ABSTRACT

PREOPERATIVE RISK ASSESSMENT FOR CARDIAC SURGERY USING PIECEWISE LINEAR REGRESSION MODEL

by
Arifa Zafar

In the present climate of quality assurance policies, rigorous requirements for informed consent, and a constantly changing patient population, a system of preoperative risk assignment for cardiovascular surgery was developed to monitor and evaluate surgical outcomes. The goal of this work is to estimate the preoperative risk associated with cardiac bypass surgery for patients in different risk categories. These risk categories are determined by the Parsonnet model which is based upon studying the severity of illness. The Parsonnet model assigns a risk value to a range of risk factors consisting of patient attributes and disease parameters. The aggregate of these risk factors is the mortality number in this thesis which is the subjective risk. After attaining posterior risk values for different risk classes we select a piecewise linear model which best estimates the risk for low, moderate and high risk cases. Confidence bands for posterior risk are also presented.

This thesis will utilize a database comprised of preoperative risk categories and their respective surgical outcomes in order to uniformly rate patient survival rates.

PREOPERATIVE RISK ASSESSMENT FOR CARDIAC SURGERY USING PIECEWISE LINEAR REGRESSION MODEL

by
Arifa Zafar

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirement for the Degree of
Master of Science in Biomedical Engineering

Biomedical Engineering Committee

January 1995

APPROVAL PAGE

PREOPERATIVE RISK ASSESSMENT FOR CARDIAC SURGERY USING PIECEWISE LINEAR REGRESSION MODEL

Arifa Zafar

Dr. Andrew Meyer, Thesis Advisor   Date
Professor of Electrical Engineering,
Electrical and Computer Engineering Dept., NJIT

Dr. Rose Dios, Committee Member   Date
Associate Professor of Mathemathics,
Mathemathics Dept., NJIT

Dr. David Kristol, Committee Member   Date
Professor of Chemistry,
Director of Biomedical Engineering Program, NJIT

BIOGRAPHICAL SKETCH

Author: Arifa Zafar

Degree: Master of Science in Biomedical Engineering

Date: January 1995

Undergraduate and Graduate Education:

- Master of Science in Biomedical Engineering,
  New Jersey Institute of Technology,
  Newark, New Jersey, 1995

- Bachelor of Engineering in Electrical Engineering,
  N.E.D University of Engineering and Technology,
  Karachi, Pakistan, 1988

Major: Biomedical Engineering

*This thesis is dedicated to
Dr. Rose Dios, Dr. Andrew Meyer and
my beloved father.*

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE SURVEY

1.1 Introduction

Risk assessment is becoming increasingly important in cardiac surgery. Patients desire an accurate statement of the risk they are to assume for the treatment of their disease. An informed consent is a medicolegal necessity. Quality-assurance programs are evolving and increasing the scrutiny applied to cardiac surgery and the need for accurate data acquisition on performance of cardiac operations. Changing methods of surgical management are probably altering risk and consequently making it important to observe how risks are defined and identified. Preoperative risk viewed as "risk factors" are dependent upon categories of severity of illness based on a wide range of patient attributes and disease parameters. These categories are identified by the well known physician Dr. V. Parsonnet [1], whose model serves as the source of the prior subjective probabilities of expiration for individual patients as shown in Table 1.1 . These prior probabilities are risk values subjectively assigned to patients for lacking (first number) or possessing (following number) a physical attribute which is recognized to be a risk factor. For example, if a patient is not diabetic, the subjective probability of expiration is 0 and if he is diabetic the subjective probability is 3%. In Table 1.1, the "coefficients of Risk" are unity for all risk factors prior to the statistical adjustment which produces the posterior risk probabilities. Essentially, these risk factors utilize the intuition of an experienced physician thus giving us rich medical information to achieve our goal.

It is our goal to estimate the preoperative risk associated with cardiac bypass

surgery for patients in different "overall risk" categories by transforming prior subjective

probabilities of expiration into posterior risk values utilizing a range of regression

procedures. Table 1.1 shows the prior subjective probabilities of expiration given by Dr.

Parsonnet. According to Parsonnet, we can view this model as the linear combination of

Table 1.1 Prior Subjective Probabilities of Expiration (Risk) [From Dr. V. Parsonnet]

i	Risk Factors (attribute)	Coefficient of Risk	Prior Subjective Probability (% of Risk): "Risk Value" x_i
1.	sexriskn (gender) (male,female)	1	(0,1)
2.	obesity (no,yes)	1	(0,3)
3.	diabetic (no,yes)	1	(0,3)
4.	hyperten (hypertension) (no,yes)	1	(0,3)
5.	efriskno (ejection fraction) (good,fair,poor)	1	(0,2,4)
6.	ageriskn (age) (0-69-,70-74,75-79,80+)	1	(0,7,12,20)
7.	reoperat(reoperation) (no,first,second,third)	1	(0,5,10,10)
8.	preopiab (intra aorta balloon) (no,yes)	1	(0,2)
9.	lva (no,yes)	1	(0,5)
10.	crashptc (no,yes)	1	(0,10)
11.	dialdepe (dialysis dependent) (no,yes)	1	(0,10)
12.	avr (no,gradient≥120,gradient<120)	1	(0,7,5)
13.	mvr (no,pressure≥160,pressure<60)	1	(0,8,5)
14.	tvr (no,yes)	1	(0,3)
15.	addedcab (no,yes)	1	(0,2)

variables (i.e., prior subjective probabilities which are additive) and coefficient of risk is

the slope of the line without using regression analysis. These coefficients of risks can be

viewed as a membership function of fuzzy set [3], where assignment of the grade of

membership is both subjective in nature, context-dependent and is a matter of definition

rather than measurement.

Then, $\hat{M} = b_0 + b_1 x_1 + \ldots + b_k x_k$

x_i = risk value of the ith risk factor

b_i = coefficient for i^{th} risk factor (i = 1, 2,........,15)

b_0 = intercept (minimum risk)

For example: Let x_3 denote the diabetic risk value, i.e.,

$x_3 = 0$ if patient is not diabetic

$x_3 = 3$ if patient is diabetic

So $0 \leq b_3 x_3 \leq 3b_3$. Hence, the patient's risk is increased by $3b_3$% if he (or she) is diabetic and is not increased at all if he (or she) is not diabetic. Thus, b_3 is the adjustment factor which transforms a prior risk value to a posterior risk value. This is repeated for all remaining risk factors. Here, we can see that the intuition of an experienced physician is helping us to achieve our goal.

1.2 Literature Survey

In this section, four papers are reviewed, which are relevant to subject of this thesis.

1.2.1. Preoperative Risk Assessment in Cardiac Surgery (Junod, Harlan, Payne, Smeloff, Miller, Kelly, Ross, Shankar, and McDermott [4])

A system of preoperative risk assignment and postoperative correlation was developed to monitor and evaluate surgical performance. Patients were categorized by operation priority (emergent, urgent, and elective).

Risk was assigned before operation using data from the Coronary Artery Surgery Study (CASS) and the recent literature. Patients were assigned by risk of operative death into one of the five categories (Table 1.2). The groups were compared using a X^2 test for significant difference.

Table 1.2 Division of Patients

No.	Categories	Criteria
1.	< 2%	Patients undergoing primary isolated coronary artery bypass grafting. (CABG)
2.	2 - 5 %	Patients had isolated valve replacement in the presence of good ventricular function, repoper CABG or primary isolated CABG.
3.	6 - 10 %	Patients had isolated valve replacement with poor ventricular function, valve replacement with CABG, and CABG in the presence of serious complicated features.
4.	11 - 50 %	Patients usually had multiple procedures or several complicated conditions.
5.	> 50 %	Patients in cardiogenic shock were certain to die without surgical intervention.

Results of Study by Junod et al:

- The observed mortality matches the predicted mortality in all risk groups except the 2 to 5 group where the observed risk was lower than the predicted risk.

- Operative deaths by surgical priority demonstrated a statistically higher risk in the emergent priority than the elective priority (Table 1.3).

- Overall, there was no increased risk associated with increased age (Table 1.4). The only

subset of patients with higher risk was women more than 70 years old. Overall, there was

no difference in risk between men and women.

Table 1.3 Operative Mortality by Surgical Priority for Patients Having Isolated Primary
CABG

Group	All Patients	Patients with Isolated Primary CABG
Elective	11/533 (2.1)*	2/329 (0.6)
Urgent	15/580 (2.6)	5/450 (1.1)
Emergent	26/190 (13.7)	7/134 (5.2)

* Numbers in parentheses are percents.

- Table 1.3 shows that patients were not given erroneously low risk. Therefore,

preoperative risk assignment is an effective method of quality assurance. Results of a

further study are given in Table 1.4.

Table 1.4. Operative Mortality by Age and Sex for Patients Having Isolated Primary
CABG

Age(yr.)	All Patients	Male	Female
<50	0/84	0/70	0/14
50-59	2/249(0.8)*	2/213(0.9)	0/36
60-69	7/348(2.0)NS**	6/250(2.4)NS	1/98(1.0)
≥70	5/232(2.2)NS	1/148(0.7)	4/84(4.8)p<0.05

*Numbers in parentheses are percents.
**NS = not significant to p < 0.05.

So the CASS researchers concluded that an age greater than 60 years and female

sex affected operative mortality. However, Junod and co-workers support the decreased

importance of age as a determinant in the seventh and eight decades. Only the class of

women showed a statistically significant difference in patients older than 79 years. Overall, there was no difference in risk between men and women.

This paper devoted a great deal of time to discussing high risk factors in surgical outcome risk. Frequently, high risk patients are all grouped together regardless of why they are considered to be high risk. High risk patients are usually compared to low risk ones. But this paper compared one high risk class to another one and so on. For this reason, it is interesting and useful for our future work.

1.2.2 Difference in Mortality Rates for Patients from Coronary Artery Bypass Graft (CABG) Surgery of Five Teaching Hospitals (Williams, Nush, Goldfurb [5])

Five teaching hospitals in Philadelphia cooperated in a project to compare information about patient outcomes . The purpose was to improve quality of care by identifying hospital-to-hospital differences in mortality rates.

The data extracted from the discharge abstract included the patient identification number, age, sex, race, discharge status (alive or dead), and whether the admission was scheduled in advance or emergency or urgent admissions and severity of illness. Beside that whether they are assigned to DRG 106 or DRG 107 (Table 1.5).

Table 1.5 Division of patients according to the type of surgery

DRG 106	Coronary artery catheterization and CABG surgery occur during the same admission.
DRG 107	Coronary artery catheterization is performed prior to admission, and only CABG surgery is performed during the admission.

Difference in DRG-specific, hospital-specific, and surgeon-specific mortality were examined using X^2 tests.

Results of Study by Williams et al:

- Mortality rates were higher in DRG 106 than in DRG 107. Patients in DRG 106 also were more severely ill.

- The hospital-to-hospital differences in mortality rates for DRG 107 were small and not statistically significant (p=0.572).

- Patients in hospitals A and D experienced lower mortality rates compared to the patients in the remaining three hospitals experienced higher mortality rates.

Although illness severity did identify patients who were more likely to expire, differences in severity of illness did not explain differences in hospital- or surgeon-specific mortality rates. Williams found inconclusive evidence for patient mortality rates associated with a surgeon's clinical skills, and, to a lesser extent, with the hospital's volume of procedures and the hospital's organization and staffing.

This encourages us to pursue the study of preoperative surgical risk for patients in different "overall risk" categories. A "prior probability of mortality" may be used to identify the primary risk groups. Hence, our work focuses upon the use of the "Parsonnet Model". (See chapter 2)

1.2.3 Multivariate Discriminant Analysis of Operative Mortality From the Collaborative Study in Coronary Artery Surgery (CASS) (Kennedy, Kaiser, Lloyd, Fisher, Maynard, Fritz, Myers, Mudd, Ryan, and Coggin[6])

The Collaborative Study in Coronary Artery Surgery (CASS) is a large multi-institutional study of the medical and surgical treatment of coronary artery disease (CAD). In an effort

to better understand the clinical and angiographic characteristics predictive of OM, Kennedy and associates have done a multivariate discirminant analysis of variables associated with OM.

The data file of CASS contains information about the clinical, angiographic, and surgical characteristics of patients enrolled in the study. The baseline data were controlled by physicians and trained data technicians at the time the patient was hospitalized for coronary arteriography.

The results of this multivariate discriminant analysis of the predictors of OM are presented for several clinical groups as shown below:

Table 1.6 Clinical Groups

Group I	All operated patients
Group II	All CABG operations
Group III	Elective CABG operations
Group IV	Urgent or emergent CABG operations
Group V	Patients in group II divided by age
Group VI	Patients in groups II, III, and IV divided by sex

The operative mortality for the total groups of patients and various subgroups is given in Table 1.7.

Table 1.7 Operative Mortality for Groups

Groups	No. of Pts.[*]	Description	OM (%)
I	6,652	All operated pts	2.9
II	6,176	All CABG pts	2.3
III	4,913	Elective CABG pts	1.7
IV	1,263	Urgent - emergent CABG pts	4.4
V	4,303	CABG only, < 60 years	1.4
	1,873	CABG only, ≥60 years	4.2
VI	5,197	Men CABG only	1.8
	979	Women CABG only	4.5

*No. of pts = Number of patients.

Results of Study by Kennedy et al:

- Clinical variables of most predictive value were age, female sex, increased heart size and congestive heart failure. Angiographic variables of importance included left ventricular wall motion abnormalities, and left main coronary disease.

There were six variables that contained the most predictive information by analysis for a group of 6,176 patients who had isolated bypass operations. They are age, left main coronary artery stenosis ≥90%, female sex, left ventricular end-diastolic pressure, and the presence of heart failure (CHF score and rales). The risk of OM for an individual patient may be estimated with the use of these clinical and angiographic characteristics.

1.2.4 To Assess the Association between Isolated Systolic Hypertension and Subclinical Cardiovascular Disease in the Elderly Aged 65 and above (Psaty, Furberg, Kuller, Borhani, Rautaharju, O'Leary, Bild, Robbins, Fried, and Reid[7])

Recent studies have drawn attention to the risk associated with SBP and ISH in particular.

The cardiovascular study is a prospective cohort study of risk factors for coronary

heart disease and stroke in men and women aged 65 years and above.

The participants of the program meet following criteria:

1) were 65 years or older

2) were non institutionalized

3) expected to remain in the area

4) gave informed consent

5) did have both DBP and SBP

6) do not have clinical cardiovascular disease and not taking hypertensive medicine

They used SPCC-PC for data analysis techniques included analysis of variance and

logistic regression. The 5-year age groups were divided and evaluation of high blood

pressure was done according to following basis:

1) Borderline Isolated Systolic hypertension: DBP of less than 90 mm Hg and SBP of

140-150 mm Hg.

2) ISH: DBP of less than 90 mm Hg and SBP of at least 160 mm Hg.

3) DHhypertension: DBP of at least 90 mm Hg.

Results of the Study by Psaty et al:

- The subsequent analysis focused on the 1322 women and the 867 men who were not

receiving antihypertensive medications, were free of clinical cardiovascular disease, and

had a DBP of less than 90 mm Hg.

- The prevalence increased with age in both men and women (P < .0001). The proportion with borderline ISH was slightly higher in women (22.7%) than in men (20.9%), as was the prevalence of ISH (8.7% for women, and 8.5% for men).

- Most of the differences among groups were significant. Both fasting blood glucose and cholesterol levels, for instance, were higher among those with ISH.

- Systolic hypertension was not only associated with unrecognized myocardial infarction and increased LVM, but also with cardiac function and carotid IMT. Based on clinical trials, the benefits of treating both isolated systolic and combined hypertension in the elderly are clear.

CHAPTER 2

MODELING DEVELOPMENT FOR RISK ESTIMATION

2.1 The Logistic Model

Frequently, one wishes to pose a model which possesses some specific asymptotic trends. In particular, when we wish to "smooth" out a step function, we recognize the need to incorporate the following conditions:

$$\lim_{x \to -\infty} f(x) = 0 \qquad \text{and} \qquad \lim_{x \to -\infty} f(x) = B \qquad (2.1)$$

where B is a constant, usually equal to 1. If we further require that f(x) be monotone increasing, then we establish two key features in the model: (1) f(x) satisfies the conditions of a distribution function (e.g., a probability distribution for $0 < B \leq 1$). (2) f(x) possesses attributes of a function, y, which satisfies the following initial value problem.

$$\frac{dy}{dx} = Ay(B-y) \qquad \text{with } y(0) = B_0 \qquad (2.2)$$

We want a function whose rate of change is (a) proportional to the dependent variable value, and (b) proportional to a constant minus the dependent value. That is

$$y' \propto y \qquad \text{and} \qquad y' \propto (B - y) \qquad (2.3)$$

Historically, $y' \propto y$ leads us to a well known model in population dynamics [The Malthusian Linear Model]. Demographically, we expect however that the population growth will level off as the maximum available space and resources are depleted. Hence $y' \propto (B - y)$ represents a leveling off of the growth function which leads us to the

12

previously outlined initial value problem referred to as the logistic model. This approach allows us to smooth a step function to a differentiable probability distribution which will model the cumulative risk and utilize the database variable.

2.2 Piecewise Linear Model

We wish to establish a piecewise model that estimates the predicted risk of cardiovascular surgery (as a percent likelihood of fatality due to surgical procedure).The purpose is to improve quality of care by giving patients a better understanding of the results of the operative treatment. By using the predicted operative mortality as a guide, quality-assurance requirement and performance can be assigned for improvement. The initial model will utilize the available data for 1021 patients. The data extracted from the patients' discharge abstract includes the patients' age, sex, etc. (as shown Table 1.1). These are called risk factors. The discharge abstract also contains information describing the outcome of the surgery . Discharge status (death rate) is 0 if the patient is alive 30 days after operation or, discharge status is 1 for expiration.

1. The discharge abstract also contains information which reflects the patient's overall health condition in the form of a subjective risk value (probabilities of fatality due to a given medical condition) as well as the aggregate of the risk values which constitute a prior probability of expiration given subjectively by an expert physician. These values M, are referred to as the *"mortality number"* in the patient database.

2. We initially divide the patients into different overall risk categories by observing the value of the mortality number. The discharge abstract also contained information about

the each patient illness-severity level.

There are 1021 patients who have had cardiovascular surgery within the last 5 years.

148 Patients with 0 < Mortality number <4: *"Low-risk"*

224 Patients with 5 <= Mortality number <9: *"Moderate-risk"*

220 Patients with 10 <= Mortality number <14: *"Intermediate-risk"*

272 Patients with 15 <= Mortality number < 25: *"High-risk"*

157 Patients with Mortality number => 25: *"Critical-risk"*

Thus, the higher stage numbers indicate greater severity of illness.

The data is organized so that each of the input variables identifies key risk contributors and quantifies their values in the form of percentages. The risk contributors are modeled to be additive and "mutually independent" in the calculation of overall risk.

2.3. Multivariate Linear Regression Model

Hence, we select a linear regression model to estimate overall risk as shown. Let,

$$\hat{M} = b_0 + b_1 x_1 + \ldots\ldots\ldots + b_k x_k \quad \{k=15\}$$

= estimated risk

x_j = value assigned by (physician to patient) of risk category $\{j = 0,\ldots,15\}$

To obtain weight of risk-category for coefficient $b_0 ,\ldots\ldots, b_{15}$, linear regression is used, based on N=1021 previous patients. A linear regression (least squares) model allows us to access the model validity (fit) and accuracy. This is a form of linearly combining the contributions to surgical risk in order to obtain an aggregate risk value. In this model, the goodness of fit is determined by the F statistic.

The ANOVA table associated with this multivariate regression process is shown in following table [8]:

Table 2.1 Analysis of Variance (ANOVA)

Source	SS	Degree of Freedom	Mean Square	F_{ratio}
Regression	SS_{reg}*	k	$MS_{reg} = \dfrac{SS_{reg}}{k}$	
Residual	SS_{res}**	(n-k-1)	$MS_{res} = \dfrac{SS_{res}}{(n-k-1)}$	$F = \dfrac{MS_{reg}}{MS_{res}}$
Totals corrected	$SS_{total\ corrected}$	(n-1)		

*reg=regression

**res=residual

The elements of this table are defined below in eqs. (2.4) - (2.10):

The correlation coefficient is R as shown

$$R = \sqrt{\frac{SS_{reg}}{SS_{total\ corrected}}} = \text{correlation coefficient} \qquad (2.4)$$

where

$$SS_{regrssion} = \sum (\hat{M} - \overline{M})^2 \qquad (2.5)$$

$$SS_{residual} = \sum (M - \hat{M})^2 \qquad (2.6)$$

$$SS_{total\ corrected} = \sum (M - \overline{M})^2 \qquad (2.7)$$

We utilize the fundamental partition equation

$$SS_{\text{total corrected}} = S_{\text{reg}} + SS_{\text{res}}$$ (2.8)

We define the variables and notations as follows:

$N = 1021$ = number of patients in the data base (sample-size)

$k = 15$ = total number of risk categories (degree of freedom for regression model)

i = patient number ($i = 1,\ldots\ldots, N = 1021$)

j = category number ($j = 1,\ldots\ldots\ldots, j = 15$)

x_{ji} = risk value of category j for patient i

b_j = weight for risk category j in the regression model (same for all patients)

b_o = minimum risk for all patients from regression model

$$M_i = \sum_{j=1}^{k} x_{ji}$$ (2.9)

= subjective probability number for patient i

$$\overline{M} = \frac{1}{N} \sum_{i=1}^{N} M_i = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} x_{ji}$$ (2.10)

= expected mortality number for entire sample of $N = 1021$ patients

$$\hat{M}_i = b_o + \sum_{j=1}^{k} b_j x_{ji}$$ (2.11)

= Regression model adjusted risk value for patient i

2.4 Bivariate Linear Regression Model

Let us now look at the death rate for surgery and compare it to the mortality number. The

death rate is proportional to $\hat{M}$, i.e.,

$$\text{Death rate} \approx \sum_{m=0}^{k} b_m \, x_m \quad \text{where } x_o = 1 \text{ and } k = 15$$

Mortality number and death rate are quantitative measures of overall risk:

the former is a subjective probability assigned by the physician or surgeon; the latter is an

aggregated discrete code for the outcomes of the surgical procedure. It is the percentage

of patient with a given mortality number who did not survive the surgery. The regression

model mortality number is the weighted sum of risk contributions. The pre-established

weights are the physician's subjective contribution to this measure.

Ideally, uncorrected death rate is exactly equal to the sum of mortality numbers for

a given class, which does not encounter minimum risk as shown in Figure 2.1 [1].
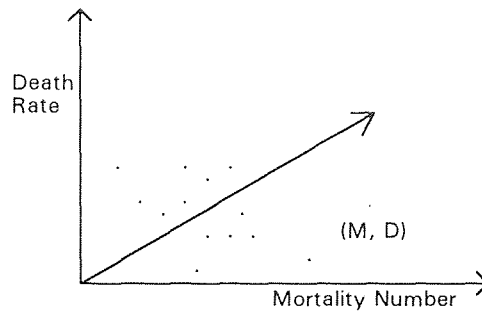
Figure 2.1 Uncorrected Death Rate as Compared to Mortality Number

Therefore, we would like to identify the relation between the Mortality number (as a percentage of risk) and the surgical outcome (survival or expiration). We fit the Mortality number with this outcome (referred to as "Corrected Death Rate") in a linear model to determine the regression coefficients as shown in Figure 2.2.
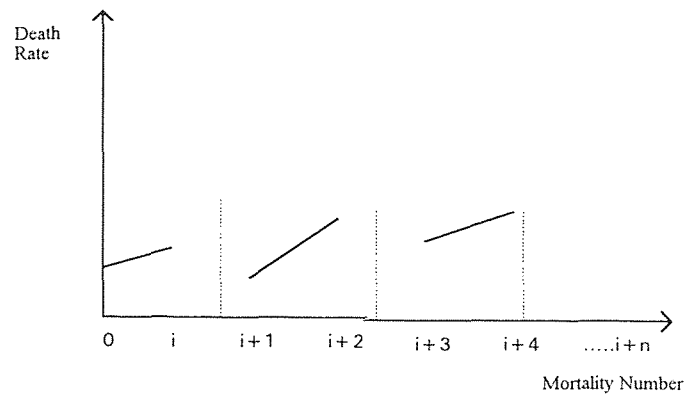
Figure 2.2 Piecewise Linear Model

$\hat{D}$ = Corrected death rate

$\quad = a_0 + a_1 M$

a_0 = intercept (minimum risk of operation, i.e., assuming 0 for everything)

a_1 = posterior coefficient of death rate (calculated from regression analysis)

M = mortality number

CHAPTER 3

STATISTICAL ANALYSIS AND RESULTS

3.1 Methods

Information on the statistical program is presented below:

Patient Population: The database being utilized contains 1021 patients who have had

cardiovascular surgery within the last 5 years.

Database: Extensive historical, physical and lab information was collected on each patient.

Each of these parameters was analyzed univariately to determine the relationship of the

variable to mortality number.

Statistical Analysis: Regression analysis is used to develop a rule to distinguish between

two or more groups, in this case between those who survive and those who do not survive

cardiovascular surgery. The first step in this process is the examination of variables that

measure characteristic which are expected to differ in the two groups (survivors and non-

survivors). The fifteen variables identified as related to mortality number and subsequently

used in regression analysis.

Goal: Our goal is to determine the posterior probability of expiration due to

cardiovascular surgery, based upon a prior "cardiac risk distribution" provided by the

"Parsonnet Model". These initial risk values will be adjusted so as to estimate an overall

risk function (at first by linear approximation). This adjustment is twofold. We seek a

linear combination of "independent" risk factors as an average computation of risk.

Table 3.1 Reduced Predictor/Maximum Information Model (RMM)

Risk Factors	Initial Weight	Final Weight	Initial Risk Contribution (Maximum Parsonett Number) (Prior Prob.)	Final Contribution (Posterior Prob.)
intercep	1	0.896634	0	0.896634
sexriskn (male, female)	1	1.196721	(0,1)	(0,1.196721)
obesity (no, yes)	1	0.876160	(0,3)	(0,2.62848)
diabetic (no, yes)	1	1.203874	(0,3)	(0,3.611622)
hyperten (no, yes)	1	0.968884	(0,3)	(0,2.906652)
efriskno (good, fair, poor)	1	1.451194	(0,2,4)	(0,2.902388,5.804776)
ageriskn (0-69, 70-74, 75-79, 80+)	1	1.045009	(0,7,12,20)	(0,7.315063,12.540108, 20.9001)
reoperat (no, first, second, third)	1	1.086016	(0,5,10,10)	(0,5.43008,10.8601, 10.86016)
preopiab (no, yes)	1	2.846297	(0,2)	(0,5.692594)
Iva (no, yes)	1	1.431998	(0,5)	(0,5.187995)
crashptc (no, yes)	1	1.037599	(0,10)	(0,10.37599)
dialdepe (no, yes)	1	1.052031	(0,10)	(0,10.52031)
avr (no, gradient ≥ 120, gradient < 120)	1	1.107784	(0,7,5)	(0,7.754488,5.53892)
mvr (no, pressure ≥ 60, pressure < 60)	1	1.498085	(0,8,5)	(0,11.9868,7.490425)
tvr (no, yes)	1	2.947985	(0,3)	(0,8.843955)
addedcab (no, yes)	1	0.777405	(0,1)	(0,0.777405)

Further, since we are utilizing regression analysis, we are minimizing the sum of squares of the errors.

The previous study used a piecewise linear model subjectively selected by the physicians involved in this study. At this point we look at alternative piecewise linear model in order to find the one which produce the highest level of statistical confidence. We are using as a model RMM (Reduced Predictor/ Maximum Information) (Table 3.1) model which gives us the maximum information according to consulting physicians. We have divided the model into four different groups. Below we construct the previous linear models for different groups and different categories. In group 1, the risk categories are of length 2 units. In group 2, the risk categories are of length 3 units. In group i, the risk categories are of length i+1 units. The critical risk category always begins with a mortality number approximately 20-22 % and ends at 100 % for all groups. The piecewise linear regression model are then presented graphically.

Group 2:

There are 1021 patients with the following value of mortality number.

RCLASS 1.00   67 patients with M = 0, 1

RCLASS 2.00   81 patients with M = 2, 3

RCLASS 3.00   80 patients with M = 4, 5

RCLASS 4.00   79 patients with M = 6, 7

RCLASS 5.00   87 patients with M = 8, 9

RCLASS 6.00   84 patients with M = 10, 11

RCLASS 7.00   88 patients with M = 12, 13

RCLASS 8.00   83 patients with M = 14, 15

RCLASS 9.00   69 patients with M = 16, 17

RCLASS 10.00   59 patients with M = 18, 19

RCLASS 11.00   53 patients with M = 20, 21

RCLASS 12.00   191 patients with 22>M

                     1021 patients

Group 3:

There are 1021 patients with the following value of mortality number.

RCLASS 1.00   108 patients with M = 0, 1, 2

RCLASS 2.00   127 patients with M = 3, 4, 5

RCLASS 3.00   127 patients with M = 6, 7, 8

RCLASS 4.00   123 patients with M = 9, 10, 11

RCLASS 5.00   136 patients with M = 12, 13, 14

RCLASS 6.00   104 patients with M = 15, 16, 17

RCLASS 7.00   83 patients with M = 18, 19, 20

RCLASS 8.00   213 patients with 21>M

                     1021 patients

Group 4:

There are 1021 patients with the following value of mortality number.

RCLASS 1.00   150 patients with M = 0, 1, 2, 3

RCLASS 2.00   159 patients with M = 4, 5, 6, 7

RCLASS 3.00   171 patients with M = 8, 9, 10, 11

RCLASS 4.00   171 patients with M = 12, 13, 14, 15

RCLASS 5.00   128 patients with M = 16, 17, 18, 19

RCLASS 6.00   242 patients with 20>M

                   1021 patients

Group 5:

There are 1021 patients with the following value of mortality number.

RCLASS 1.00   172 patients with M = 0, 1, 2. 3, 4

RCLASS 2.00   224 patients with M = 5, 6, 7, 8,9

RCLASS 3.00   220 patients with M = 10, 11, 12, 13, 14

RCLASS 4.00   163 patients with M = 15, 16, 17, 18, 19

RCLASS 5.00   242 patients with 20 >M

                   1021 patients

3.2 Results

3.2.1 Evaluation of Piecewise Linear Model

By plotting the results (Appendix B) we observe:
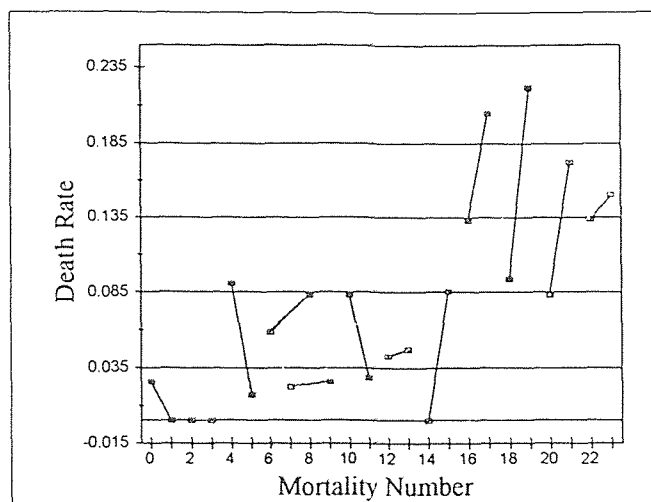
Figure 3.1 Plot of Death Rate and Mortality Number for Group 2

Group 2 (Figure 3.1):

i) Low risk	Not good
ii) Moderate risk	Not good
iii) Intermediate risk	Not good
iv) High risk	Good
v) Critical risk	Good
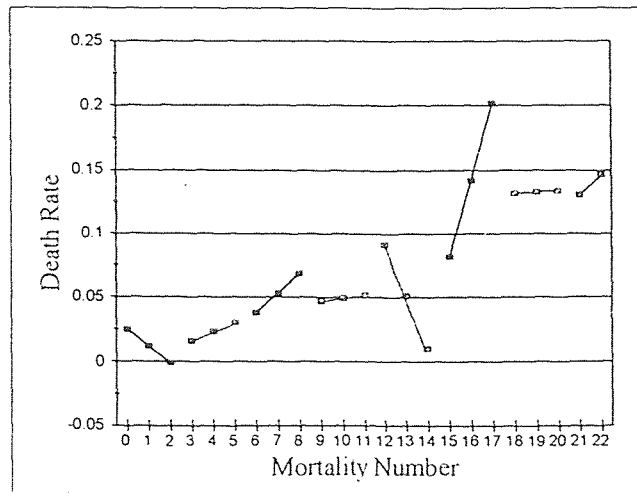
Figure 3.2 Plot of Death Rate and Mortality Number for Group 3

Group 3 (Figure 3.2):

i) Low risk	Good
ii) Moderate risk	Good
iii) Intermediate risk	Not good
iv) High risk	Not good
v) Critical risk	Good

Figure 3.3 Plot of Death Rate and Mortality Number for Group 4

Group 4 (Figure 3.3):

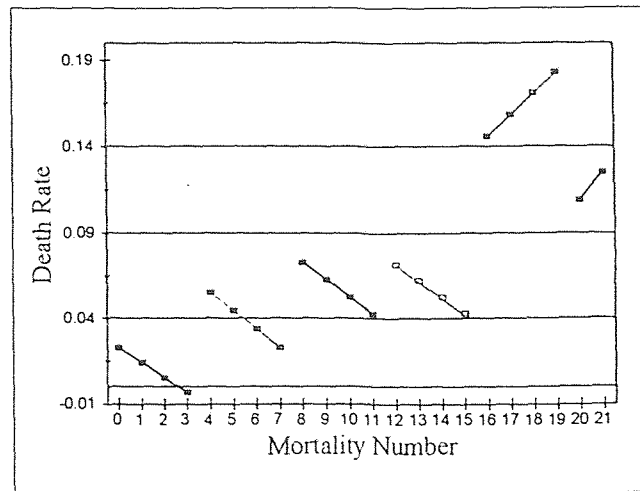|	
i) Low risk	Not good
ii) Moderate risk	Not good
iii) Intermediate risk	Not good
iv) High risk	Not good
v) Critical risk	Good
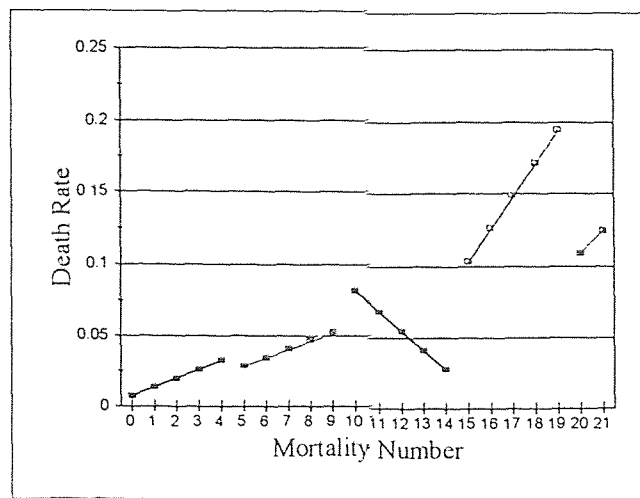
Figure 3.4 Plot of Death Rate and Mortality Number for Group 5

Group 5 (Figure 3.4):

i) Low risk	Good
ii) Moderate risk	Good
iii) Intermediate risk	Not Good
iv) High risk	Good
v) Critical risk	Good

We evaluate the piecewise linear models heuristically knowing the preoperative risk should be monotone increasing function with positive slopes and continuity at interval endpoints. Hence, we may select Group 5 as most plausible except for the intermediate risk category. The intermediate risk category should be studied separately within the context of logistic model.

3.2.2 Equations for Confidence Bands

We observe that the graph of the confidence bands (Appendix B) is governed by the equation / inequality.

$$y \leq y_0 \pm t \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2}} \cdot s \qquad (3.1)$$

where s is the standard error from the ANOVA table and t is the 95% t-score from the student's t distribution and we may rewrite this equation as

$$(y - y_0)^2 \leq t^2 s^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2} \right] \qquad (3.2)$$

which can be simplified as follows.

$$(y - y_0)^2 - \frac{(x_0 - \bar{x})^2 t^2 s^2}{\Sigma(x - \bar{x})^2} \leq t^2 s^2 \left[1 + \frac{1}{n} \right] \qquad (3.3)$$

Hence,

$$\frac{(y - y_0)^2}{a^2} - \frac{(x_0 - \bar{x})^2}{b^2} \leq 1 \qquad (3.4)$$

where

$$a^2 = t^2 s^2 (1 + \frac{1}{n}) \qquad \text{and} \qquad (3.5)$$

$$b^2 = (1 + \frac{1}{n}) \Sigma(x_0 - \bar{x})^2 \qquad (3.6)$$

This is the equation of a hyperbola (which opens upward & downward). This equation locates the confidence bands on the risk function graphs.

The values of a^2 & b^2 for the preferred risk function are given below :

$$a^2 = (1.96)^2 (0.19705)^2 [1+\frac{1}{223}] = 0.15$$

$$b^2 = (463.25) (1+\frac{1}{223}) = 465.33$$

for

$\alpha = .05$ indicating 95% confidence

for $5 \leq$ mortality number ≤ 9

for Group 5.

CHAPTER 4

SUMMARY & CONCLUSIONS

The piecewise linear models which we have considered in this thesis allow us to establish a patient's risk for different overall risk categories. The fourth solution appears to be the optimal one with separation of risk into the following intervals or risk categories :

0	$\leq$	Overall Risk	$\leq$	4	Low risk
5	$\leq$	Overall Risk	$\leq$	9	Moderate risk
10	$\leq$	Overall Risk	$\leq$	14	Intermediate risk
15	$\leq$	Overall Risk	$\leq$	19	High risk
20	$>$	Overall Risk			Critical Risk

This estimation process gives best result for moderate risk.

We further explored the confidence which enclose the given risk function. These were found to be very wide for 95% confidence. The width of the band is

$$W_\alpha \leq 2.t_\alpha \cdot \sqrt{1+\frac{1}{n}+\frac{(x_o - \overline{x})^2}{\Sigma(x-\overline{x})^2}} \quad .s \qquad (4.1)$$

Since the radical is very close to 1 for large n and variance of $\overline{X}$, we may approximate the width to be

$$W_\alpha \approx 2.t_\alpha .1.s \qquad (4.2)$$

which was found to be quite large for our data. If we reduce confidence from 95% to 70% , t_α is reduced from 2 to 1.

30

But what must be small for tight confidence bands is s = Root Mean Square

Residual. If a piecewise model has b > 0 and s ≈ 0.05 we obtain confidence bands which

enclose the risk function with width ≈10% or 20% for $\alpha \approx 0.3$ or $\alpha \approx 0.05$ respectively.

This thesis has presented the twofold process for determining risk: the sequential

calculation of estimated risk as presented by Teng [2] followed by the calculation of death

rate. Both elements of this sequential calculation could be subjected to logistic smoothing

with smaller confidence bands.

The smoothing of the piecewise models into logistic functions may produce such

an appropriate value for s and we recommend that to be the next step in the on-going

research of this problem.

APPENDIX A

Components of the Additive Model

Risk Factor	Weight	Disasters and Rare Conditions	Weight
Age at operation		0. None	0
0-69	0		
70-74	7	CARDIAC CONDITIONS	
75-79	12		
80+	20	1. Left Main Disease, Unstable Angina	3
Sex		2. Ventricular Tachycardia / Ventricular Fibrillation (VT/VF), aborted sudden death	5
Male	0	3. Shock/Cardiogenic (urinary output < 10 cc/hr, mean BP 40 without vasopressors)	25
Female	1	4. Transmural Acute MI within 48 hrs	7
Ejection Fraction		5. CHF, chronic (with peripheral edema, plural effusion)	5
Good or 50%+	0	6. Pacemaker Dependent	2
Fair or 30%-49%	2	7. AR, acute (endocarditis)	10
Poor or 1%-29%	4	8. MR, acute (endocarditis, papillary muscle rupture, etc.)	10
Morbid Obesity		9. VSD, Acute	20
No	0	10. Constrictive Pericarditis	5
Yes	3	11. Congenital Heart Disease in adult, cyanotic	10
Diabetes			
No	0	HEPATO-RENAL CONDITIONS	
Yes	3		
Hypertension		12. Renal Failure, Chronic (CR>2, w/out dialysis)	5
No	0	13. Renal Failure, Acute	25
Yes	3	14. Cirrhosis of liver, (serum bilirubin > 5)	
Reoperation			
No	0	PULMONARY CONDITIONS	
First	5		
Second	10	15. COPD, sever	5
Third	10	16. Pulmonary Hypertension (mean pressure > 30)	10
Preoperative IABP		17. Idiopathic Thrombocytopenci Purpura (ITP)	10
No	0		
		18. Endotrachial Tube, pre-operation	5
Yes	2	19. Asthma (peak expiratory flow rate < 100)	20
LV Aneursym		20. Asthma (peak expiratory flow rate < 200)	10
No	0		
Yes	5		

APPENDIX A
(Continued)

Risk Factor	Weight	Disasters and Rare Conditions	Weight
Dialysis-dependent			
No	0	21. PVD, severe	2
Yes	10	22. Carotid Disease, unilateral occlusion	5
PTCA or			
Catherization "crash"		23. Carotid Disease, bilateral	1
No	0	24. AAA, Asymptomatic	1
Yes	10	25. Dissecting Thoracic Aneurysm	
MV procedure			
No	0	MISCELLANEOUS CONDITIONS	
Yes			
PA pressure >=60	8	26. Severe neurologic disorder (healed CVA,	
PA pressure < 60	5	paraplegia, muscular dystrophy,	
		hemoparesis)	
		27. Diabetes, Juvenile	
AV procedure		28. Hyperlipidemia (cholesterol > 300,	
No	0	HDL < 30)	
		29. Jehovah's Witness	
Yes		30. Cold Agglutinins	
Gradient>=120	7	31. Aspirin Rx (ASA Rx)	
Gradient < 120	5	32. Substance abuse (alcohol, drugs), severe	
TV Procedure		33. AIDS, active disease (HIV positive	
No	0	excluded)	
		34. Active Neoplasm (leukemia, lymphoma,	
Yes	3	etc.)	
		35. High-dose steroids, active	
Added CABG			
No	0		
Yes	2		

APPENDIX B

Calculations for Corrected Death Rate

Table B.1 Calculations for Corrected Death Rate for Group 2

	RCLASS 1:00		RCLASS : 2:00	
X - Axis	0	1	2	3
b_0	0.025641	0.025641	0.00	0.00
b_1	-0.02564	-0.02564	0.00	0.00
Death Rate	0.025641	0.00	0.00	0.00
	RCLASS 3:00		RCLASS : 4:00	
X - Axis	4	5	6	7
b_0	0.385580	0.385580	0.630952	0.630952
b_1	-0.073668	-0.073668	-0.054762	-0.05476
Death Rate	0.0909	0.05407	0.083332	0.02857
	RCLASS : 5:00		RCLASS : 6:00	
X - Axis	8	9	10	11
b_0	0.544872	0.544872	0.6122245	0.385580
b_1	-0.057692	-0.057692	-0.05061	-0.073668
Death Rate	0.083336	0.025644	0.025644	0.028574
	RCLASS : 7:00		RCLASS : 8:00	
X - Axis	12	13	14	15
b_0	0.020894	0.020894	0.00	0.00
b_1	0.011429	0.011429	0.00	0.00
Death Rate	0.020894	0.03232	0.00	0.00
	RCLASS : 9:00		RCLASS : 10:00	
X - Axis	16	17	18	19
b_0	0.385580	0.385580	0.278431	0.27843
b_1	-0.073668	-0.073668	-0.036601	-0.03660
Death Rate	0.0909	0.05407	0.058825	0.02222
	RCLASS : 11:00		RCLASS : 12:00	
X - Axis	20	21	22	23
b_0	0.544872	0.544872	0.227635	0.227635
b_1	-0.057692	-0.057692	0.016441	0.016441
Death Rate	0.083336	0.025644	0.134067	0.150508

Table B.2 Calculations for Corrected Death Rate for Group 3

	RCLASS	1:00	
X - Axis	0	1	2
b_o	0.024730	0.024730	0.024730
b_1	-0.01288	-0.01288	-0.01288
Death Rate	0.024730	0.011843	0-1.044x10
	RCLASS	2:00	
X - Axis	3	4	5
b_o	-0.005449	-0.005449	-0.00544
b_1	0.007114	0.007114	0.007114
Death Rate	0.015893	0.023007	0.030121
	RCLASS	3:00	
X - Axis	6	7	8
b_o	-0.053435	-0.053435	-0.05343
b_1	0.015267	0.015267	0.015267
Death Rate	0.038167	0.053434	0.068701
	RCLASS	4:00	
X - Axis	9	10	11
b_o	0.022636	0.022636	0.022636
b_1	0.002663	0.002663	0.002663
Death Rate	0.046603	0.049266	0.051929
	RCLASS	5:00	
X - Axis	12	13	14
b_o	0.580396	0.580396	0.580396
b_1	-0.040710	-0.040710	-0.040710
Death Rate	0.091876	0.051166	0.010456
	RCLASS	6:00	
X - Axis	15	16	17
b_o	-0.82034	-0.82034	-0.82034
b_1	0.059896	0.059896	0.059896
Death Rate	0.082034	0.14193	0.201826
	RCLASS	7:00	
X - Axis	18	19	20
b_o	0.111911	0.111911	0.111911
b_1	0.001091	0.001091	0.001091
Death Rate	0.131549	0.13264	0.133731
	RCLASS	8:00	
X - Axis	21	22	
b_o	-0.198985	-0.198985	
b_1	0.015695	0.015695	
Death Rate	0.13061	0.146305	

Table **B.3** Calculations for Corrected Death Rate for Group 4

	RCLASS	1:00		
X - Axis	0	1	2	3
b_o	0.002263	0.002263	0.002263	0.002263
b_1	-0.00861	-0.00861	-0.00861	-0.00861
Death Rate	0.002587	0.006345	0.010103	0.013861
	RCLASS	2:00		
X - Axis	4	5	6	7
b_o	0.098894	0.098894	0.098894	0.098894
b_1	-0.01084	-0.010841	-0.01084	-0.010841
Death Rate	0.05553	0.044689	0.033848	0.023007
	RCLASS	3:00		
X - Axis	8	9	10	11
b_o	0.01552	0.01552	0.01552	0.01552
b_1	-0.01022	-0.01022	-0.01022	-0.010221
Death Rate	0.07325	0.063037	0.052816	0.042595
	RCLASS	4:00		
X - Axis	12	13	14	15
b_o	0.188894	0.188894	0.188894	0.188894
b_1	-0.009730	-0.009730	-0.009730	-0.009730
Death Rate	0.072134	0.062404	0.052674	0.042944
	RCLASS	5:00		
X - Axis	16	17	18	19
b_o	-0.053157	-0.053157	-0.053157	-0.053157
b_1	0.012457	0.012457	0.012457	0.012457
Death Rate	0.146155	0.158612	0.171069	0.183526
	RCLASS	6:00		
X - Axis	20	21		
b_o	-0.21119	-0.21119		
b_1	0.016021	0.016021		
Death Rate	0.109229	0.12525		

Table B.4 Calculations for Corrected Death Rate for Group 5

|	RCLASS	1:00			
X - Axis	0	1	2	3	4
b_0	0.007448	0.007448	0.007448	0.007448	0.007448
b_1	0.006263	0.006263	0.006263	0.006263	0.006263
Death Rate	0.007448	0.013711	0.019974	0.026237	0.0325
	RCLASS	2:00			
X - Axis	5	6	7	8	9
b_0	-0.00353	-0.00353	-0.00353	-0.00353	-0.00353
b_1	0.006342	0.006342	0.006342	0.006342	0.006342
Death Rate	0.028176	0.034518	0.04086	0.047202	0.053544
	RCLASS	3:00			
X - Axis	10	11	12	13	14
b_0	0.218016	0.218016	0.218016	0.218016	0.218016
b_1	-0.01358	-0.01358	-0.01358	-0.01358	-0.01358
Death Rate	0.082206	0.068625	0.055044	0.041463	0.027882
	RCLASS	4:00			
X - Axis	15	16	17	18	19
b_0	-0.23954	-0.23954	-0.23954	-0.23954	-0.23954
b_1	0.022867	0.022867	0.022867	0.022867	0.022867
Death Rate	0.103465	0.126332	0.149199	0.172066	0.194933
	RCLASS	5:00			
X - Axis	20	21			
b_0	-0.21119	-0.21119			
b_1	0.016021	0.016021			
Death Rate	0.109229	0.12525			

APPENDIX C

Calculations for Confidence of Interval for Group 5

The formula used is as follows:

$$y \leq y_0 \pm t \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \overline{x})^2}{\Sigma(x - \overline{x})^2}} \cdot s$$

- Group 5/ Rclass 2.00 (Low Risk):
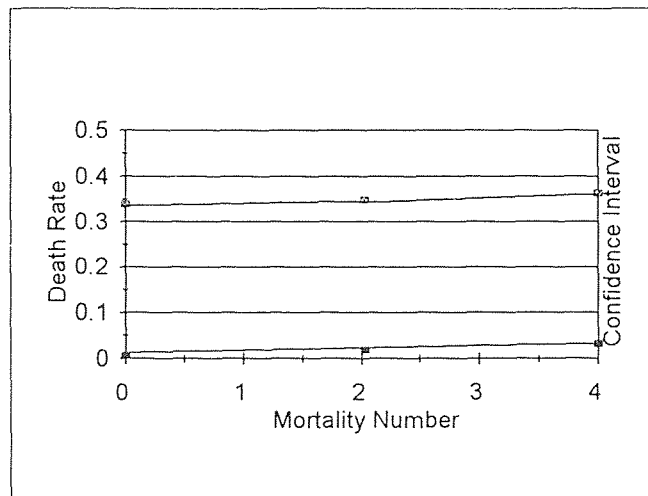
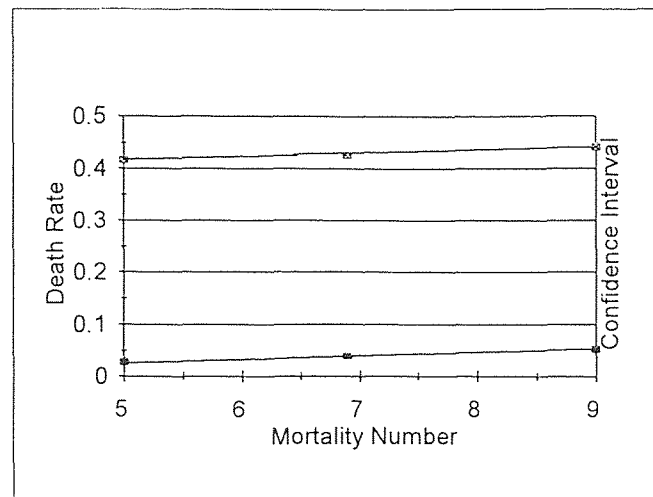$y_0 = 0.007440, 0.0202703, 0.0325$

$t = 1.96, n = 149$

$x_0 = 0, 2.033557, 4$

$\overline{x} = 2.0335$

After calculations:

$y = 0.287846, 2.0335, 0.31277$

- Group 5/ Rclass 2.00 (Moderate Risk):

y_o = 0.028176, 0.0401786, 0.053544
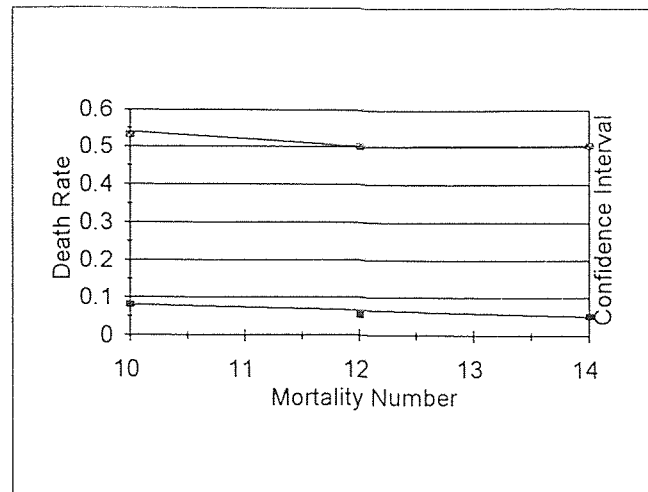
t = 1.96

n = 224

x_o = 5, 6.8928571, 9

$\bar{x}$ = 6.8928571

s = 0.19705

After calculations:

y = 0.04167397, 0.4272615, 0.4424611

-Group5 / RCLASS 3.00 (Intermediate Risk):

y_o = 0.082206, 0.0547945, 0.0.053544
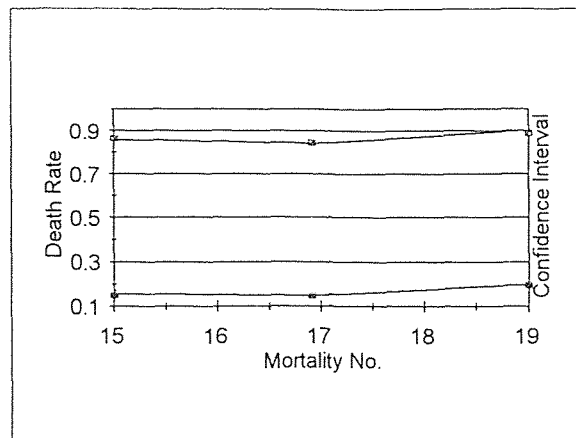
t = 1.96

n = 120

x_o = 10, 12.0090909, 14

$\overline{x}$ = 12.0090909

s = 0.22732

After calculations,

y = 0.5332043, 0.5022098, 0.504478

-Group 5 / RCLASS 4.00 (High Risk):

y_o = 0.149199, 0.1472393, 0.194933

t = 1.96

n = 162

x_o = 15, 16.9141104, 19

$\overline{x}$ = 16.9141104

s = 0.35513

After calculations,

y = 0.864115, 0.8454391, 0.8980345

REFERENCES

[1] Parsonnet, Victor, David Dean, Alan D. Bernstein: "A Method of Uniform Stratification of Risk for Evaluating the results of Surgery in Acquired Adult Heart Disease", *Circulation*, vol. 79, n.6, June 1989, (Supplement I), pp. I-3-I-12.

[2] Teng, Huey-Chung: *"Bayseian Methods in Preoperative Risk Assessment for Cardiac Surgery"*, Masters Thesis, New Jersey Institute of Technology, Newark, NJ, May 1993.

[3] Zadeh, L. A.: "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes", *IEEE- Transactions on Systems, Man and Cybernetics*, vol. SMC-3, 1973, pp. 28-44.

[4] Junod, Forrest L., Bradley J. Harlan, Janie Ayne, et al: "Preoperative Risk Assessment in Cardiac Surgery", *J ANN THORAC SURG*, vol. 43, n.1, January 1987, pp. 59-64.

[5] Williams, Sankey V., David B. Nash, Neil Goldfarb: "Differences in Mortality from Coronary Artery Bypass Graft Surgery at Five Teaching Hospitals", *JAMA* , vol. 266, n.6, August 1991, pp. 810-815.

[6] Kennedy, J. Ward, George C. Kaiser, Lloyd D. Fisher, et al: "Multivariate Discriminant Analysis of the Clinical and Angiographic Predictors of Operative Mortality from the Collaborative Study in Coronary Artery Surgery (CASS)", *J THORAC CARDIOVASC SURG*, vol. 80, n.6, December 1980, pp. 876-887.

[7] Psaty, Bruce M. ,Curt D. Furberg, Lewis H. Kuller, et al: "Isolated Systolic Hypertension and Subclinical Cardiovascular Disease in Elderly", *JAMA* , vol. 268, n.10, September 1992, pp. 1287-1292.

[8] Draper N.R, H. Smith , *"Applied Regression Analysis"*, John Wiley and Sons, NY, 1981.