefore, mixed models can be used for modelling repeated measures or longitudinal data:

$R$  matrix is the place to model the covariance structure of a subject's data: e.g., compound symmetry, AR(1), unstructured.

- Note that the estimating equations have the same form as GEE

## Other Comparisons

---

### **GEE / SUDAAN: Comparisons to other procedures and packages:**

#### **Vs. SAS GLM:**

- Idea is that of MIXED with exchangeable correlations (if you use the univariate approach) with fewer options for the correlation matrix (i.e., fits the compound symmetry structure only).
- Throws out entire clusters with missing observations
- Method-of-moments estimation for variance components; least squares for parameter estimation

#### **Vs. SAS CATMOD**

- Log-linear modelling of contingency tables
- GSK and WLS methods  
See Koch, et al, 1977; Grizzle, Starmer, Koch, 1969
- Analogy: CATMOD : GEE as ANOVA : REGRESSION  
*i.e.*, good for categorical covariates that are not sparsely distributed

See Zeger, 1988

## Other Comparisons

---

### **GEE / SUDAAN: Comparisons to other procedures and packages:**

#### **Vs. Conditional Logistic Regression and Mixed Logistic Models**

- Mixed effects models and random regression models for binary and ordinal responses:

Stiratelli, Laird, and Ware (1984)

Gibbons and Hedeker (1994)

Hedeker and Gibbons (1994)

- Conditional likelihood approach for matched pair data:  
Breslow and Day (1980).
- Cluster-specific parameters in model; intercept terms allowed to vary between clusters according to a specified distribution
- Regression coefficients have different meaning than for marginal models; tells how covariates effect responses for particular "latent" risk groups represented by the random intercepts
- Random effects and mixed models better for within-cluster covariates

Comparisons more fully discussed in:

Neuhaus, Kalbfleisch, and Hauck, 1991

Zeger, Liang, and Albert, 1988

Park, 1993