1994

A Study of the Two Major Causes of Neonatal Deaths: Perinatal Conditions and Congenital Anomalies

Felipe Lorenzo-Luaces
University of North Florida

A STUDY OF TWO MAJOR CAUSES OF NEONATAL DEATHS: PERINATAL CONDITIONS AND CONGENITAL ANOMALIES

by

Felipe Lorenzo-Luaces

A thesis submitted to the Department of Mathematics and Statistics in partial fulfillment of the requirements for the degree of

Master of Science in Mathematical Sciences

University of North Florida
College of Arts and Sciences

July, 1994

This thesis of Felipe Lorenzo-Luaces is approved:

(Date)

Signature Deleted

7/19/94

Signature Deleted

July 19, 1994

Signature Deleted

July 19, 1994

Committee Chairperson

Accepted for the Department:

Signature Deleted

July 19, 1994

Accepted for the College:

Signature Deleted

July 27, 1994

Dean

Accepted for the University:

Signature Deleted

8/1/94

Dean of Graduate Studies

Acknowledgments

I wish to thank Dr. Pali Sen, the advisor of this thesis, for her motivation and guidance in completing the work. I wish to thank the other review committee members, Dr. Donna Mohr and Dr. Adel Boules, for their time and useful comments to make revisions for this paper. I would also like to thank the faculty, staff and students of this University for their input and assistance with this paper.

I also wish to thank my family for their love and support while I worked on this project. Without their help I could not have completed it.

TABLE OF CONTENTS

List of Tables and Figures

Abstract

Infant mortality is a public health concern in the United States. We concentrate on neonatal mortality for its high accountability of infant mortality. In this paper we study the neonatal mortality of Florida's 1989 live birth cohort.

The data has been analyzed for two major causes of deaths: perinatal conditions and congenital anomalies. We use the KAPLAN-MEIER method to estimate the survival probabilities. For each cause, data were fit to the Weibull models and Extreme Value models to estimate the parameters of the survival curves. The results indicate that primary factors for each cause of neonatal deaths are very low birth weight, prior pregnancies of the mother, and late initiation of prenatal care when the variables are considered separately. The conclusion still remains the same for perinatal conditions when the interaction effects of the factors are considered, but we do not conclude similarly for the congenital anomalies at the same interaction level.

Chapter 1 - Introduction

In this paper, we study neonatal mortality among the 1989 live birth cohort in the state of Florida, specific to two major categories of underlying causes of death: perinatal conditions and congenital anomalies. The neonatal period for an infant is defined from birth to 27 days. This is a critical period for a birth because the hazard rate or force of mortality is much higher during this period (Elandt-Johnson and Johnson, 1980).

For this paper, neonatal deaths are classified into two major categories of underlying causes of death. These categories are established in the International Classification of Diseases (ICD; World Health Organization, 1979). Perinatal conditions (ICD codes 760-779) are defined as conditions originating closely surrounding the time of birth which result in the death of the infant. Two examples of perinatal conditions are intrauterine growth retardation and respiratory distress syndrome. Congenital anomalies (ICD codes 740-759) are genetic conditions which affect fetus development. Two examples of congenital anomalies are spina bifida and Down's syndrome.

Infant mortality is a subject of great interest for demographers and public health researchers for obvious

reasons. The main area of interest is to identify both biological and social variables which influence the infant mortality rate. A review of literature shows that the principal area of interest lies in the identification of the main demographic variables and their interaction effects on infant mortality. Cramer (1987) discovered in his study that race/ethnicity interacts with education and timing of the prenatal care. In addition, Cramer observed that birth weight and race/ethnicity interact differently in various causes of death. A study by Eberstein, Nam, and Hummer (1990) observed similar demographic main and interaction effects by causes of death. Their study found evidence which suggests that interactions among variables relating to infant mortality are complex and attention needs to be given to interactions in further research on this topic.

This study is different from prior research in that we identify independence among some of the main effects used by Eberstein, Nam, and Hummer for the neonatal deaths only. These variables are then used to create stratification levels in the live birth cohort to introduce the interaction effects among the variables. For each stratification level combination the survival distribution function is estimated using a nonparametric model. We define the dependent variable, age of infant at time of death, as failure time. There are eight explanatory variables. Out of these, seven of the explanatory variables are related to the mother:

maternal race/ethnicity (RACE), AGE, education (EDUC), marital status (UNWED), time when prenatal care began (CARE), rank order of this birth (ORDER), and number of prior pregnancies to the mother (PREG). The last explanatory variable, birth weight measured in grams (WEIGHT), is related to the infant.

The explanatory variables are measured in ordinal or in categorical levels, as appropriate. The four categorical variables are RACE, UNWED, PREG, and ORDER. RACE has three levels black, white/hispanic, and other. UNWED has two levels married and unmarried. The variable PREG has two levels: no prior pregnancy or at least one prior pregnancy. ORDER has two levels: first birth and at least second birth. The ordinal variables are AGE, EDUC, CARE, and WEIGHT with their levels defined as intervals. AGE has three intervals: less than 20 years, 20-29 years, and 30 or more years. EDUC has three intervals: 0-8 grade, 9-12 grade, and 13+ grade. CARE has three intervals: 1-3 months, 4-6 months, and 7-9 months. The last interval of CARE also includes no care during the pregnancy. WEIGHT has three intervals: less than 1500 grams, 1500-2499 grams, and 2500 grams or more. The levels for WEIGHT are chosen because of their current use in public health research. The first interval defines very low birth weight, the second defines low birth weight, and the third one is considered as normal weight.

The Life Table model is used as a nonparametric estimator of the survivor function (Kalbfleisch and Prentice, 1980). The proper exponential form of the survival function is determined by graphical analysis (Cox and Oaks, 1984). The Weibull model is used with some of the explanatory variables as covariates to fit the data using Life Regression procedure given in SAS. We also use an extreme value model (Agresti, 1990, Elandt-Johnson and Johnson, 1980) for different strata levels of the explanatory variables and use the Logistic Regression to fit the models to the data using SAS.

In Chapter 2 we discuss the methods for testing independence among explanatory variables conditioning on neonatal deaths and determining strata levels as well for model selection. Chapter 3 deals with the results obtained from the various hypotheses tested and a discussion is included in Chapter 4.

Chapter 2 - Methods

The independence among the explanatory variables is necessary for the assumptions of the different models used in the analysis. Independence tests of explanatory variables will identify important variables to be considered for our analysis. We also would like to find differences in the levels of these variables for further stratification. Here we use the nonparametric logrank test for right adjustment and Wilcoxon tests for left adjustment for both causes of death.

Section 1 Test of Independence for Explanatory Variables

For independence testing we use the Pearson Chi-Square test to test the hypothesis that the levels of one variable are independent of the levels of the second variable. In two-way contingency tables with multinomial sampling, the null hypothesis tested for statistical independence is $H_0: p_{ij} = p_{i+}\, p_{+j}$ for all i and j.

The probability distribution $\{p_{ij}\}$ is the joint distribution of variables X and Y. The marginal distributions are the row and column totals obtained by summing the joint probabilities. These are denoted by $\{p_{i+}\}$ for the row variable and $\{p_{+j}\}$ for the column variable,

where the subscript "+" denotes the sum over the index it replaces (Agresti, 1990).

Section 2 Nonparametric Model

We define the variable infant's age at death as the number of days from birth until death. This variable has the same meaning as failure time. An infant is at risk of dying for a random length of time (T_i) having a probability distribution $F(t_i)$ of survival. We censor infants who died from other causes during the neonatal period and those who survive past the neonatal period.

For our first model we will use the Life Table as a nonparametric model to estimate the survivor function. The advantage of this model is that no assumptions of normality for the data are needed. The KAPLAN-MEIER (KM) estimator $\hat{F}(t)$ will be used to estimate the survivor function. $\hat{F}(t)$ is defined as:

$$\hat{F}(t) = \prod_{j|t_j<t} \frac{(n_j - d_j)}{n_j}, \qquad (1)$$

where $n_j = (m_j + d_j) + \ldots + (m_k + d_k)$, is the number at risk just prior to t_j ($j = 1, \ldots, k$). Here $t_1 < t_2 < t_3 < \ldots < t_k$ represent the observed failure times or age at death for the neonatal deaths from the homogeneous birth cohort with survivor function $\hat{F}(t)$. Next we define d_j to be the number of neonatal deaths at time t_j and m_j as the

number of censored observation in the time interval $[t_j, t_{j+1})$.

To test for differences in mortality between the levels of the variables, we use the nonparametric logrank and the Wilcoxon tests. The null hypothesis tested for statistical difference is $H_0: F_1(t)=F_2(t)=\ldots=F_r(t)$, where r will determine the population stratification level of the explanatory variable(s).

Section 3 Parametric Models

The analysis of the Life Regression procedure was done for all of the exponential, Weibull, gamma, log-normal models for both causes of death and for each covariate. For these models, the dependent variable is the failure time. The vector $\mathbf{z}=(z_1, z_2, \ldots, z_s)$ of explanatory variables (or covariates) is considered for both censored and uncensored observations. The probability distribution of $T_i>0$ can be specified in three ways:

1) Probability density function $g(t_i)$
2) Hazard function $h(t_i)$
3) Survivor function $F(t_i) = 1-P(T_i<t)$, $0 < t < \infty$

By definition the survival function is given by

$$F(t_i) = g(t_i)\ [h(t_i)]^{-1}.$$

For Weibull distribution the probability density function and the hazard function, conditional upon **z**, are given by:

$$1)\ g(t_i;\mathbf{z}) = \alpha\rho(\alpha t_i)^{\rho-1}e^{\mathbf{zb}}\ \exp[-(\alpha t_i)^{\rho}e^{\mathbf{zb}}] \qquad (2)$$

$$2)\ h(t_i;\mathbf{z}) = \alpha\rho(\alpha t_i)^{\rho-1}e^{\mathbf{zb}} \qquad (3)$$

Hence, the survival function is

$$F(t_i,\mathbf{z}) = \alpha\rho(\alpha t_i)^{\rho-1}e^{\mathbf{zb}}\ \exp[-(\alpha t_i)^{\rho}e^{\mathbf{zb}}]\ [\alpha\rho(\alpha t_i)^{\rho-1}e^{\mathbf{zb}}]^{-1}$$

$$= \exp[-(\alpha t_i)^{\rho}e^{\mathbf{zb}}]. \qquad (4)$$

Here α, ρ, and **b** parameters are estimated by the SAS procedure. The parameter α is an adjustable parameter with the dimension of the reciprocal of time. The parameter ρ is an index which changes with the shape of the hazard function. The parameter $\mathbf{b}' = (b_1, b_2, \ldots, b_s)$ is a vector of regression parameters.

The test for the Weibull distribution is done graphically by plotting $\ln(-\ln(F(t_i;\mathbf{z})))$ against $\ln(t_i)$. If T_i has the Weibull distribution, the graph should be a straight line. The estimate of the parameter ρ is the slope of the line and the estimate of α is the intercept of the line.

Section 4 Extreme Value Model

It is suggested that in modeling adult mortality, a more rapidly increasing hazard function than that represented by the Weibull distribution is sometimes necessary. One such model is known as the extreme value model. An extreme value model is defined for the survival probabilities which depart from 1 more sharply than those probabilities which depart from 0. The extreme value model uses the Gompertz hazard function (Kalbfleisch and Prentice, 1980). The Gompertz hazard function is defined as

$$h(t)=\lambda\exp(\gamma t), \qquad (5)$$

where γ is the Euler's constant. This distribution is widely used in actuarial sciences to model human mortality. The Gompertz distribution is a form of an extreme value distribution (type 1) as described in Elandt-Johnson and Johnson (1980). They state that the name Extreme Value Distributions applies to three types of limiting distributions which approximate the shape of distributions of extreme values (the least or the greatest) in large random samples. If T_i has a Weibull distribution then $Y_i=\ln(T_i)$ has a type 1 extreme value distribution, where $t_i>0$. From the previous model we know that T_i has a Weibull distribution; therefore, our next model is an extreme value

model for $\ln(T_i)$. A proof is given in the Appendix. For each cause of death, we treat the explanatory variables, WEIGHT, PREG, and CARE, as strata variables and the variable Y_i as the independent variable. The survival probability at each Y_i is used as the response variable and the extreme value model is obtained by using the Log-Log Link function in the Logistic Regression procedure given in SAS. The inverse cumulative distribution function,

$$G_r(y_i)=1-\exp[-e^{\alpha+\beta y_i}], \qquad (6)$$

where r identifies the combination of the strata level, is used to model the data (Agresti, 1990). Then the survival function equals $1-G_r(y_i)$.

Using the general linear model for complementary log-log link, the regression equation becomes

$$\ln(-\ln(1-G_r(y_i)))=\alpha+\beta y_i. \qquad (7)$$

Here the parameters α and β are estimated by the Logistic procedure in SAS. The graphs of the observed versus the predicted survival function estimates illustrate the fit for both the Weibull and extreme value models.

Chapter 3 - Results

Data for this analysis are linked birth to infant death certificates for the cohort of live births in Florida during 1989. The data is provided by the Department of Health and Rehabilitative Services (HRS). The data contained 192,581 live births of which 1,228 resulted in neonatal deaths. Among the neonatal deaths, 890 died from perinatal conditions and 254 died from congenital anomalies.

Performing a frequency procedure of the entire birth cohort file to test independence among the explanatory variables revealed that all the variables are dependent. These results are shown in Table 1. A large Chi-square value and a p-value less than 0.05 for the test indicates dependence between the variables. However, when we use data for neonatal deaths only, we find some of the variables are conditionally independent. The results of the Chi-square test for conditional independence are in Table 2. The test revealed independence among the variables WEIGHT, PREG, and CARE with p-values greater than 0.05. UNWED, ORDER, and EDUC were found to be independent in some comparisons, but all three were dependent when compared to CARE. In conclusion, the best possible explanatory variables are WEIGHT, PREG, and CARE.

The Lifetest procedure in SAS is performed to obtain the KM survival probability estimators and to test for homogeneity between the causes of death, perinatal conditions, and congenital anomalies. Table 3 shows results of the nonparametric Logrank and Wilcoxon tests. These tests determine homogeneity between the survivor functions of perinatal conditions and that of congenital anomalies. The tests show that the survivor functions are significantly different at $p<0.05$. The plot of survival probability for each cause of death against age at death (Figure 1) reveals the difference between the two survival curves. The survival curve for the perinatal deaths is lower than that of congenital anomalies over time. This is most apparent during the earlier survival times.

Figure 2 reveals that the exponential model is not appropriate for modeling the probabilities of survival curves. The curves are not linear for the graph of $-\ln(F(t))$ versus age at death. However, the graph of $\ln(-\ln(F(t)))$ versus ln(age at death) (Figure 3) show linearity for both curves. Hence, the Weibull model seems to be appropriate for modeling the data.

Also, the Lifetest procedure is used to determine if any of the independent explanatory variables WEIGHT, PREG, and CARE can be used as strata for each cause of death. Table 3 shows the results of this numerical analysis of both

nonparametric tests: Logrank and Wilcoxon. The results show that some of the levels of each variable can provide significantly different survival functions for each strata for both perinatal conditions and congenital anomalies. The p-value for each test is less than 0.05. In addition to the numerical analysis, the three diagnostic plots, survival probability versus age at death, -ln(F(t)) versus age, and ln(-ln(F(t))) versus ln(age), are created for the variables WEIGHT, PREG, and CARE for each cause of death.

For neonatal deaths a Lifetest procedure is performed to obtain a data set containing the Wilcoxon rank statistics to be used as input in a SAS Regression procedure. The objective here is to use the stepwise regression method to determine how much of the variability in age at death can be explained by the variables WEIGHT, CARE, and PREG. The results have been abstracted and are shown in Table 4. This table shows that 90 percent of the variability in the variable age at death is explained by the variable WEIGHT. Each of the remaining variables accounted for less than 2 percent of the variability in age at death. The variables WEIGHT, PREG, and CARE can explain 93 percent of the variability in the variable age at death for neonatal deaths. Similar Lifetest and Regression procedures with the entire 192,581 births revealed that only 10% of the variability was explained by these variables. However, the death rate is only 0.6 % for the entire birth file.

For both causes of death, the survival probability plots show that the curves of the 2nd and 3rd intervals of the variable WEIGHT are significantly different from that of the 1st interval (Figure 4 and 7). The non-linearity of -ln(F(t)) for both causes indicates that the exponential model is not appropriate for modeling the WEIGHT variable (Figure 5 and 8). The linear plots of ln(-ln(F(t))) for both causes indicates that the Weibull model is appropriate for this data.(Figure 6 and 9). These graphs also show that the higher birth weights give higher probability of survival. As a result, we decided to combine the 2nd and 3rd intrevals of the variable WEIGHT into one interval, 1500 grams or more.

The survival probability plots of the variable CARE for both causes of death also reveal some non-linear trend in the data (Figures 10 and 13). The plots also show that the births where the CARE begins in the first and second trimester of the pregnancy have a much higher survival probability over time than those beginning care in the third trimester. The non-linear plots of -ln(F(t)) versus age, for each cause, indicate that the exponential model is not appropriate for this variable (Figure 11 and 14). The plots of ln(-ln(F(t))) versus the ln(age at death) show linearity for both causes of death, indicating the Weibull model is appropriate for modeling the variable CARE (Figures 12 and

15). The variable CARE shows the first and second trimester survival probabilities are similar. Therefore, we shall combine these two categories of the variable CARE. The combined variable now indicates if care was received late or early in the pregnancy.

The variable PREG revealed for both causes that the births to mothers with no prior pregnancies had a higher survival probability over time than those which had at least one prior pregnancy. This is shown in the survival probability plots for both causes (Figures 16 and 19). These plots show some type of exponential model is appropriate. For both causes the non-linearity of the plot -ln(F(t)) versus age (Figures 17 and 20) indicates the exponential model is not appropriate. The Weibull model is appropriate, however, because of the linearity of the plots ln(-ln(F(t))) versus ln(age at death) (Figures 18 and 21).

The Life regression procedures in SAS give estimates of survival probabilities for each stratum, created by the combination of the explanatory variables, under each cause of death. The data for the strata is described in Table 5. The survival probability estimates are plotted along with the observed survival probability estimates obtained from the Lifetest procedure. The corresponding graphs show the variables WEIGHT, CARE, and PREG with different combinations. Figures 22(a) through 29(a) provide the

observed and the predicted model probabilities for perinatal conditions. Figures 22(a), 24(a), 26(a), 28(a), and 29(a) show the Weibull model provided a better fit for later times than for earlier times. Figures 23(a) and 25(a) show that when the variable CARE indicates late care the predicted probabilities become proportionately displaced by some constant. This effect is not as evident when the birth weight of the infant is less than 1500 grams. The effect of the variables WEIGHT and CARE are shown in Figures 27(a) and 29(a). For perinatal conditions, the Weibull model tends to over estimate the probability of survival at earlier times regardless of which variables are present in the model.

For congenital anomalies, figures 30(a) through 37(a), provide the observed Lifetest and predicted Life regression procedures survival probability estimates for the Weibull model. Figures 30(a) shows that when all the covariates are present in the model, with WEIGHT greater than 1500 grams, CARE received early, and the mother has had no prior pregnancies, the model provides a good fit. Figures 31(a) through 37(a) show that when any other levels of the variables are present in the model the predicted curves show more variations from the observed. For congenital anomalies, the Weibull model frequently over estimates the probabilities of survival at earlier times, as it did for perinatal conditions.

For deaths due to perinatal conditions, figures 22(b) through 29(b), plot the observed survival probabilities against the predicted by the Logistic procedure in SAS using the complementary log-log link for the extreme value model for each of the eight strata levels. Figures 22(b), 24(b), 26(b), 27(b), 28(b), and 29(b) illustrate a much better fit than the corresponding Weibull model. These curves show less deviation between the observed and the predicted probability estimates.

For deaths due to congenital anomalies, figures 31(b) through 37(b) show the observed survival probabilities with the predicted by the complementary log-log link for the extreme value model for each of the eight strata levels. The figures indicate that the model is extremely good at predicting the survival probabilities. The extreme value model provides a better fit of the survival curves than the Weibull model for all strata levels for deaths due to perinatal conditions as well as for deaths due to congenital anomalies.

Chapter 4 - Discussion

The results are significant because they reveal that infants born to mothers with prior pregnancies tend to be at a higher risk of a neonatal death from both perinatal conditions and congenital anomalies. The infants with low birth weight and with late prenatal care were expected to have low survival probabilities for both causes of death. The survival probabilities were higher for infants with high birth weight and with early prenatal care for both causes of deaths.

The analysis of independence among the explanatory variables revealed that the three variables, birth weight, number of prior pregnancies, and time when prenatal care began, are conditionally independent among the neonatal deaths but not among the live births cohort. These three variables all provided significant difference between their levels so as to create strata levels within each cause of death. The plots of survival probabilities with age at death show that infants who died from perinatal deaths have shorter survival times than infants who died from congenital anomalies with life expectancy of 2.4 days for one and about 3.5 days for the other.

odels provided good fit of the survival functions for two causes. However, near the end points of age at death the Weibull model failed to provide good estimates for the survival probabilities. Since the complementary log-log link model is an extreme value model it provided a better fit of the data at all times including the beginning and the ending points. Both models performed well when the variable combination was birth weight 1500 grams or more, no prior pregnancies, and early care, for both causes of death.

For neonatal deaths, the three most significant factors are birth weight, prior pregnancies, and prenatal care. The birth weight of the infant had an affect for both causes of death, but was more pertinent to perinatal conditions. Deaths due to perinatal conditions can be controlled by mother's demographic variables. These variables are timing of prenatal care and prior pregnancies which are of social interest, and can be controlled by careful attention. Further study should address how other social behaviors, such as tobacco and alcohol use, affect infant mortality. A recent data from the years 1990 through 1992, which were not available at the beginning of our study, should be considered for these factors.

The extreme value model performed extremely well for all times of age at death and for all strata levels. The

model choice for this type of analysis should be the extreme value model, not the Weibull model. This study is significant in determining the best model for interaction among the explanatory factors.

Table 1

Results of Independence Test Between Explanatory Variables Using Entire Birth Cohort

Constant Variable	By Variable	df	Chi-square	p-value
WEIGHT	PREG	2	32.448	0.000*
	CARE	4	1004.670	0.000*
	UNWED	2	1971.385	0.000*
	ORDER	2	7.896	0.019*
	AGE	4	221.494	0.000*
	EDUC	4	450.383	0.000*
	RACE	4	2419.428	0.000*
PREG	CARE	2	452.348	0.000*
	UNWED	1	762.714	0.000*
	ORDER	1	125967.940	0.000*
	AGE	2	18614.494	0.000*
	EDUC	2	77.761	0.000*
	RACE	2	532.353	0.000*
CARE	UNWED	2	21201.246	0.000*
	ORDER	2	1211.795	0.000*
	AGE	4	10392.452	0.000*
	EDUC	4	14307.231	0.000*
	RACE	4	9697.899	0.000*
UNWED	ORDER	1	708.123	0.000*
	AGE	2	24618.384	0.000*
	EDUC	2	18349.462	0.000*
	RACE	2	36508.524	0.000*
ORDER	AGE	2	17285.759	0.000*
	EDUC	2	547.697	0.000*
	RACE	2	1569.836	0.000*
AGE	EDUC	4	24234.575	0.000*
	RACE	4	5743.996	0.000*
EDUC	RACE	4	12422.533	0.000*

* indicates non-independence.

Table 2

Results of Independence Test Between Explanatory Variables Using Neonatal Deaths Only

Constant Variable	By Variable	df	Chi-square	p-value
WEIGHT	PREG	2	1.094	0.579
	CARE	4	8.266	0.082
	UNWED	2	2.357	0.308
	ORDER	2	0.491	0.782
	AGE	4	5.783	0.216
	EDUC	4	4.168	0.384
	RACE	4	22.042	0.000*
PREG	CARE	2	1.096	0.578
	UNWED	1	0.764	0.382
	ORDER	1	590.859	0.000*
	AGE	2	85.480	0.000*
	EDUC	2	1.191	0.551
	RACE	2	16.455	0.000*
CARE	UNWED	2	154.309	0.000*
	ORDER	2	13.397	0.001*
	AGE	4	58.443	0.000*
	EDUC	4	72.027	0.000*
	RACE	4	56.211	0.000*
UNWED	ORDER	1	0.031	0.859
	AGE	2	110.745	0.000*
	EDUC	2	111.069	0.000*
	RACE	2	246.523	0.000*
ORDER	AGE	2	92.346	0.000*
	EDUC	2	1.609	0.447
	RACE	2	19.464	0.000*
AGE	EDUC	4	130.598	0.000*
	RACE	4	29.404	0.000*
EDUC	RACE	4	93.169	0.000*

* indicates non-independence.

Table 3

Results of Stratification Testing
Using Neonatal Deaths

Strata	Test	Chi-square	df	p-value
CAUSE	Logrank	9.870	1	0.0017*
CAUSE	Wilcoxon	23.782	1	0.0001*
CAUSE: Perinatal conditions				
WEIGHT	Logrank	46818.806	2	0.0001*
WEIGHT	Wilcoxon	46809.507	2	0.0001*
PREG	Logrank	14.130	1	0.0002*
PREG	Wilcoxon	14.125	1	0.0002*
CARE	Logrank	105.017	2	0.0001*
CARE	Wilcoxon	105.171	2	0.0001*
CAUSE: Congenital anomalies				
WEIGHT	Logrank	1760.599	2	0.0001*
WEIGHT	Wilcoxon	1762.058	2	0.0001*
PREG	Logrank	4.794	1	0.0286*
PREG	Wilcoxon	4.809	1	0.0283*
CARE	Logrank	15.987	2	0.0003*
CARE	Wilcoxon	16.016	2	0.0003*

* indicates significantly different.

Table 4

Selection of Variables Using Stepwise Regression
for Neonatal Deaths

Number in Model	R-Square	Variables in Model
1	0.90345349	WEIGHT
1	0.01920223	CARE
1	0.00602512	PREG
2	0.92765334	WEIGHT CARE
2	0.90812360	WEIGHT PREG
2	0.02404078	PREG CARE
3	0.93116111	WEIGHT PREG CARE

Table 5

Summary of the Number of Censored an Uncensored Values by Cause

Legend for Strata Symbol

Strata	WEIGHT	PREG	CARE
A	1500+	NO	EARLY
B	1500+	NO	LATE
C	1500+	1+	EARLY
D	1500+	1+	LATE
E	<1500	NO	EARLY
F	<1500	NO	LATE
G	<1500	1+	EARLY
H	<1500	1+	LATE

CAUSE: Congenital Anomalies

Stratum	Total	Failed	Censored
A	57801	41	57760
B	3794	4	3790
C	116971	116	116855
D	11208	21	11187
E	684	16	668
F	114	5	109
G	1663	44	1619
H	346	7	339
Total	192581	254	192327

CAUSE: Perinatal Conditions

Stratum	Total	Failed	Censored
A	57801	36	57765
B	3794	7	3787
C	116971	78	116893
D	11208	17	11191
E	684	167	517
F	114	26	88
G	1663	462	1201
H	346	97	249
Total	192581	890	191691

Table 6

Parameter Estimates for Wiebull and Extreme Value Models
by Causes of Death

	Perinatal Conditions	Congenital Anomalies
Variables	Parameter Estimates	Parameter Estimates
intercept	32.6050	29.0108
WEIGHT	-24.3717	-12.4130
PREG	- 0.5781	- 0.9214
CARE	- 0.4266	- 1.3926
scale	3.9912	3.5830

Weibull Model

	Perinatal Conditions		Congenital Anomalies	
	Parameter Estimates		Parameter Estimates	
Strata	intercept	Y	intercept	Y
A	2.0702	- 0.0216	2.0699	- 0.0277
B	1.8899	- 0.0164	1.9723	- 0.0191
C	2.0769	- 0.0254	2.0095	- 0.0235
D	1.8945	- 0.0086	1.8755	- 0.0140
E	0.4766	- 0.0391	1.3880	- 0.0362
F	0.4801	- 0.0301	1.2823	- 0.0546
G	0.4110	- 0.0510	1.3425	- 0.0254
H	0.3710	- 0.0396	1.4332	- 0.0310

Extreme Value model (refer to Table 5)

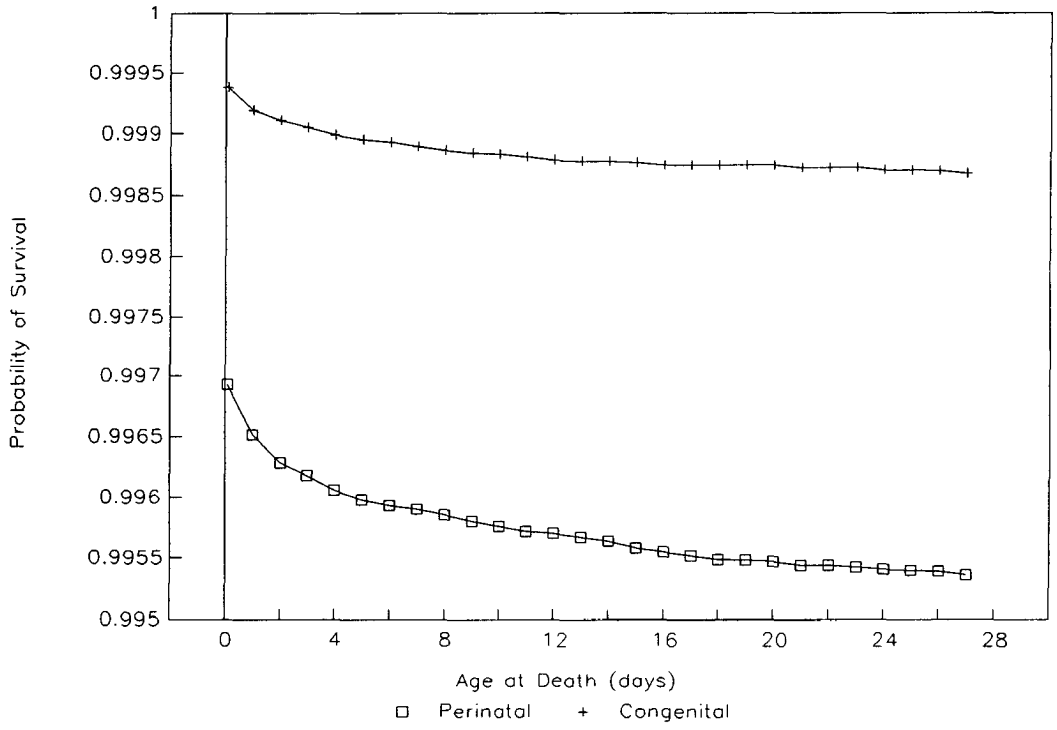


Figure 1. Probability of survival versus age at death (stratified by cause of death).

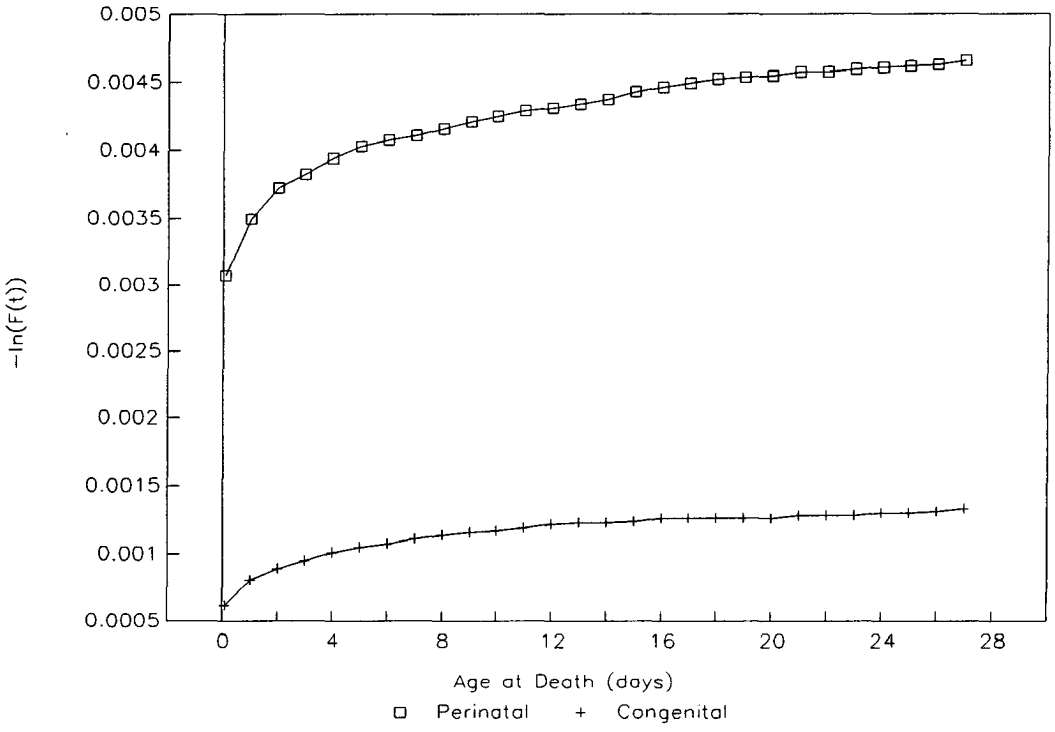


Figure 2. -ln(F(t)) versus age at death (stratified by cause of death).

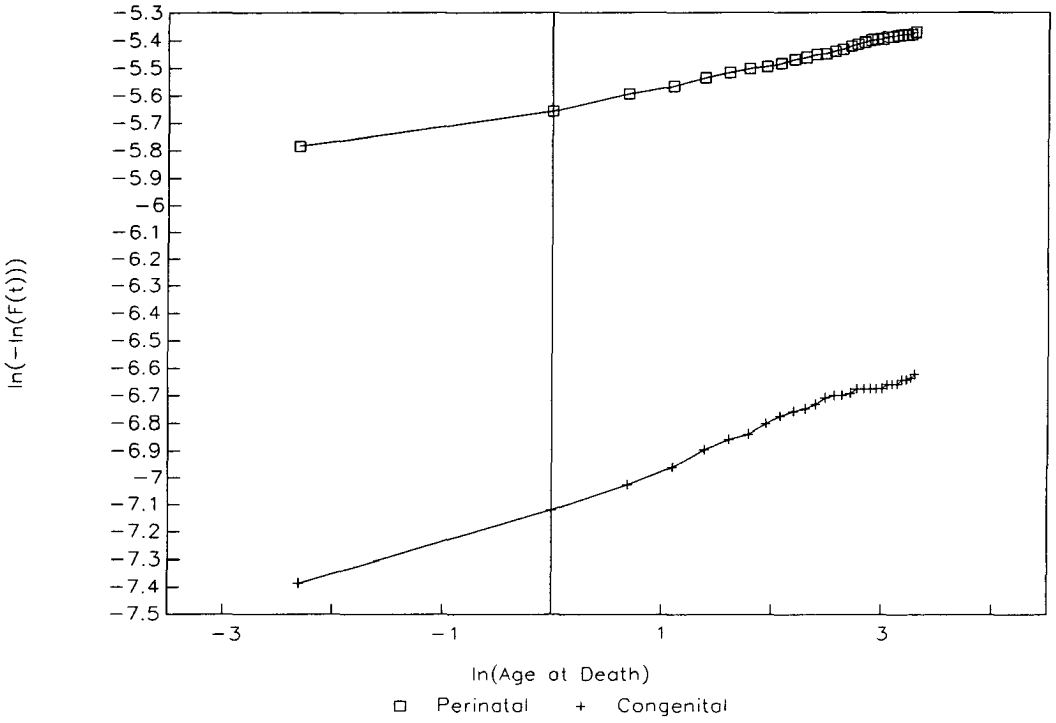


Figure 3. ln(-ln(F(t))) versus ln(age at death) (stratified by cause of death).

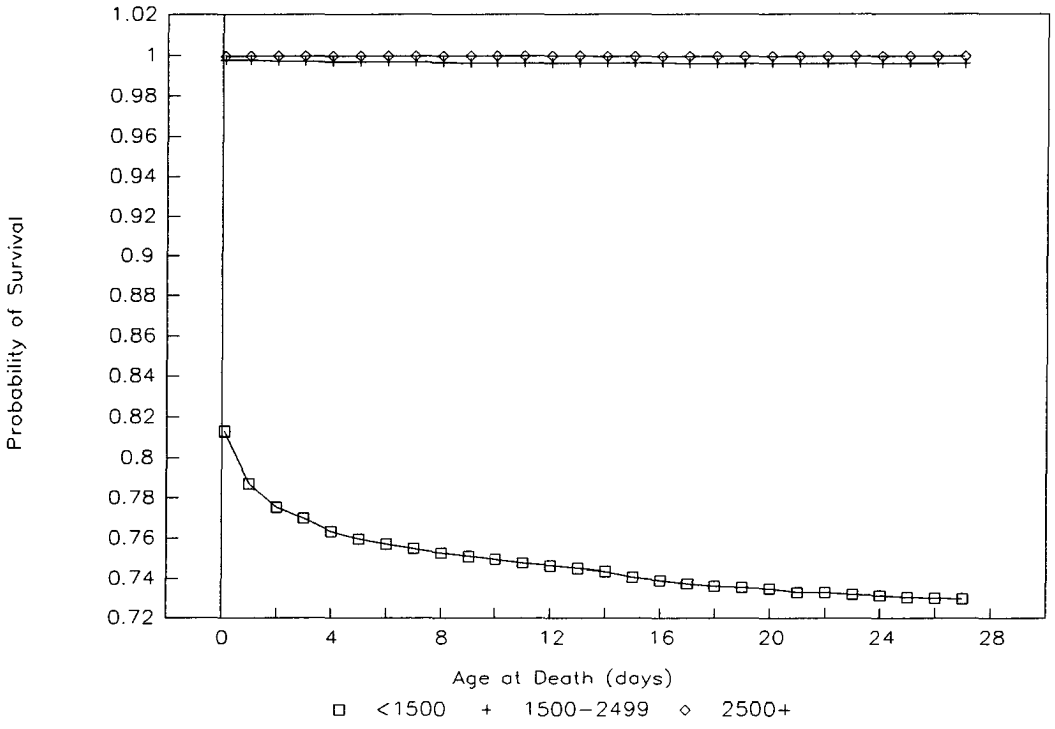


Figure 4. Probability of survival versus age at death (stratified by WEIGHT) for perinatal conditions.

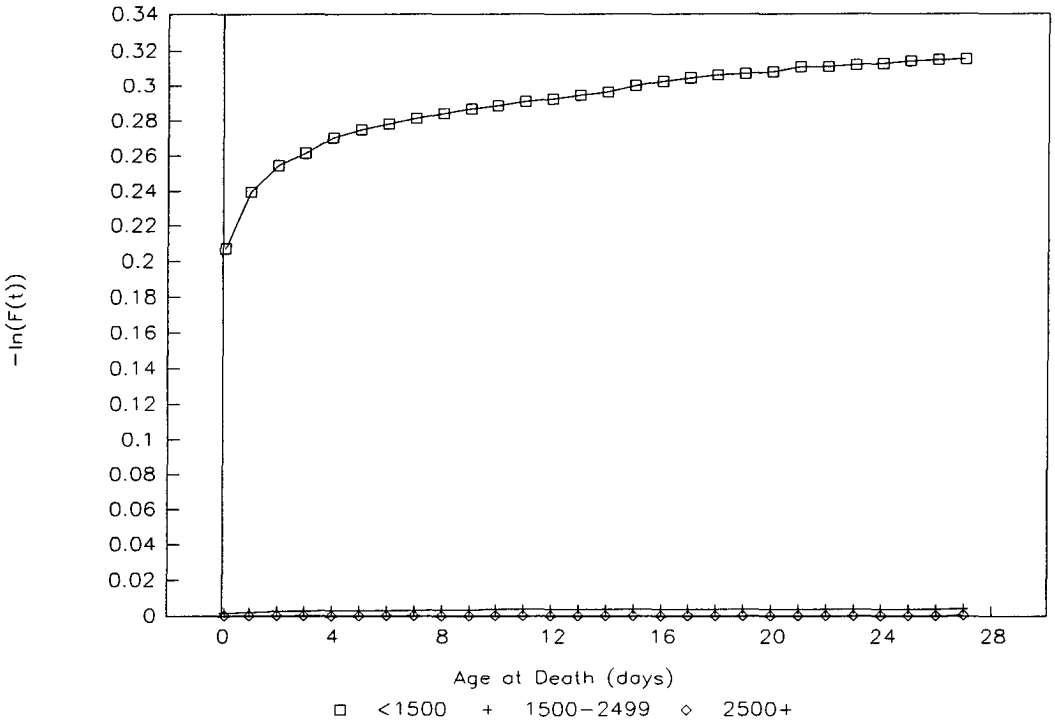


Figure 5. -ln(F(t)) versus age at death (stratified by WEIGHT) for perinatal conditions.

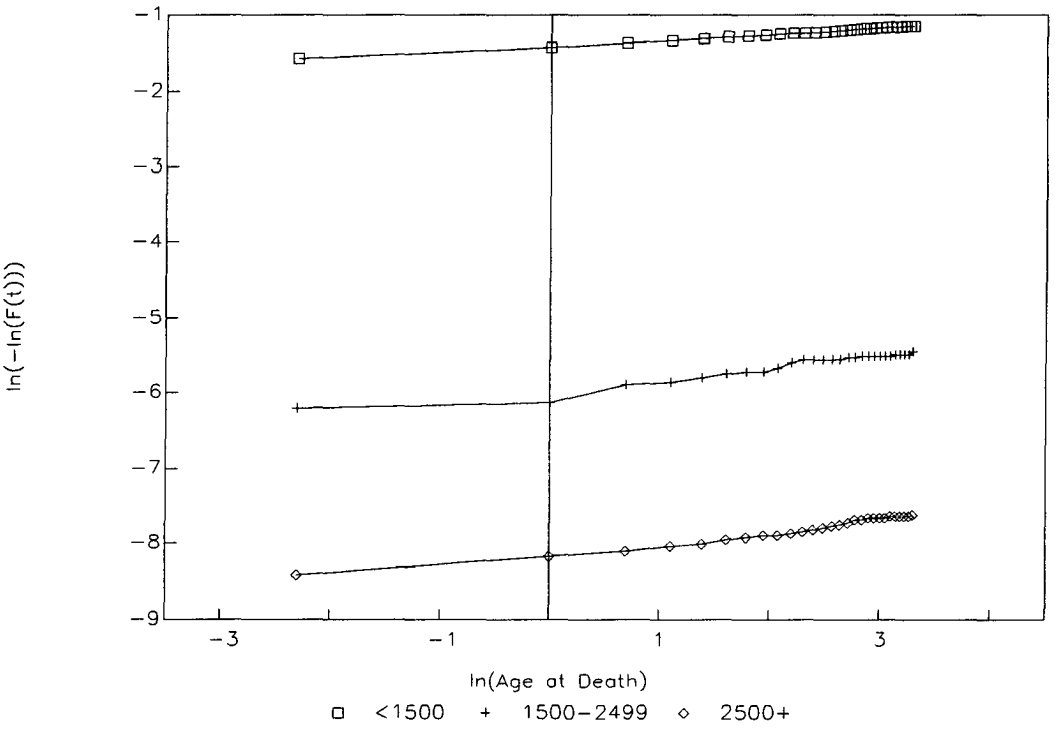


Figure 6. ln(-ln(F(t))) versus ln(age at death) (stratified by WEIGHT) for perinatal conditions.

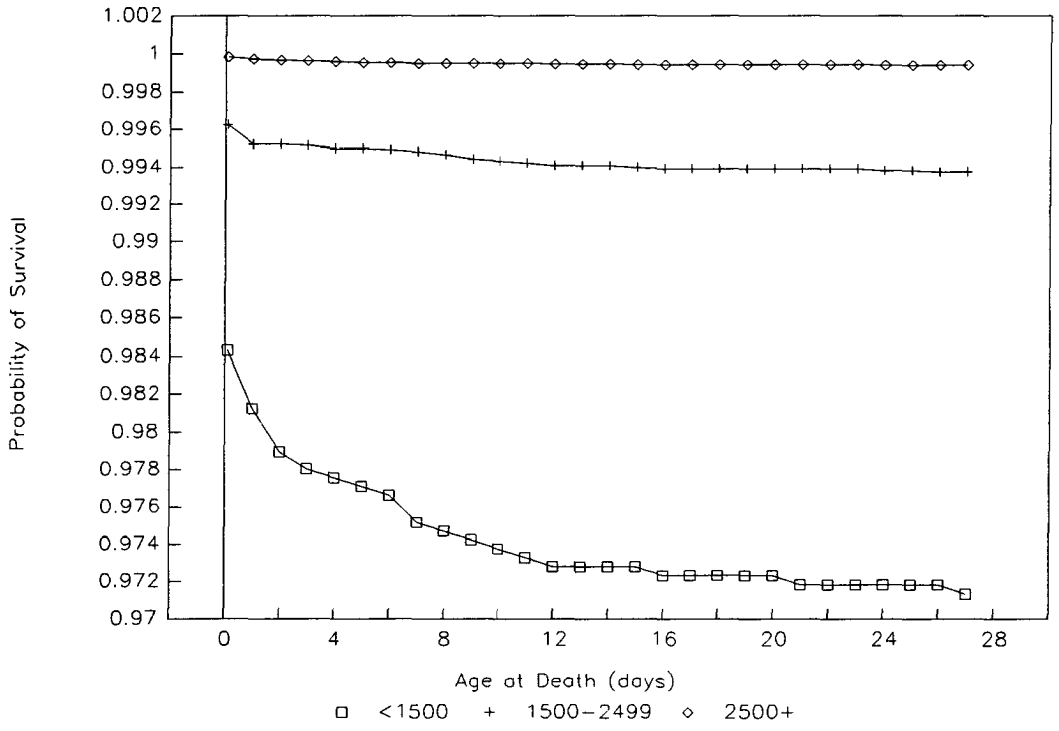


Figure 7. Probability of survival versus age at death (stratified by WEIGHT) for congenital anomalies.

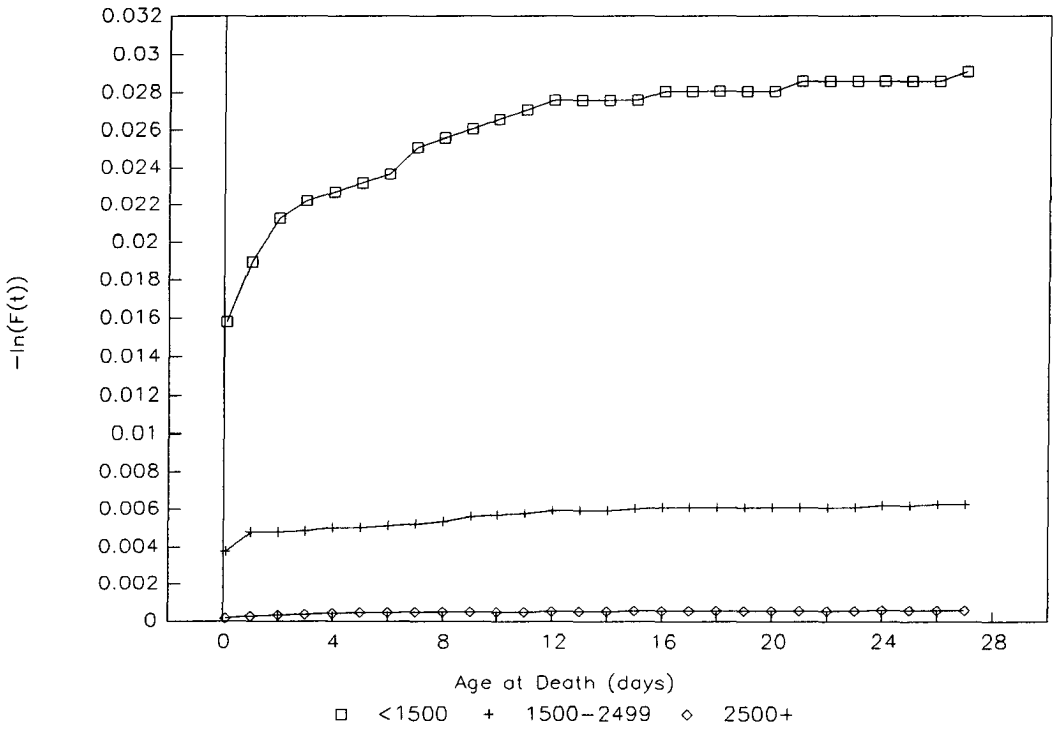


Figure 8. -ln(F(t)) versus age at death (stratified by WEIGHT) for congenital anomalies.

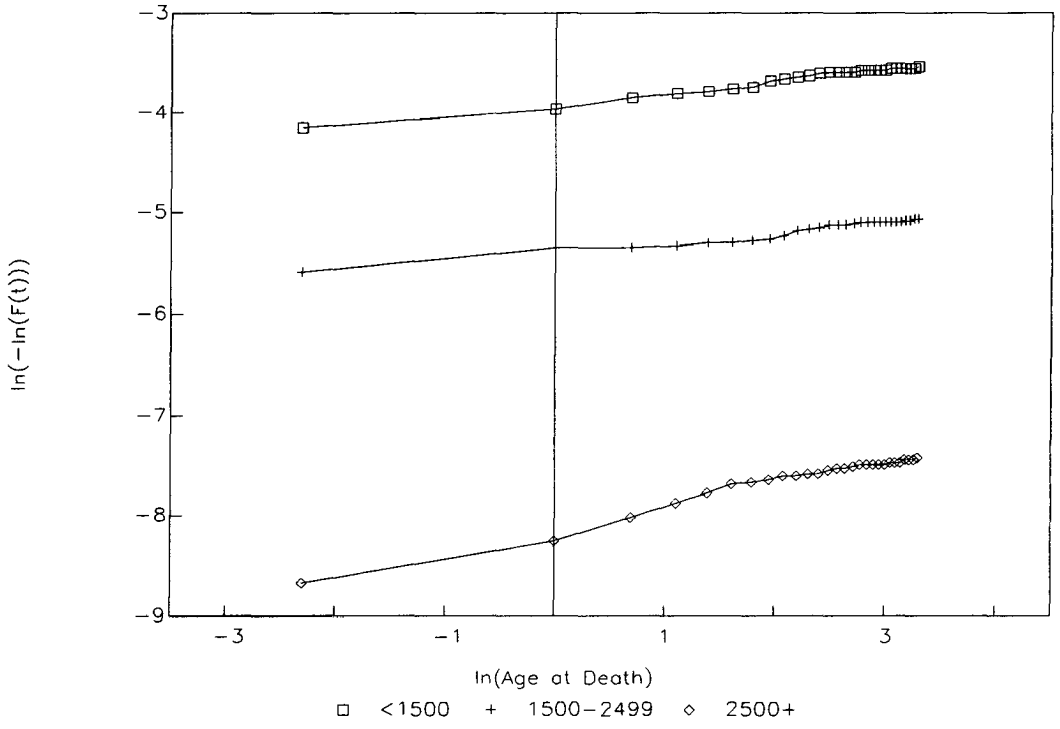


Figure 9. ln(-ln(F(t))) versus ln(age at death) (stratified by WEIGHT) for congenital anomalies.

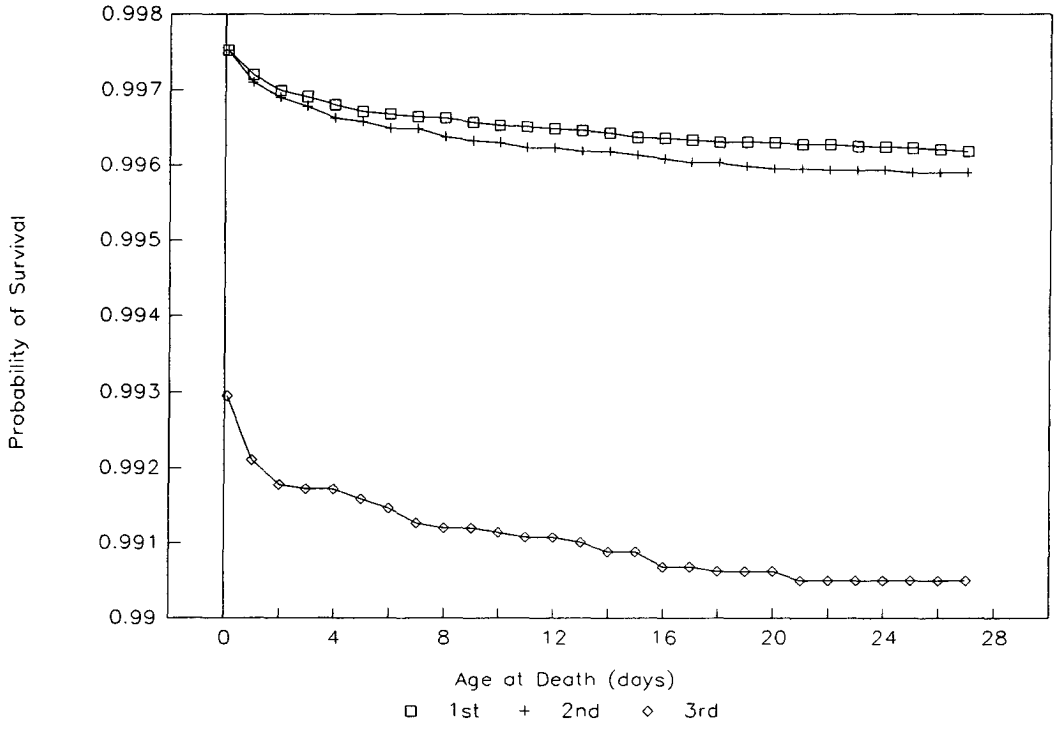


Figure 10. Probability of survival versus age at death (stratified by CARE) for perinatal conditions.

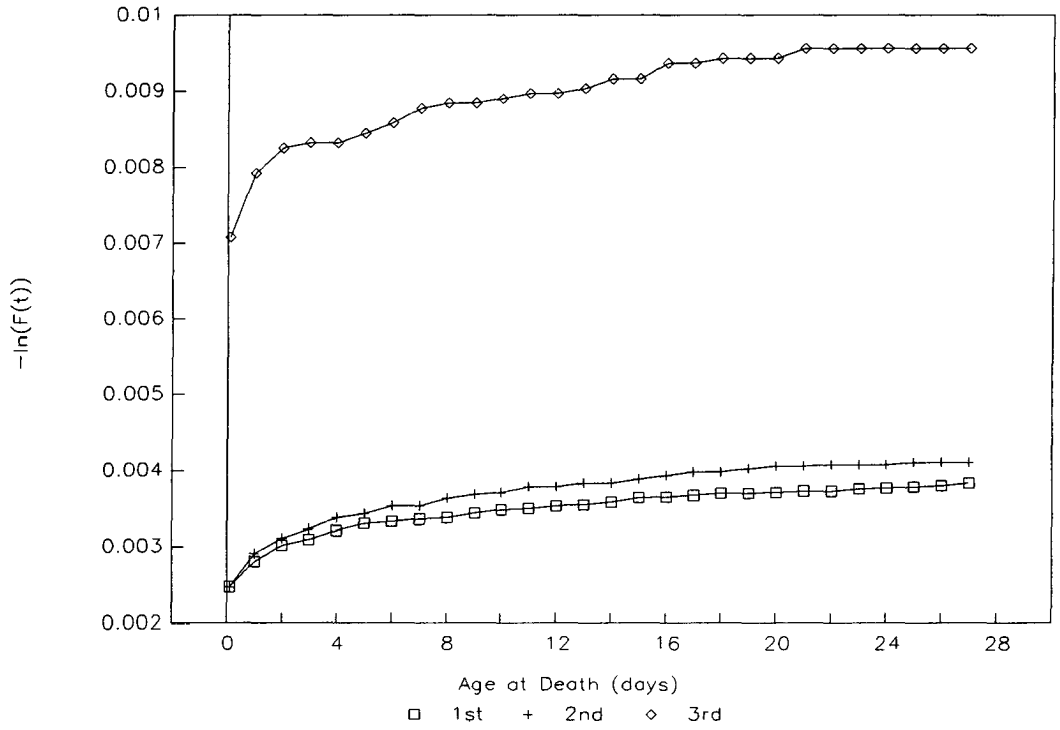


Figure 11. -ln(F(t)) versus age at death (stratified by CARE) for perinatal conditions.

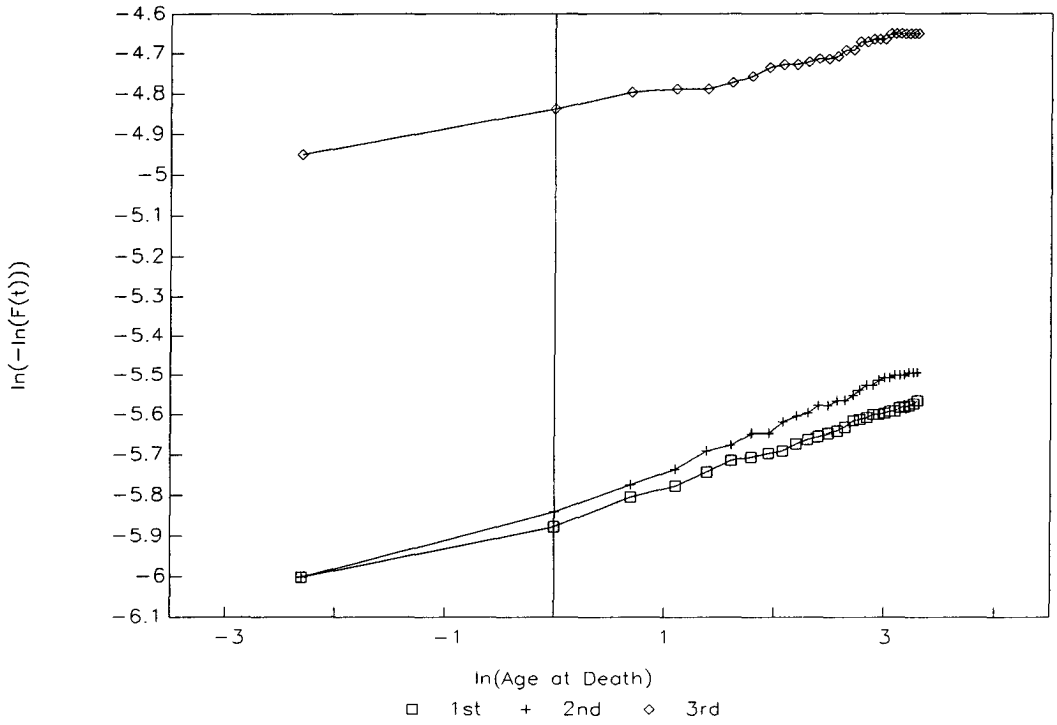


Figure 12. ln(-ln(F(t))) versus ln(age at death) (stratified by CARE) for perinatal conditions.

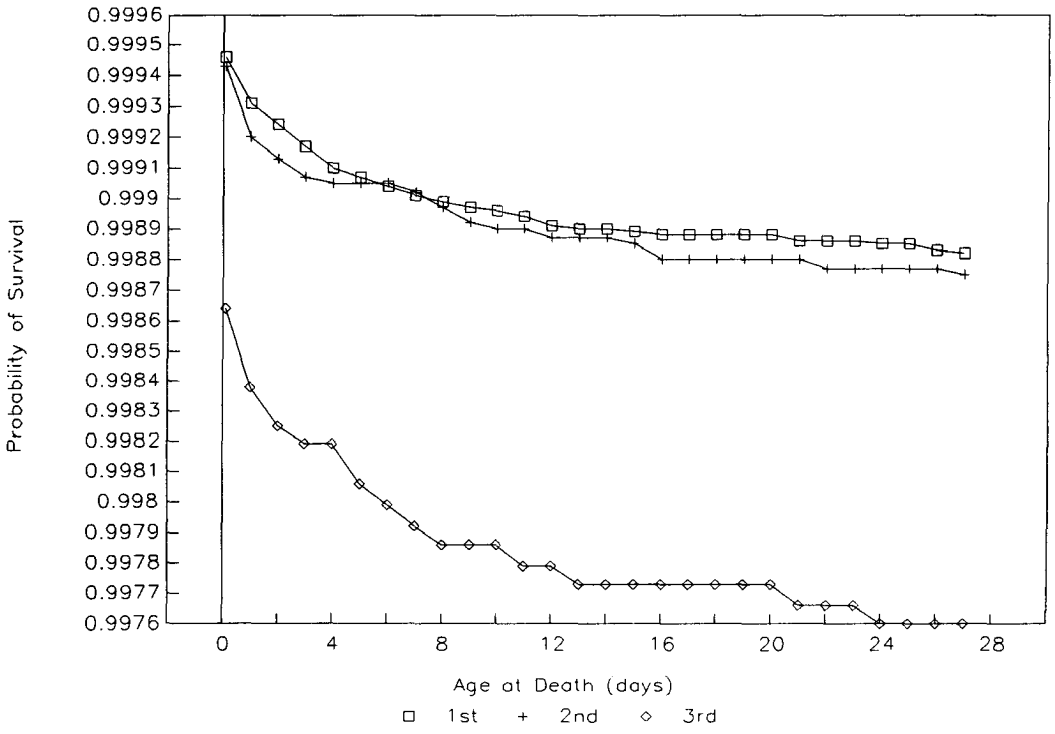


Figure 13. Probability of survival versus age at death (stratified by CARE) for congenital anomalies.

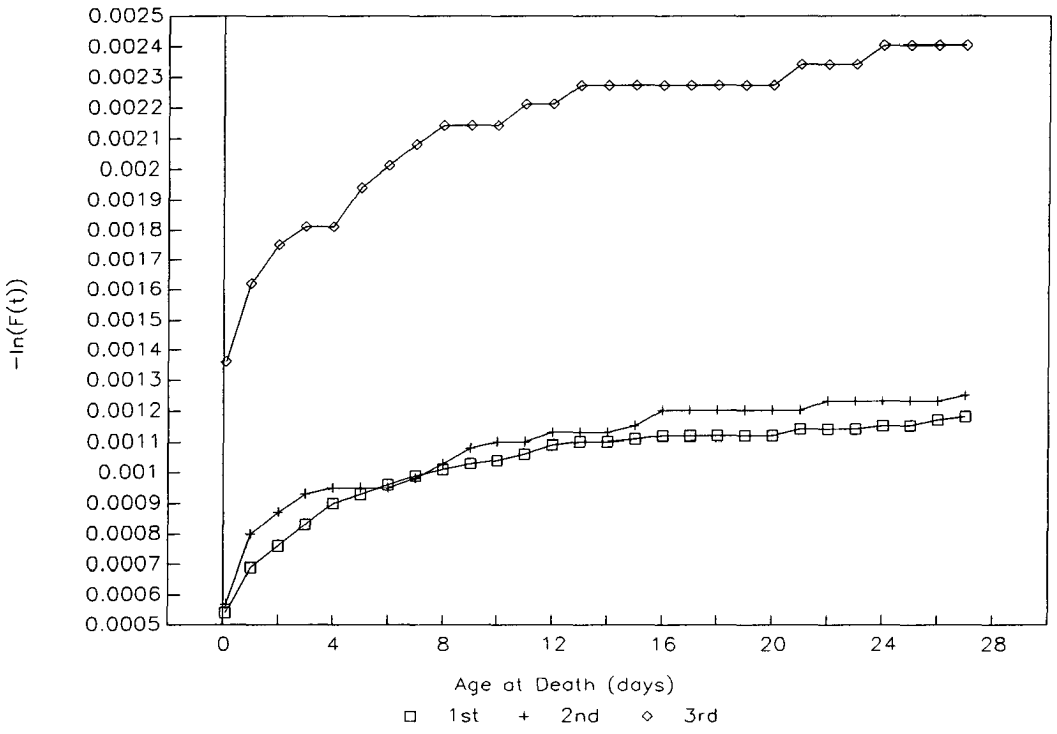


Figure 14. -ln(F(t)) versus age at death (stratified by CARE) for congenital anomalies.

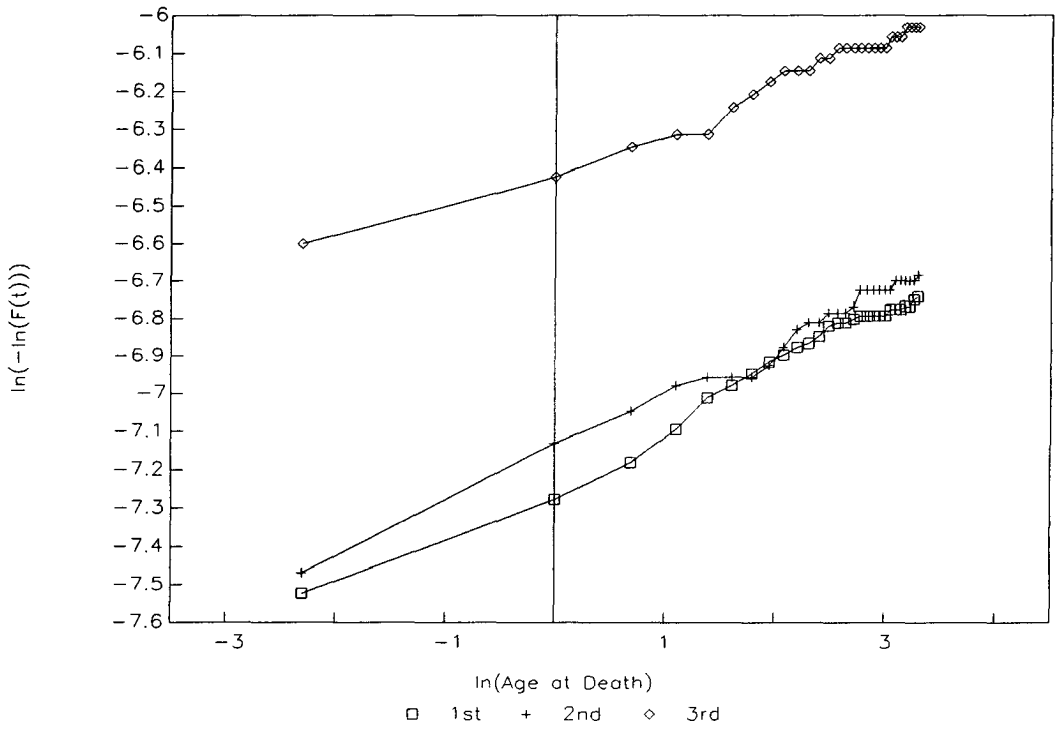


Figure 15. ln(-ln(F(t))) versus ln(age at death) (stratified by CARE) for congenital anomalies.

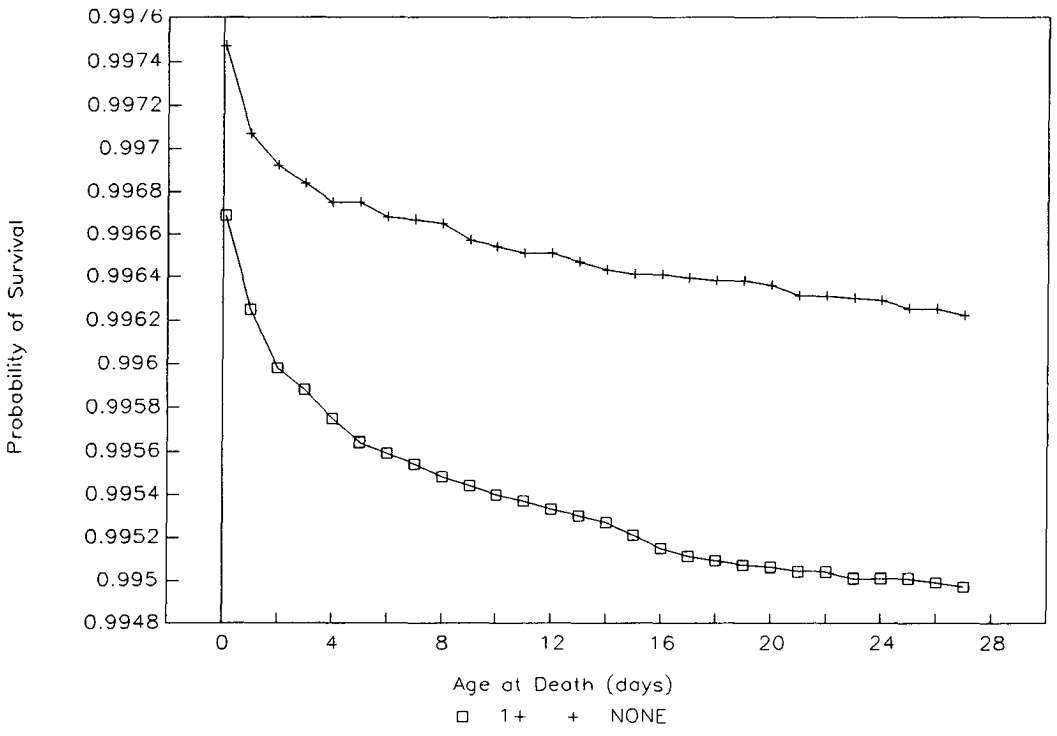


Figure 16. Probability of survival versus age at death (stratified by PREG) for perinatal conditions.

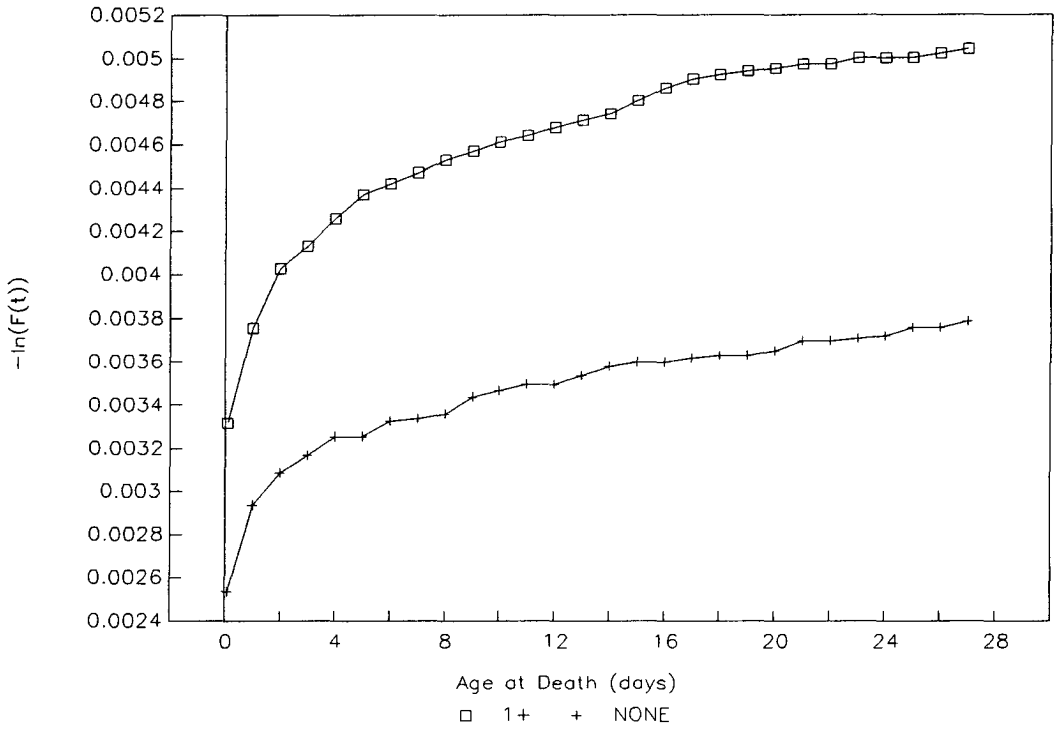


Figure 17. -ln(F(t)) versus age at death (stratified by PREG) for perinatal conditions.

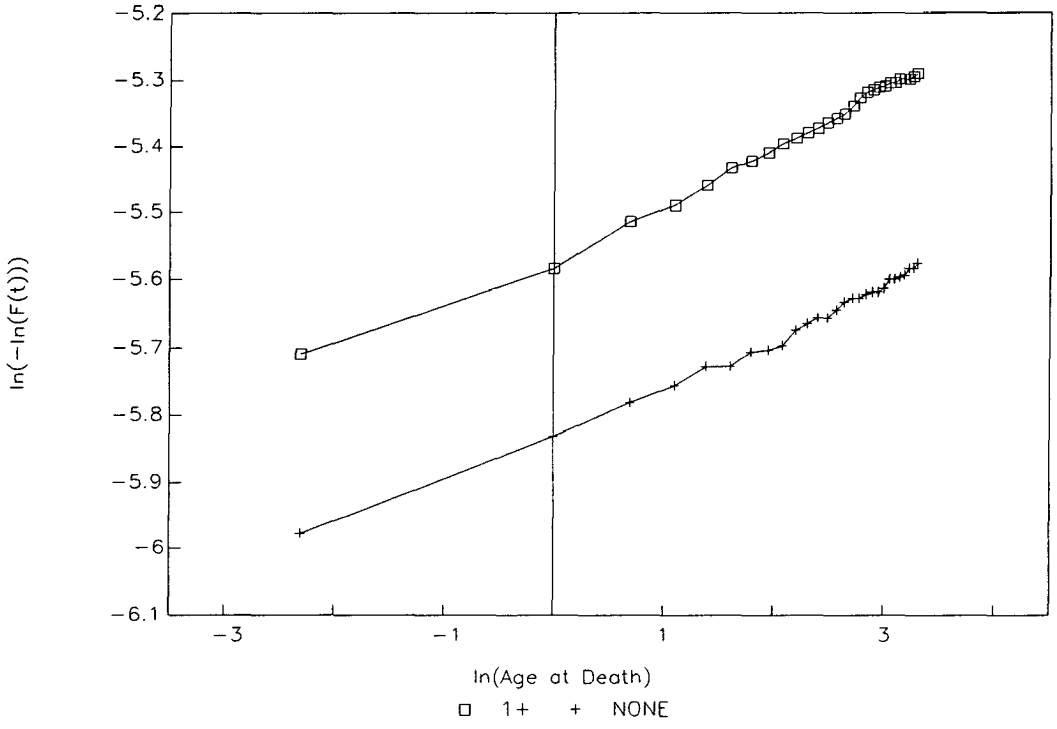


Figure 18. ln(-ln(F(t))) versus ln(age at death) (stratified by PREG) for perinatal conditions.

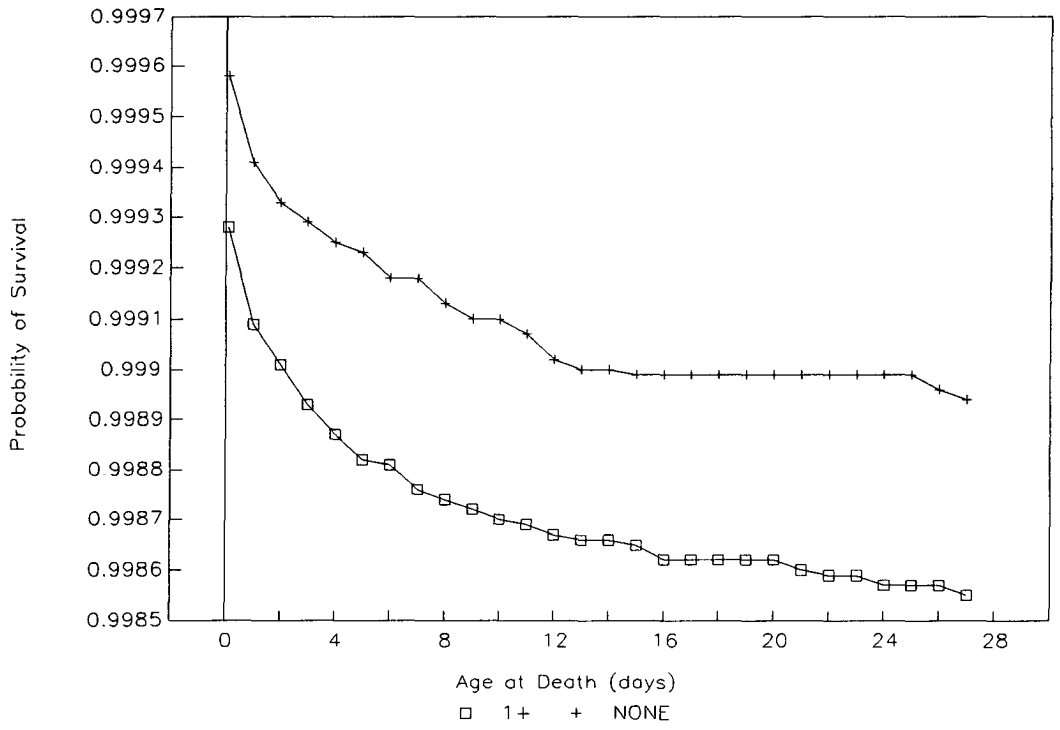


Figure 19. Probability of survival versus age at death (stratified by PREG) for congenital anomalies.

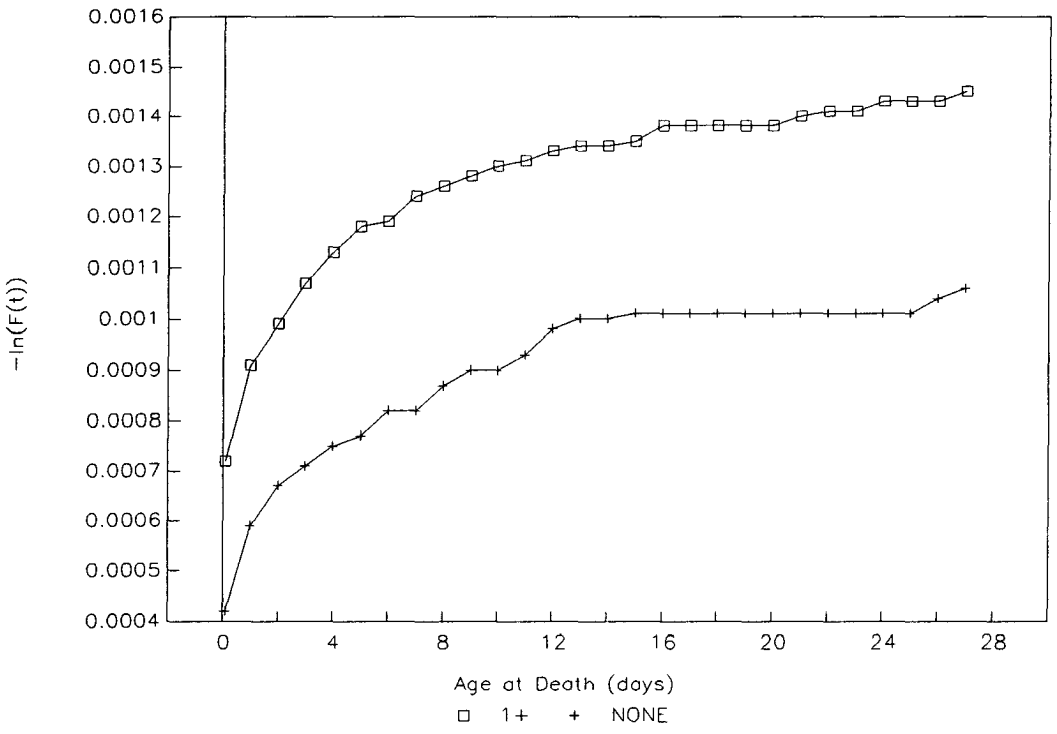


Figure 20. -ln(F(t)) versus age at death (stratified by PREG) for congenital anomalies.

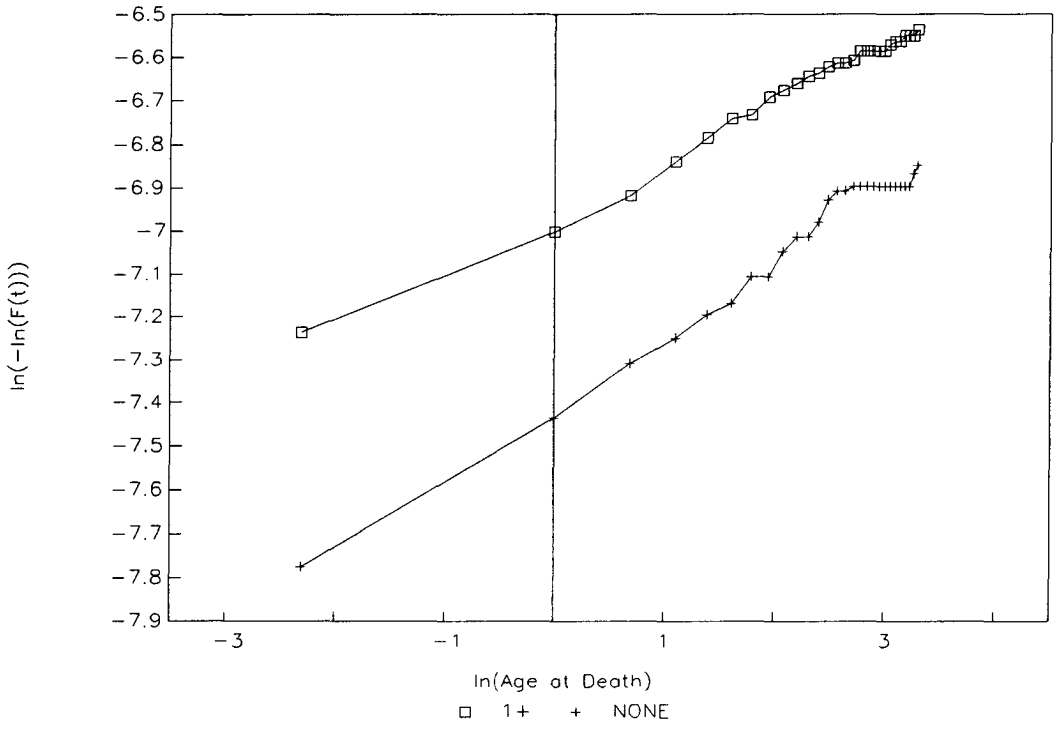
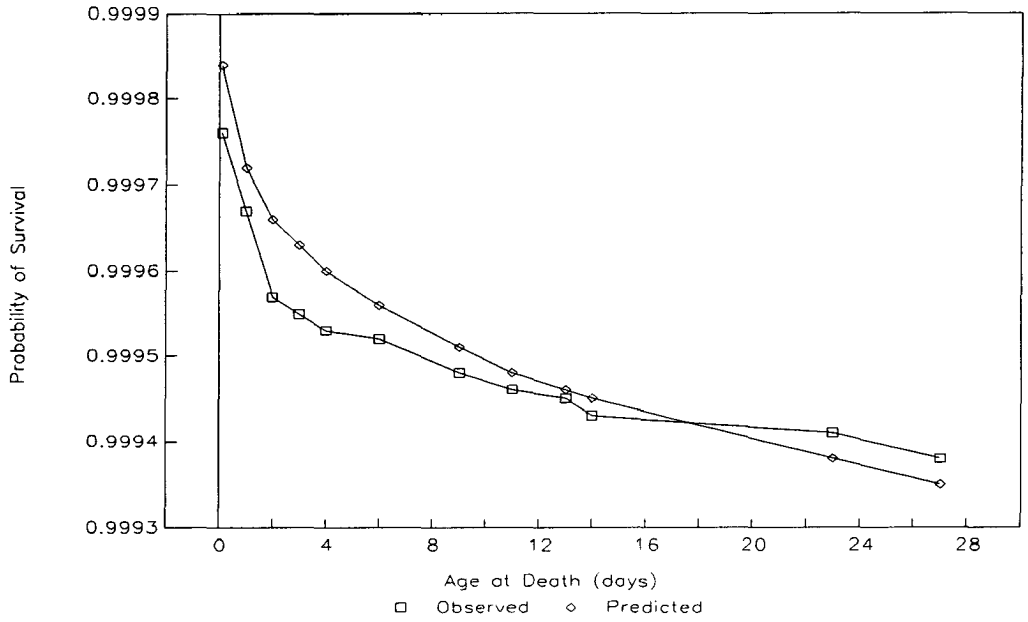
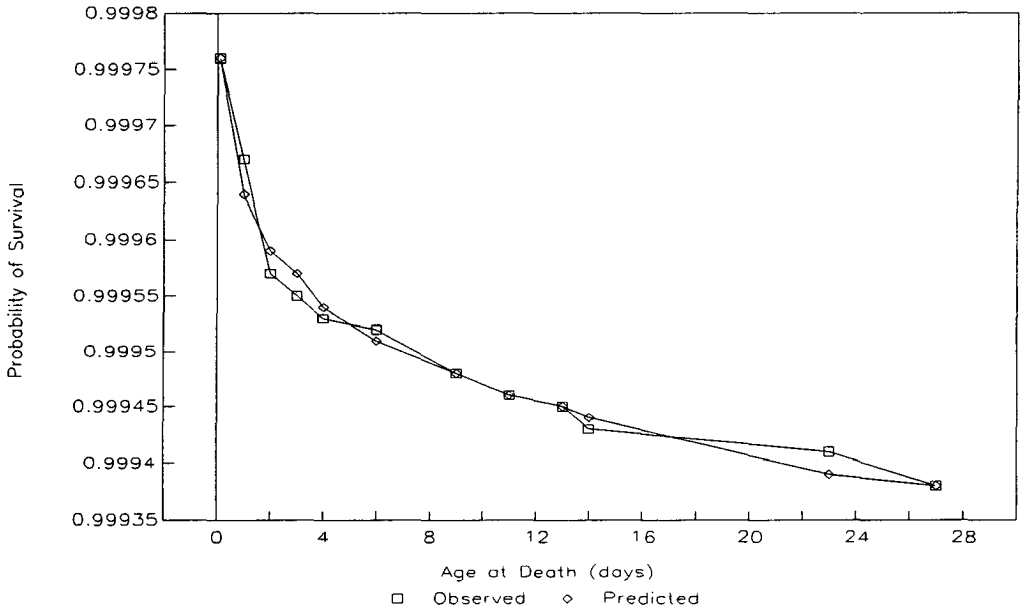


Figure 21. ln(-ln(F(t))) versus ln(age at death) (stratified by PREG) for congenital anomalies.

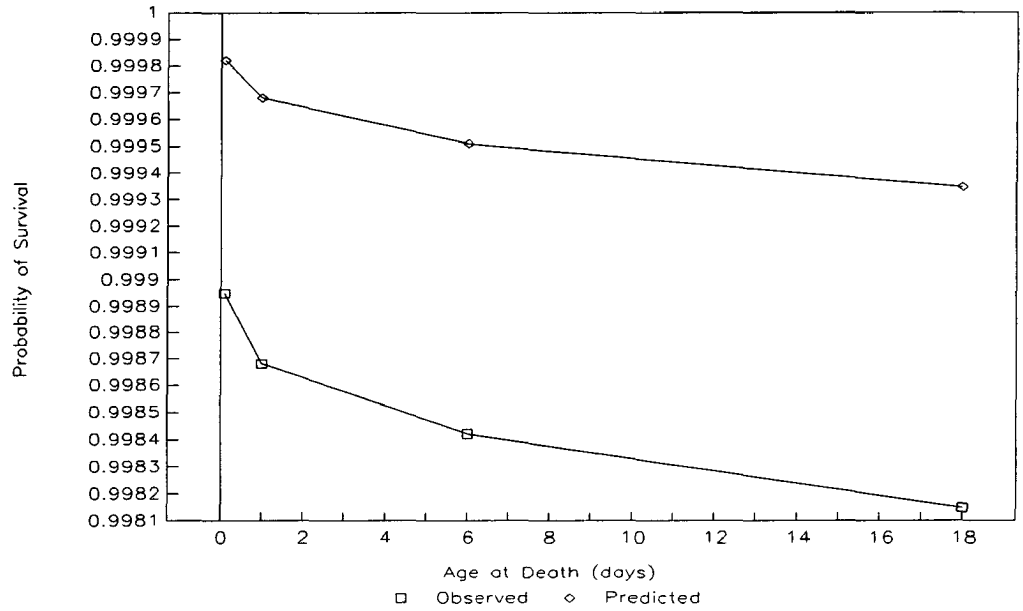


(a) Weibull model

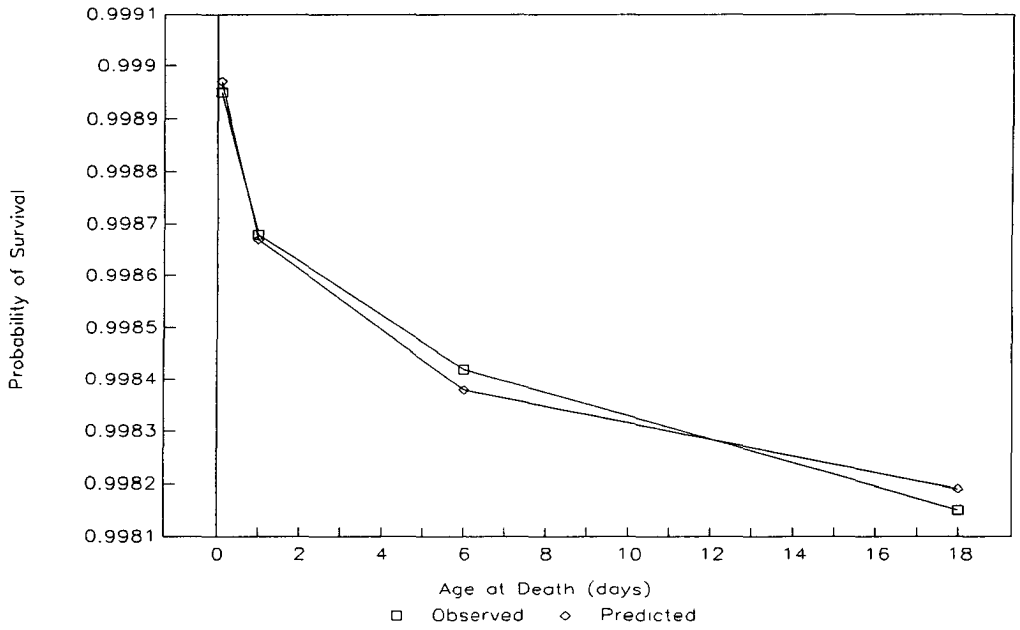


(b) Extreme Value model

Figure 22. Probability of survival with variable levels 1500+ WEIGHT, no PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

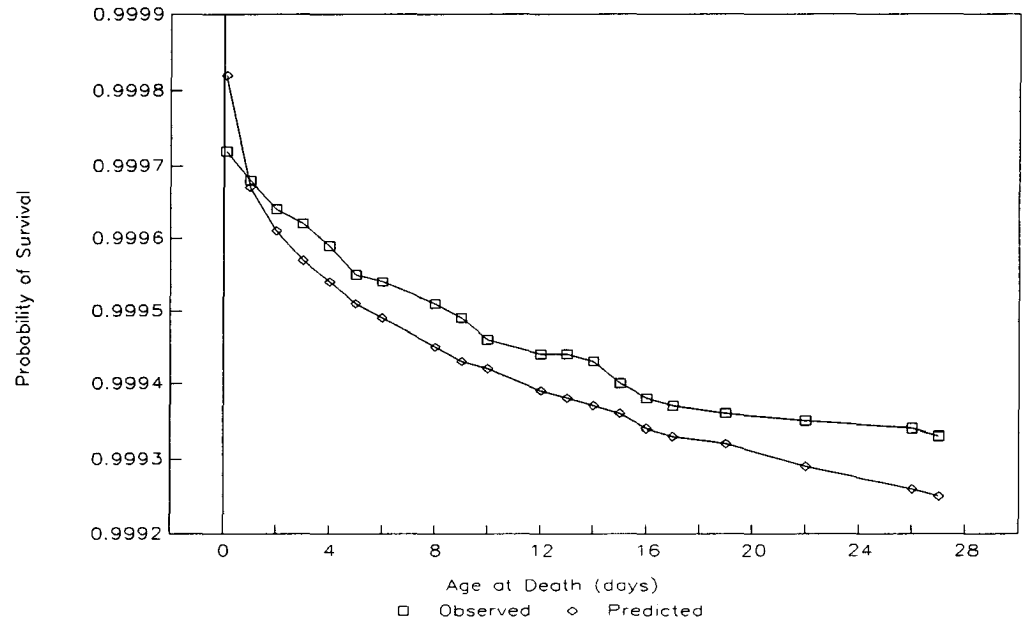


(a) Weibull model

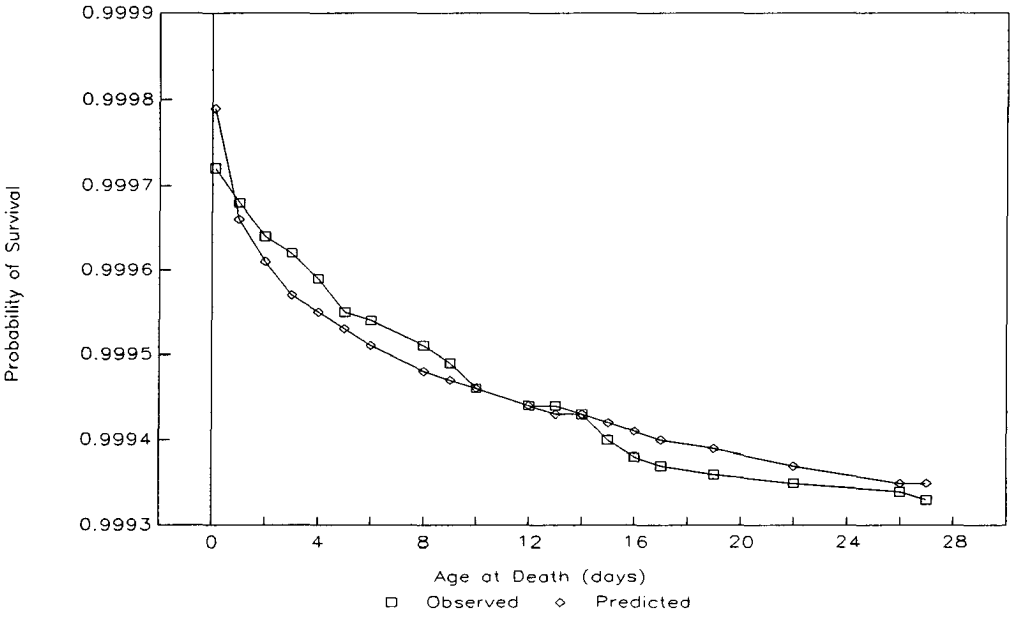


(b) Extreme Value model

Figure 23. Probability of survival with variable levels 1500+ WEIGHT, no PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions

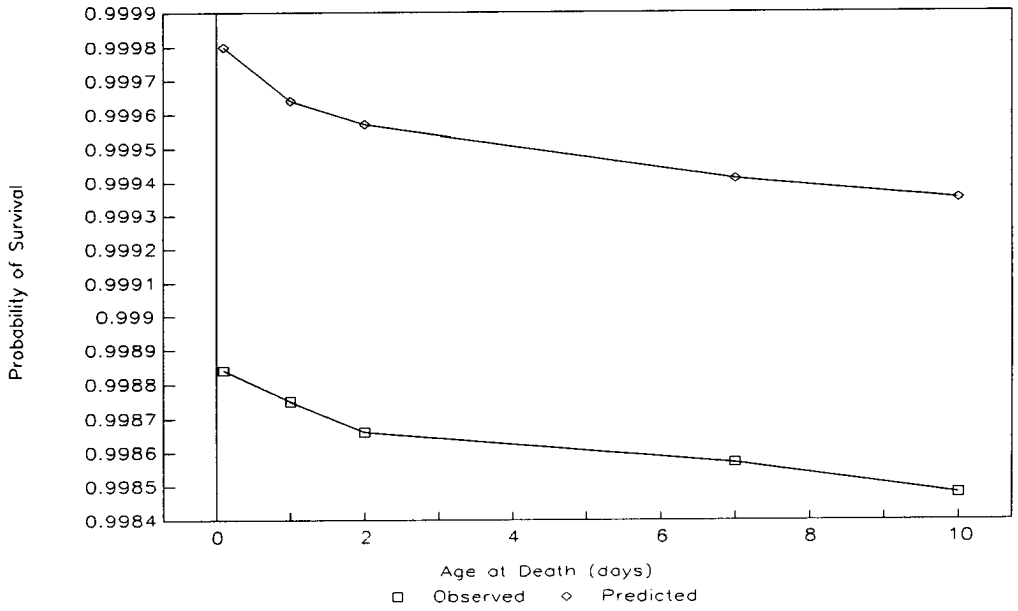


(a) Weibull model

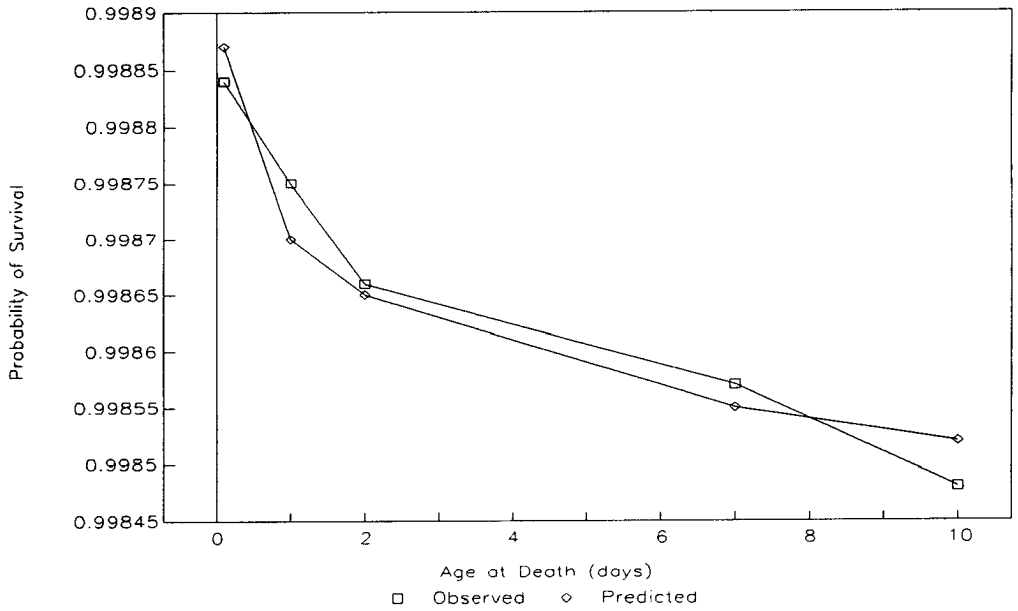


(b) Extreme Value model

Figure 24. Probability of survival with variable levels 1500+ WEIGHT, 1+ PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

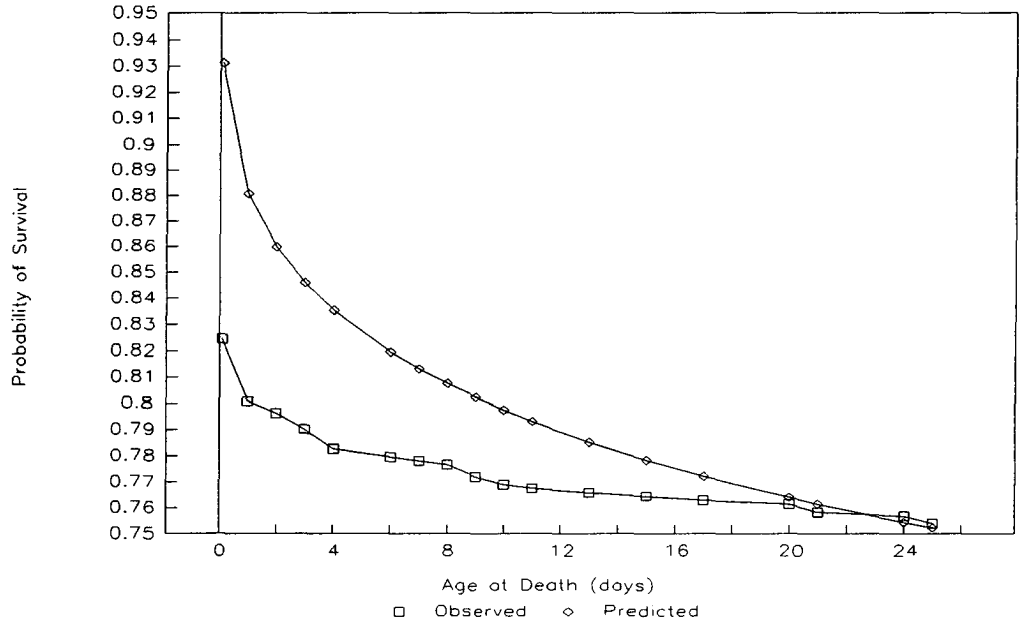


(a) Weibull model

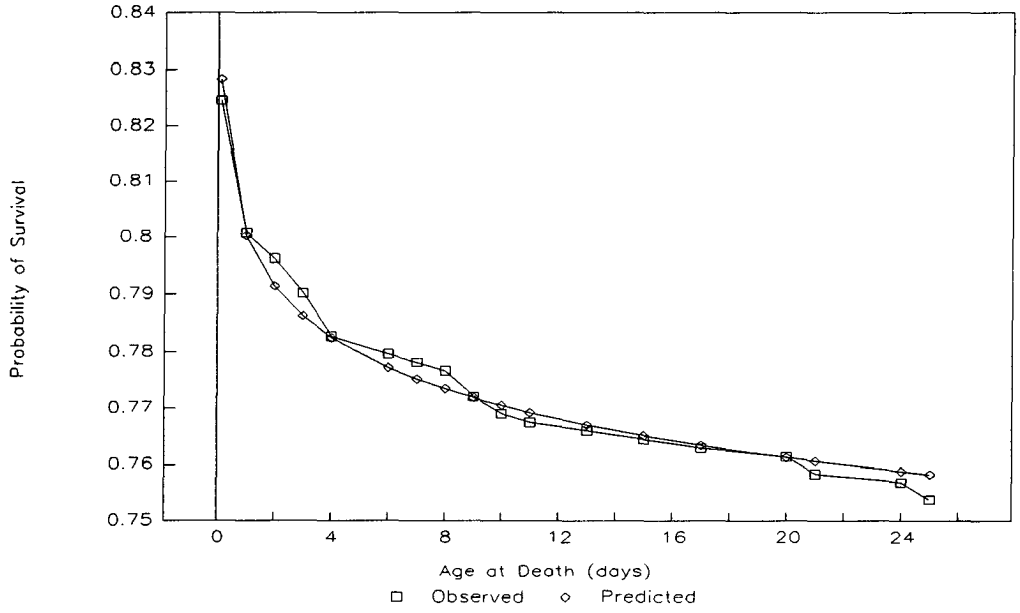


(b) Extreme Value model

Figure 25. Probability of survival with variable levels 1500+ WEIGHT, 1+ PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

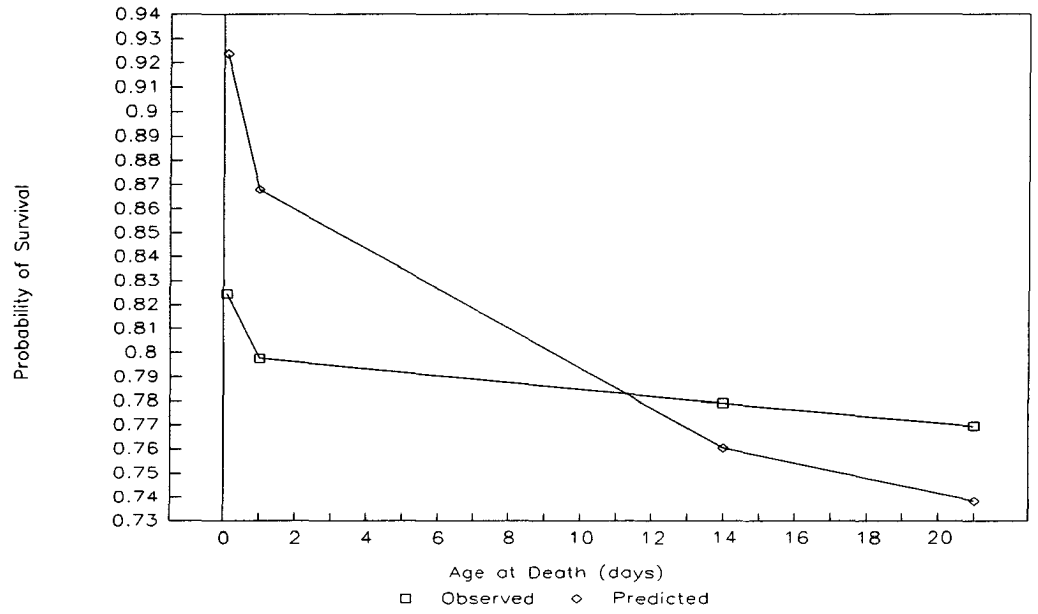


(a) Weibull model

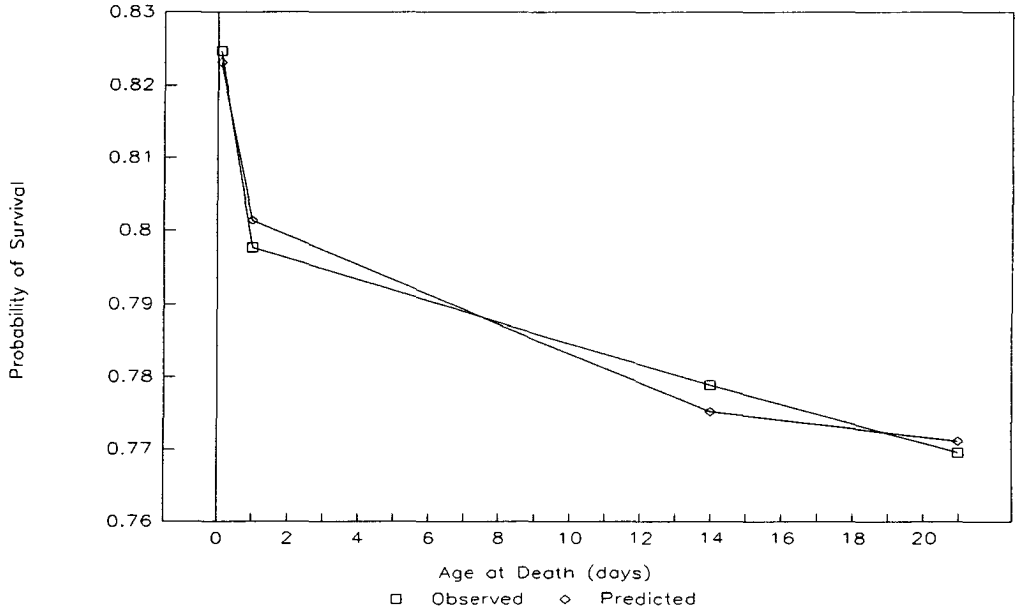


(b) Extreme Value model

Figure 26. Probability of survival with variable levels <1500 WEIGHT, no PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

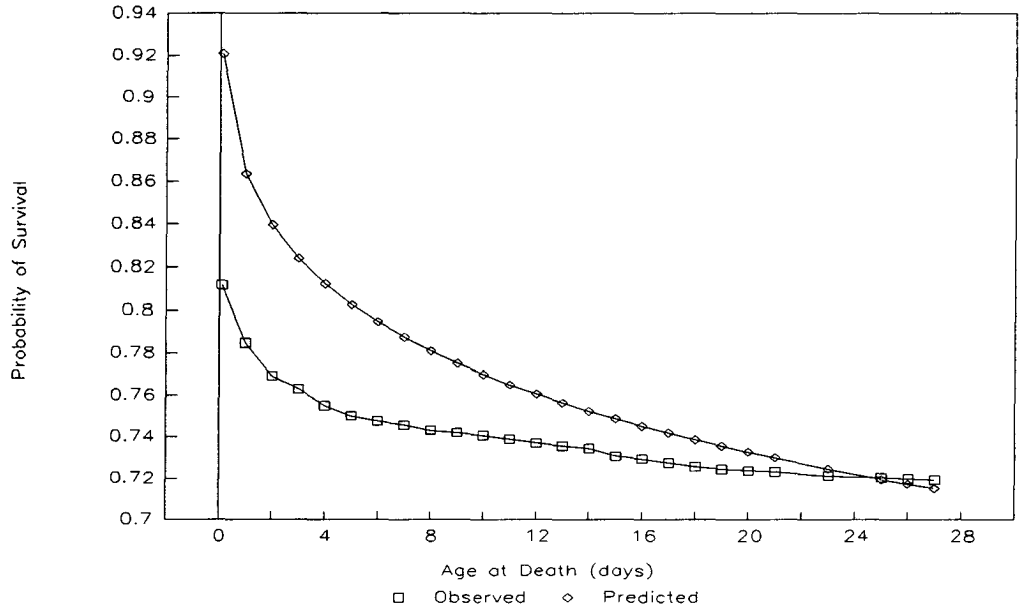


(a) Weibull model

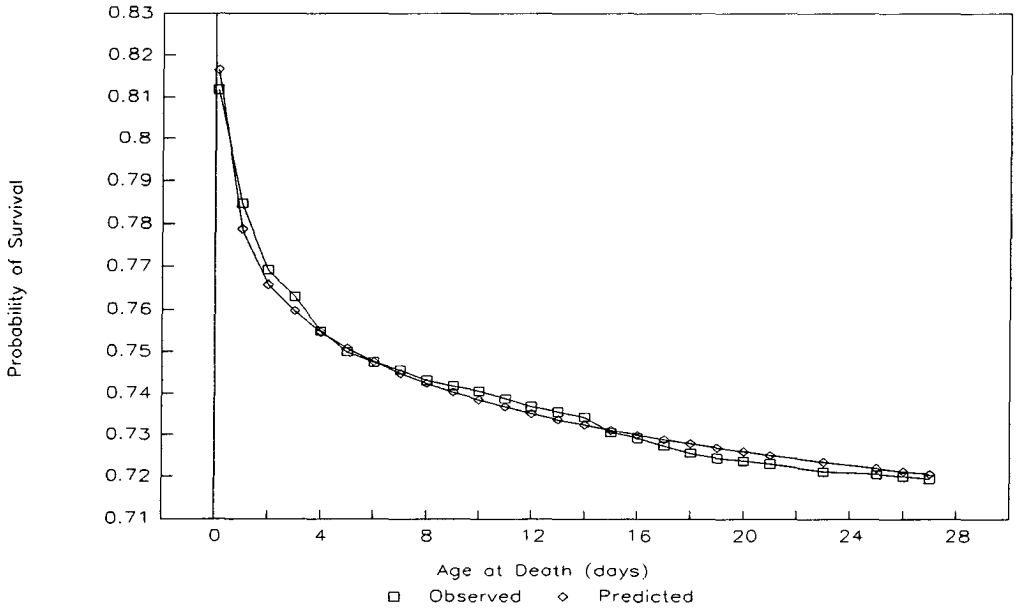


(b) Extreme Value model

Figure 27. Probability of survival with variable levels <1500 WEIGHT, no PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

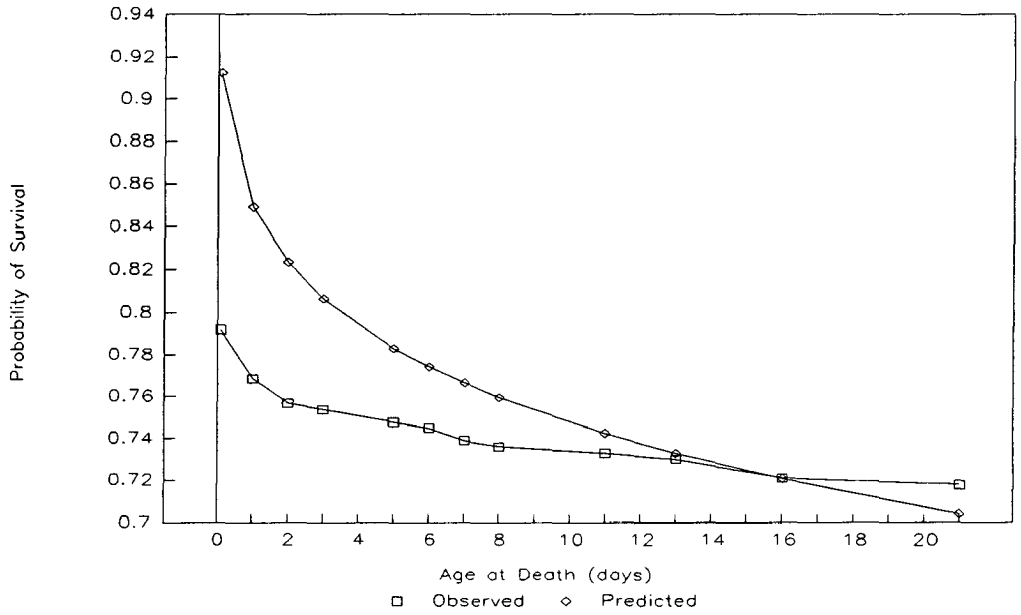


(a) Weibull model

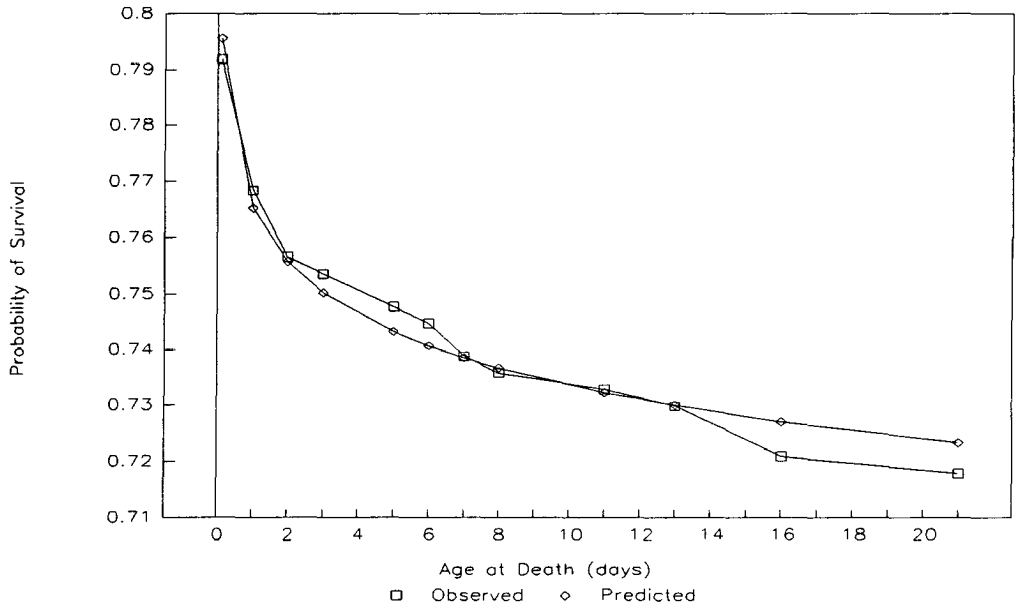


(b) Extreme Value model

Figure 28. Probability of survival with variable levels <1500 WEIGHT, 1+ PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

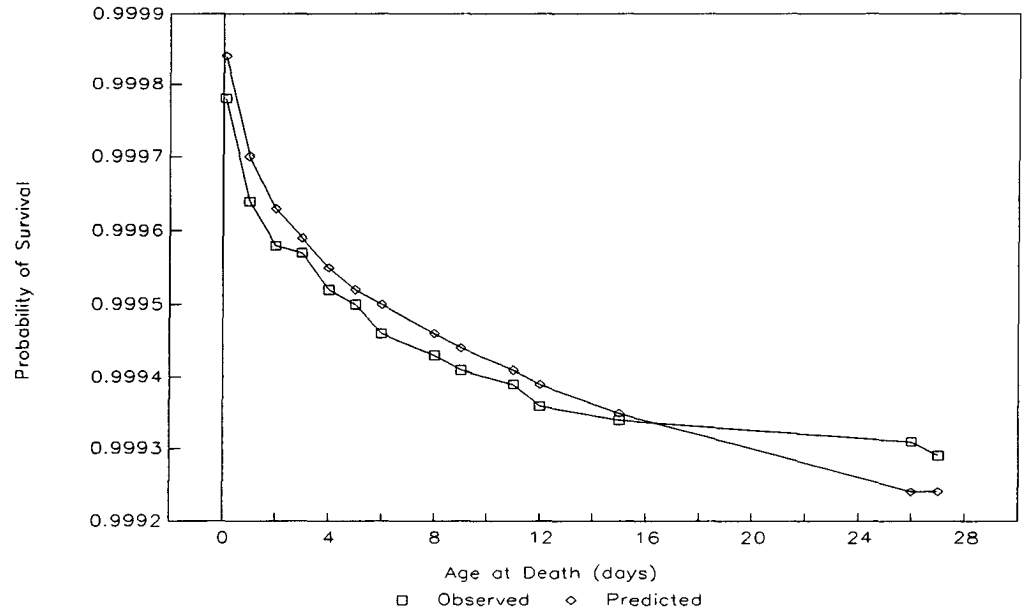


(a) Weibull model

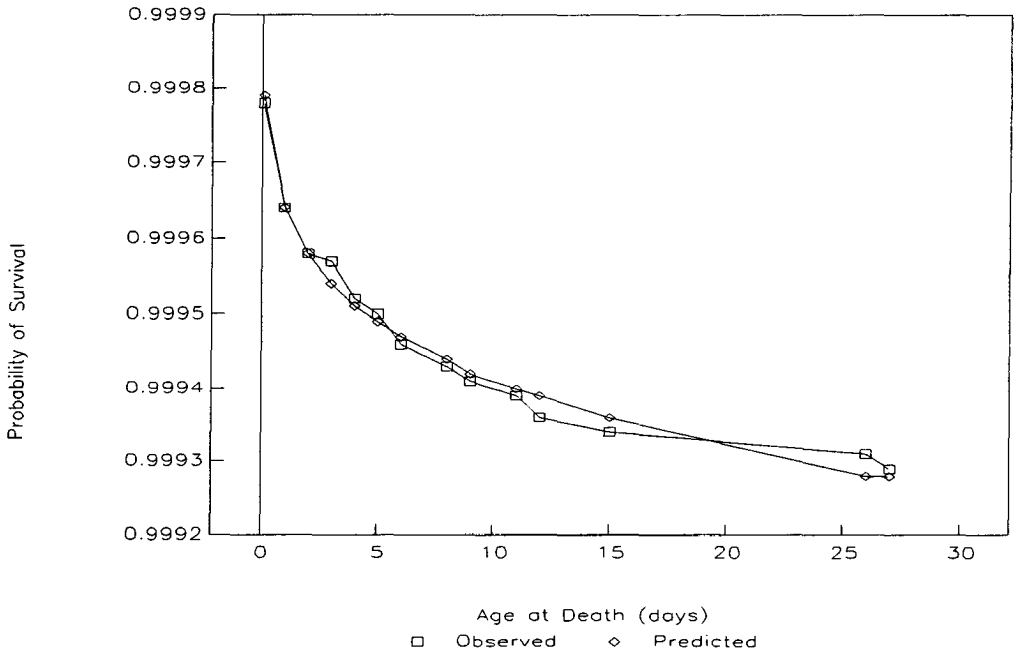


(b) Extreme Value model

Figure 29. Probability of survival with variable levels <1500 WEIGHT, 1+ PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for perinatal conditions.

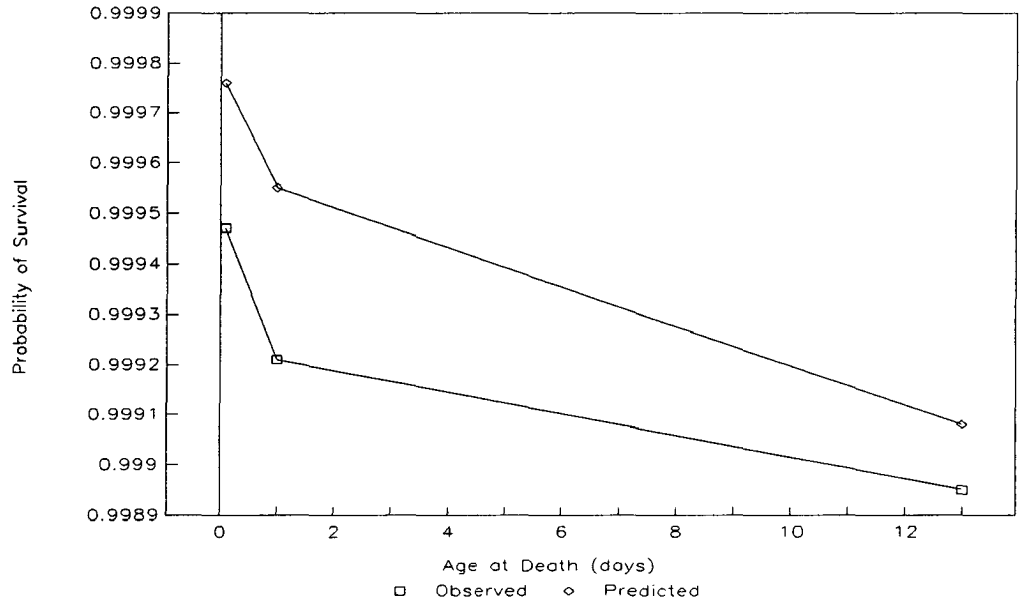


(a) Weibull model

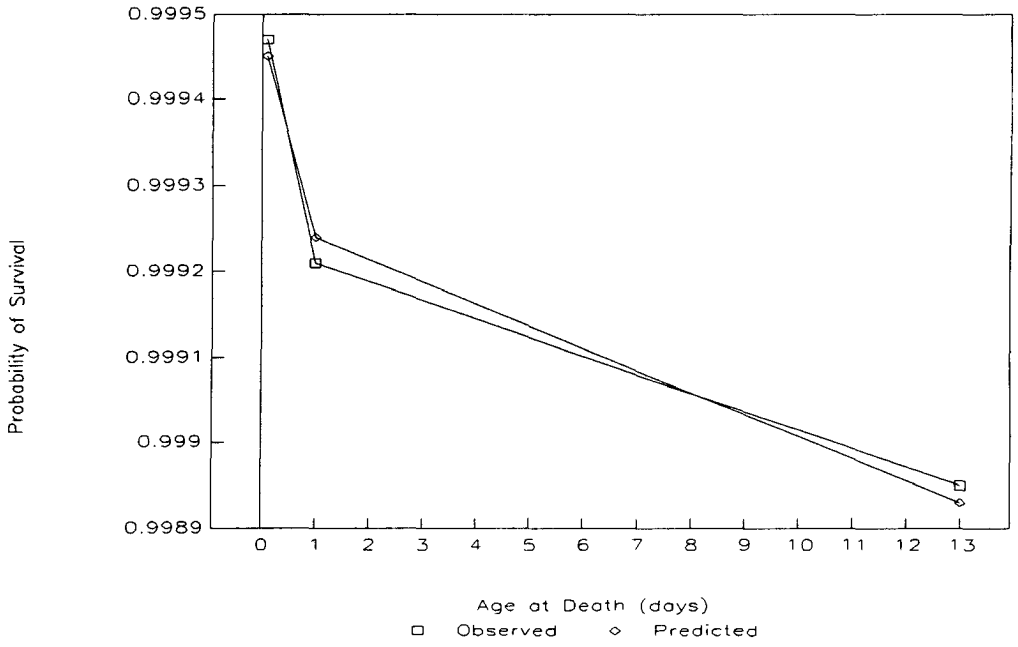


(b) Extreme Value model

Figure 30. Probability of survival with variable levels 1500+ WEIGHT, no PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

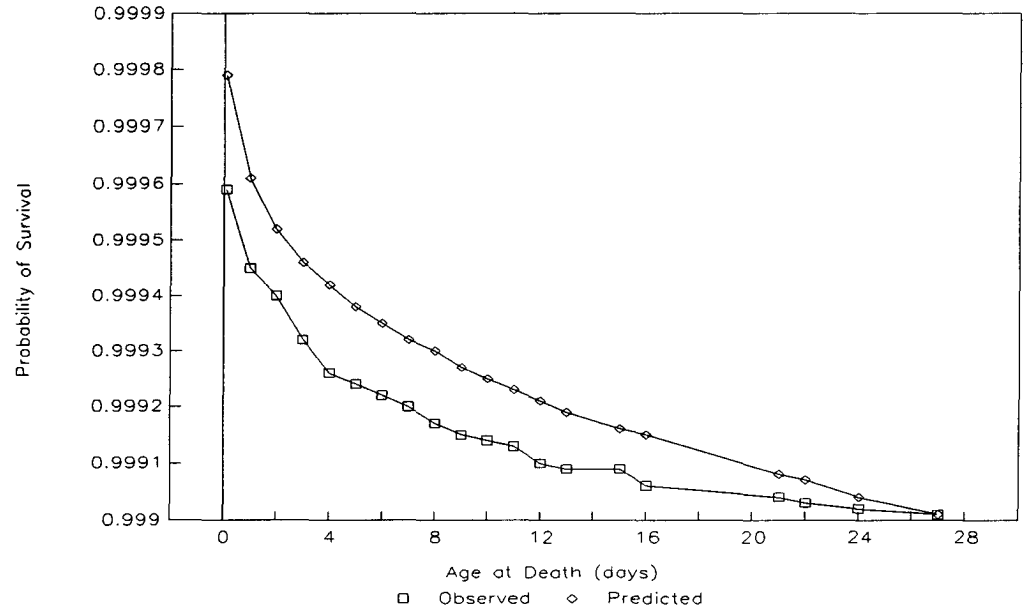


(a) Weibull model

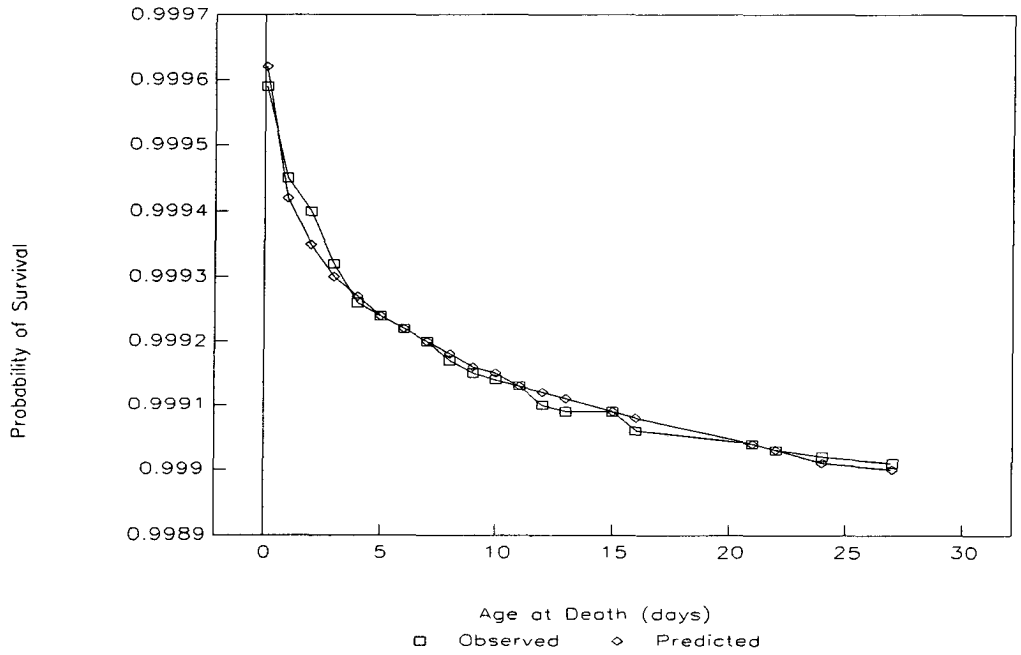


(b) Extreme Value model

Figure 31. Probability of survival with variable levels 1500+ WEIGHT, no PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

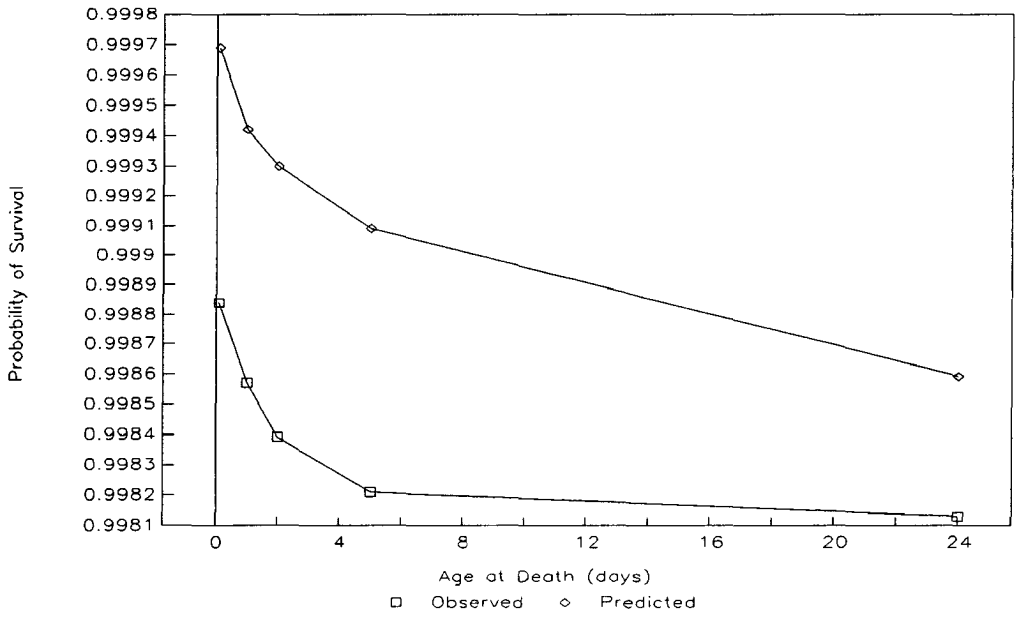


(a) Weibull model

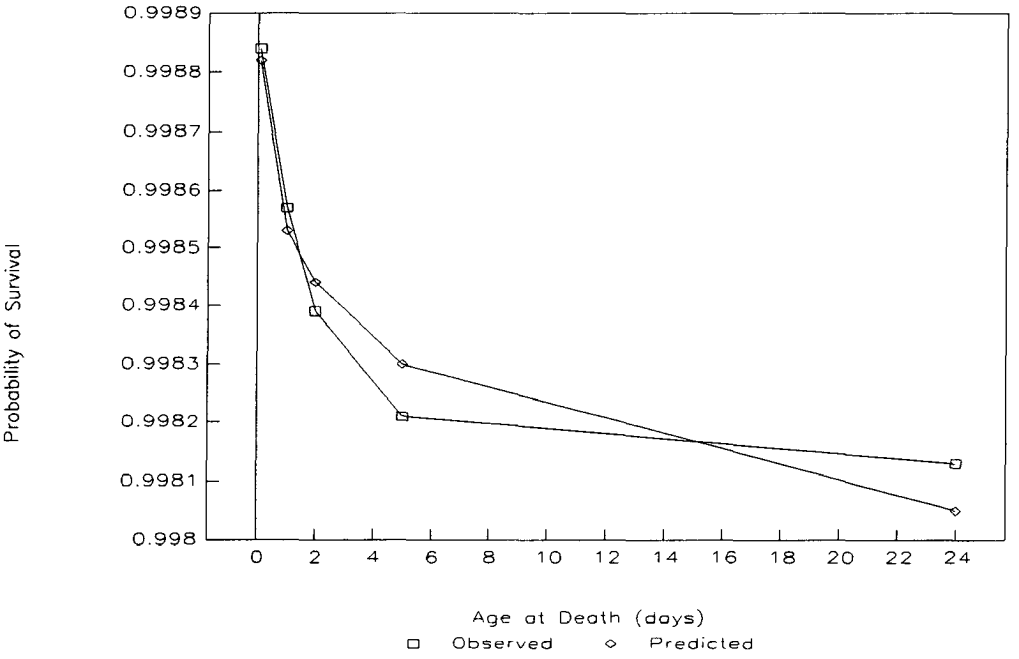


(b) Extreme Value model

Figure 32. Probability of survival with variable levels 1500+ WEIGHT, 1+ PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

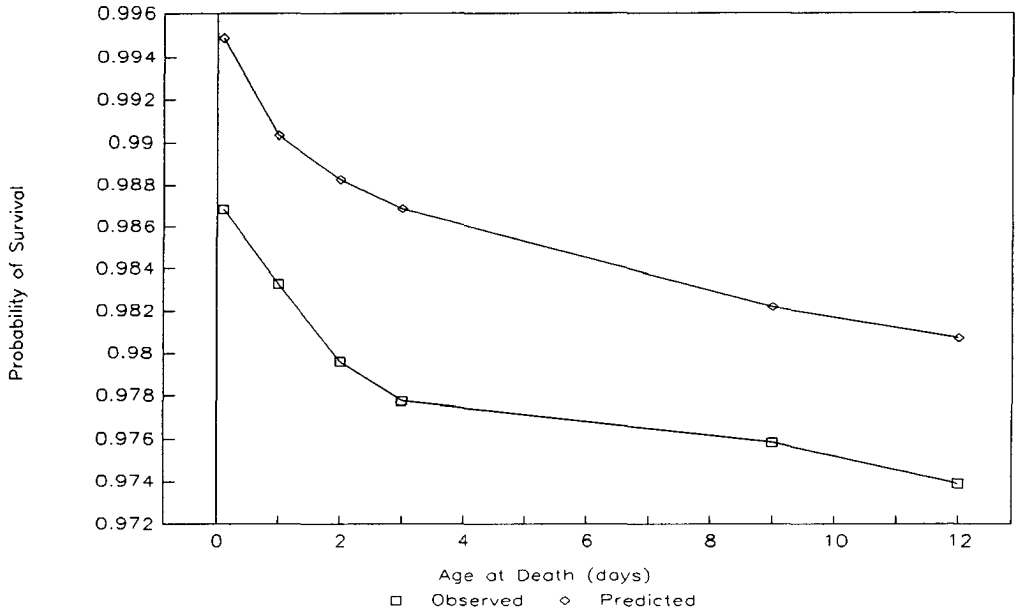


(a) Weibull model

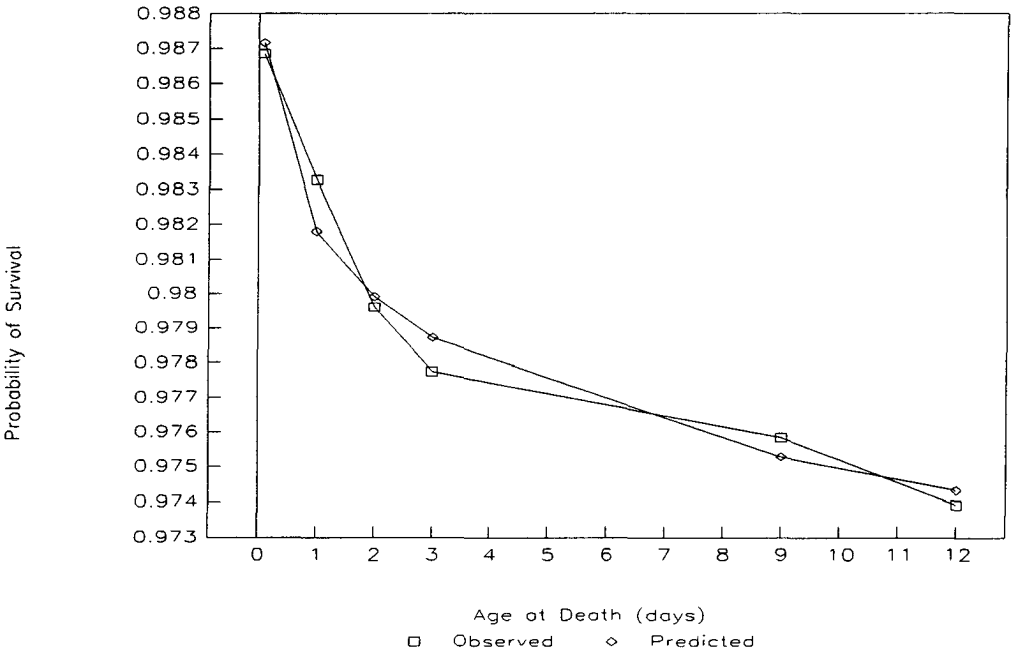


(b) Extreme Value model

Figure 33. Probability of survival with variable levels 1500+ WEIGHT, 1+ PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

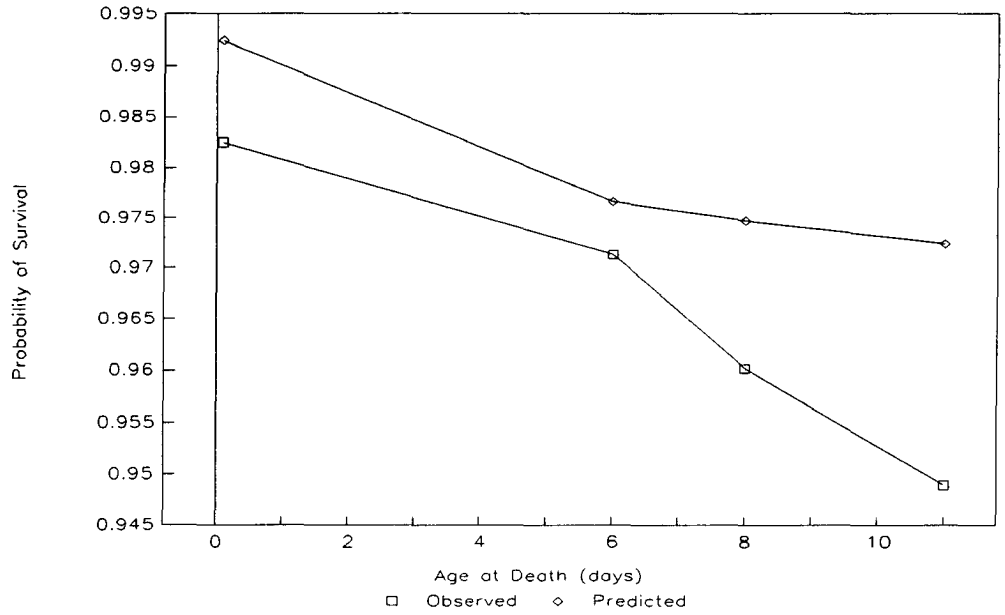


(a) Weibull model

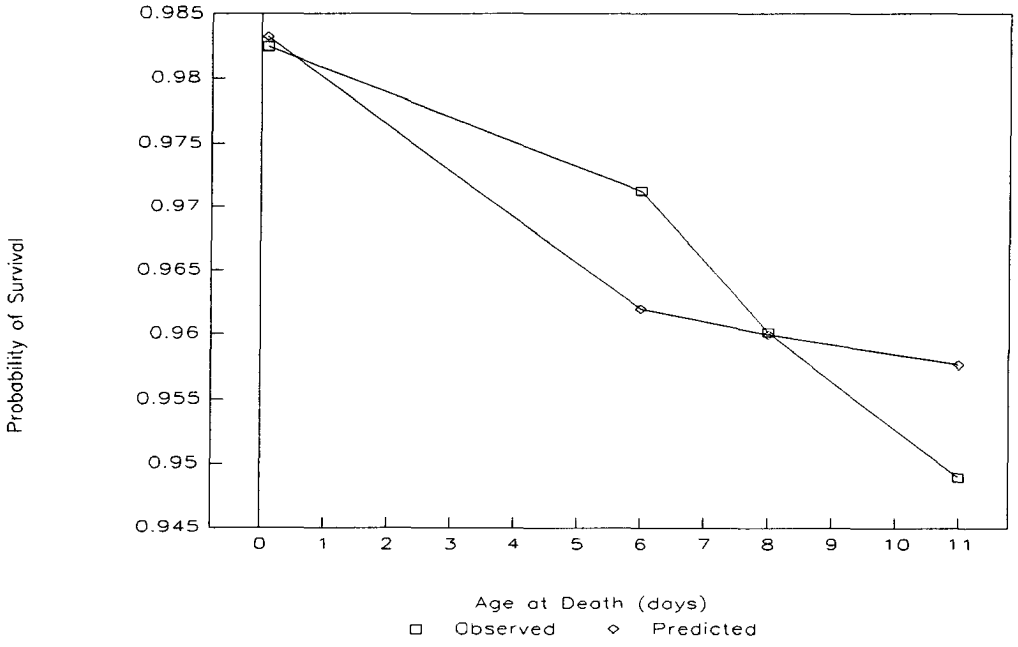


(b) Extreme Value model

Figure 34. Probability of survival with variable levels <1500 WEIGHT, no PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

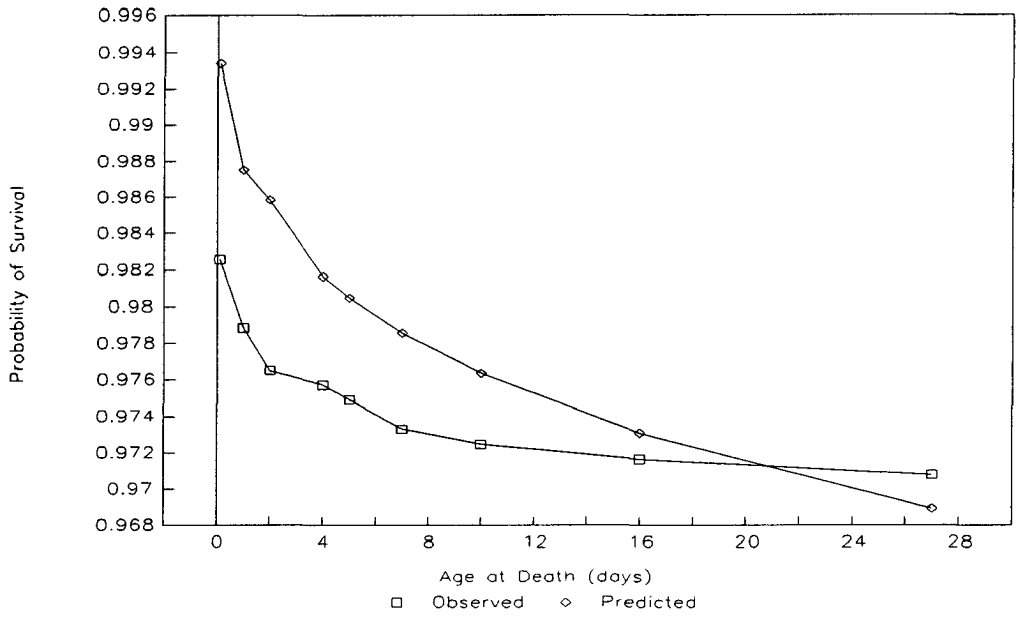


(a) Weibull model

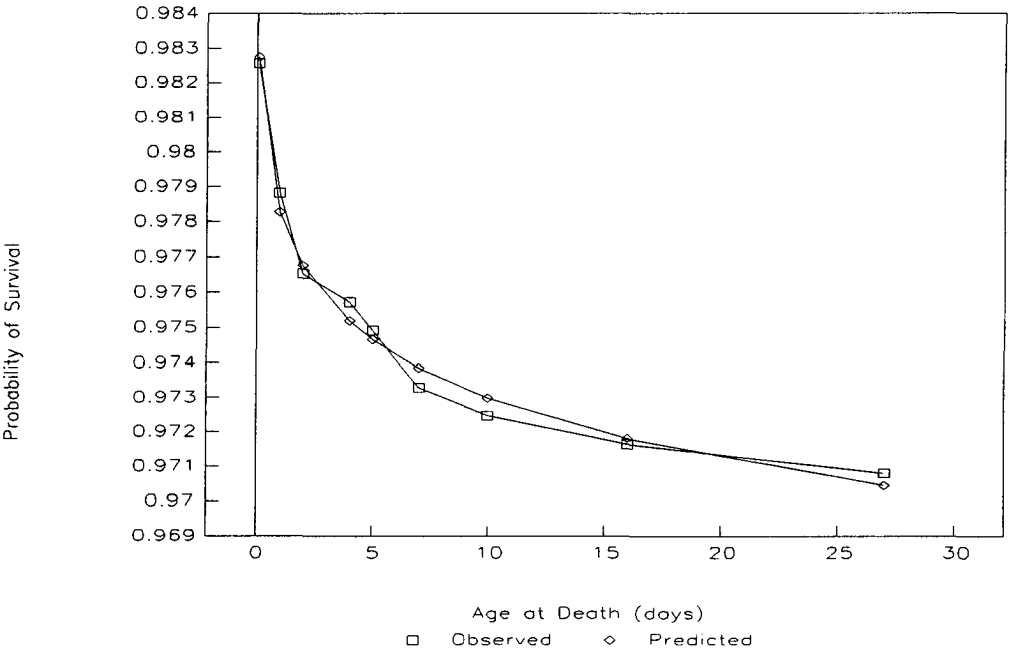


(b) Extreme Value model

Figure 35. Probability of survival with variable levels <1500 WEIGHT, no PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

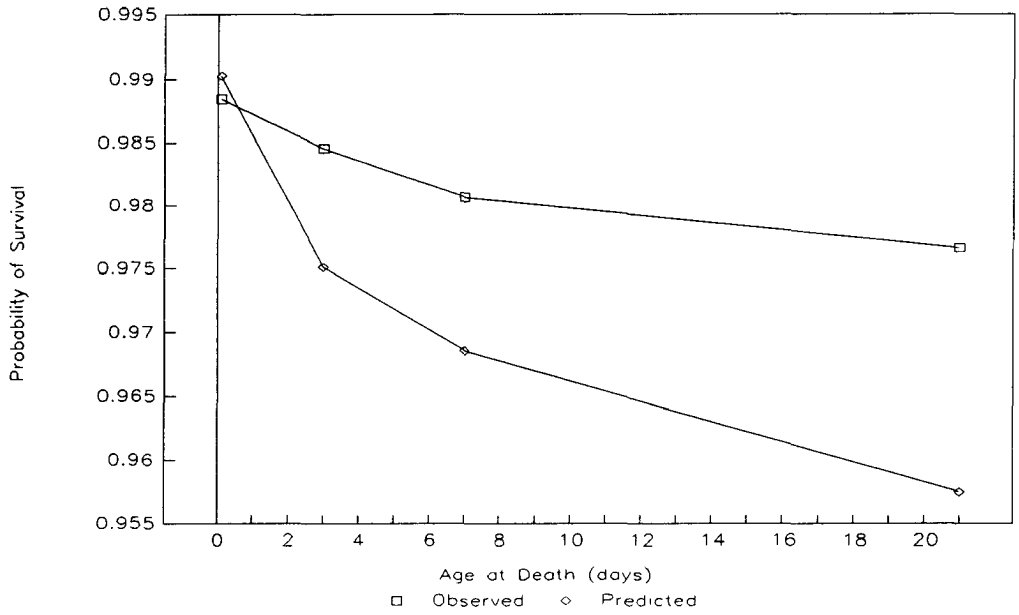


(a) Weibull model

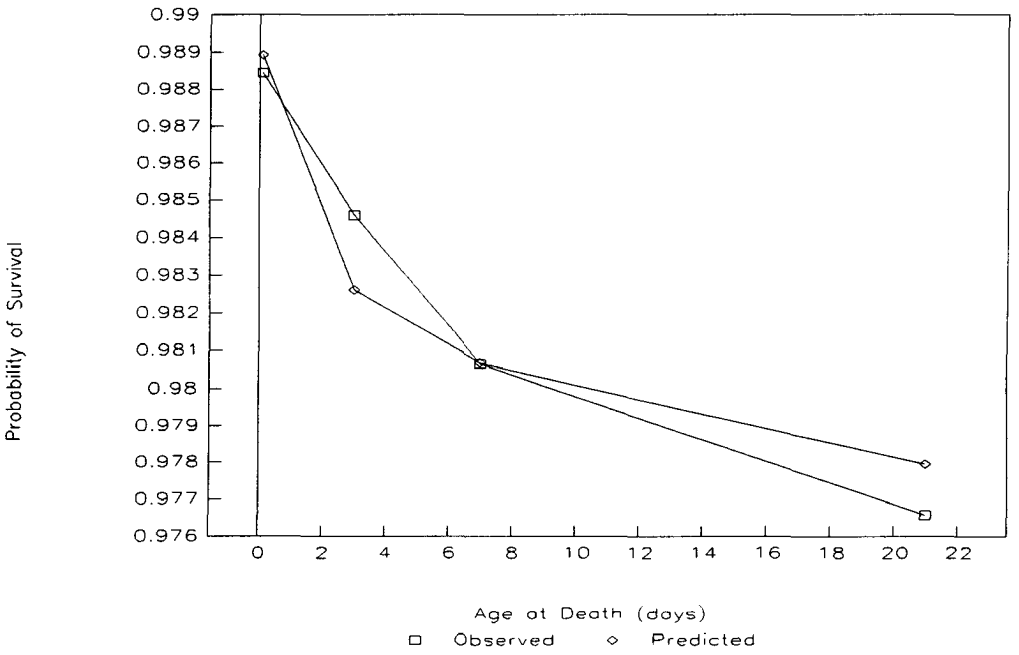


(b) Extreme Value model

Figure 36. Probability of survival with variable levels <1500 WEIGHT, 1+ PREG, and early CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.



(a) Weibull model



(b) Extreme Value model

Figure 37. Probability of survival with variable levels <1500 WEIGHT, 1+ PREG, and late CARE used as (a) covariates in Weibull model and (b) strata in Extreme Value model for congenital anomalies.

Bibliography

1. Elandt-Johnson, R. and Johnson, N. (1980). *Survival Models and Data Analysis*. John Wiley and Sons, New York.

2. World Health Organization (1977). *The International Classification of Diseases*. 9th rev., Geneva.

3. Cramer, J. (1987). *Social Factors and Infant Mortality: Identifying High Risk Groups and Proximate Causes*. Demography, 24, 299-322.

4. Eberstein, I., Nam, C., and Hummer, R. (1990). *Infant Mortality by Cause of Death: Main and Interaction Effects*. Demography, 27, 413-430.

5. Kalbfleisch, J., and Prentice, R. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, New York.

6. Cox, D., and Oaks, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London.

7. Agresti, A. (1990). *Categorical Data Analysis*. John Wiley and Sons, New York.

Appendix

Let

$$S_T(t) = \exp[-((t-\xi)/\theta)^c], \qquad (A.1)$$

be the survival distribution function (SDF) of a Weibull distribution for the random variable T, where $t>\xi$, $\theta>0$, and $c>0$. Then $Y = \ln(T-\xi)$ has a type 1 extreme value distribution.

Proof:

Let $\exp(y)+\xi = t$. (A.2)

Then substituting in (A.1) we have

$$S_T(\exp(y)+\xi) = \exp[-(1/\theta)^c \exp(cy)]$$

$$= \exp[-b \exp(ay)], \qquad (A.3)$$

where $b=(1/\theta)^c$ and $c=a$. Therefore, $S_T(\exp(y)+\xi)$ is of the form type 1 extreme value. The SDF of a type 1 extreme value distribution is

$$S_Y(y) = \exp[-\exp[(y-\psi)/\rho]], \qquad (A.4)$$

where $-\infty < y < \infty$, and $\rho > 0$. Equation (A.4) can be written as

$$S_Y(y)=\exp[-\exp(-\psi/\rho)\ \exp(y/\rho)]. \qquad (A.5)$$

By comparing equations (A.3) and (A.5) we can state that $c=1/\rho$ and $\theta=\exp(\psi)$. Which is of the form

$$\exp[-\theta^{-c}\ \exp(cy)].$$

VITA

Educational

University of North Florida - B.A. in Mathematical Sciences
1984

Professional

Research Associate - Department of Health and Rehabilitative Services, Office of Vital Statistics
1984 - present