

ABSTRACT

The problem addressed is whether patient-assessed outcomes can be used to improve subsequent patient outcomes through altering procedures and protocols. Despite a growing awareness amongst some health professionals of the need for 'real time science in medicine' - that is, routinely linking patient-assessed outcomes to the processes of care - there are no such studies reported in the literature.

A continuous quality improvement (CQI) model for surgical procedures incorporating patient-assessed outcomes is proposed. The principles of CQI are demonstrated with a model formulated, tested and refined with data collected for coronary artery bypass graft surgery (CABG) at a major Australian teaching hospital.

This database records pre-surgery, peri-surgery and post-surgery details for 3979 CABG between 1984 and 1993. The perfusion process is described with emphasis on the main measures - blood gases, blood pressure and temperature - under the control of the perfusionist. A follow-up survey administered to patients one year after surgery has two aggregate scores of patient health outcomes - one assesses neurological outcome (NSUM) whilst the other assesses physical outcome (PSUM).

Descriptive, univariate and multivariate analyses of these data have been conducted to determine associations with explanatory variables. Graphical procedures are used to illustrate associations between outcome measures and physiological variables. The univariate analyses has been carried out

using SPSSX for MAC and the multivariate analyses uses the SAS LOGIST procedure on a VAX mainframe.

As NSUM and PSUM are classified into ordered categorical data, a multivariate statistical model to explain variations using ordered polytomous logistic regression is employed to analyse these data. The variables that have association with increasing NSUM or neurological deficit and increasing PSUM or physical deficit are presented and discussed.

The results of these statistical analyses are used in a CQI model. The process of improvement is demonstrated with improved neurological and physical outcomes over time. A general model for other surgical procedures is described. Various ways of presenting the relationship between outcome and process are presented for different methods of initiating change in CABG surgery practice and other surgical practice.

Areas of research significance are: multivariate statistical techniques (ordered categorical polytomous logistic regression) applied to a large database demonstrates the attribution of patient-assessed outcomes to surgical procedure.; validation of self-assessed measures (ordinal and categorical) provides an important contribution to the outcomes movement; and an outcome management process, where specific post-discharge measures of outcome have been used to continually develop protocols for the surgical process. All of the above represent significant contribution to both health outcome and health information systems research.

**PATIENT-ASSESSED HEALTH OUTCOMES IN A CYCLE OF
CONTINUOUS QUALITY IMPROVEMENT: A CASE STUDY
OF CORONARY ARTERY BYPASS GRAFT SURGERY**

DEBORAH ANN BLACK, B Sc, Dip Ed, M Stat

A thesis submitted in fulfilment of the requirements for the degree
of

DOCTOR OF PHILOSOPHY

**Central Queensland University
Faculty of Health Sciences**

November, 1996

**CENTRAL QUEENSLAND
UNIVERSITY - LIBRARY**

TABLE OF CONTENTS

TABLE OF CONTENTS.....	(i)
LIST OF TABLES	(vi)
LIST OF FIGURES	(viii)
ACKNOWLEDGMENTS.....	(x)
DECLARATION	(xi)

CHAPTER 1

Introduction

1.1 Problem Statement.....	1
1.2 Methodology and Significance.....	5
1.3 Organisation of Thesis.....	10

CHAPTER 2

Literature Review

2.1 Introduction.....	13
2.2 Defining and Measuring Outcomes.....	14
2.3 Health Outcomes for Coronary Artery Bypass Surgery	30
2.4 Health Information Systems.....	33
2.5 Health Information Systems Directly Linked to Health Outcomes.....	45
2.6 Damage Relating to Perfusion	49
2.7 Summary of Gaps in the Literature.....	51

CHAPTER 3

The Perfusion Process

3.1	Introduction.....	54
3.2	Definition of Perfusion.....	54
3.3	Background to the Perfusion Process	55
3.4	Description of the Process at the Study Hospital.....	55
3.5	The Oxygenator	61
3.6	Acid-Base Management.....	64
3.7	Summary.....	65

CHAPTER 4

Method

4.1	Introduction.....	66
4.2	Development of an Instrument to Measure Patient Assessed Outcomes for CABG Patients.....	67
4.3	Pilot of Instrument.....	76
4.4	Administration of the Instrument and Data Entry.....	76
4.5	Perfusionist Data Analyses	78
4.6	Data Transfer.....	79
4.7	Summary.....	80

CHAPTER 5

Univariate Analyses of Coronary Artery Bypass Graft Patient Data

5.1	Introduction.....	82
5.2	Description of Respondents.....	82
5.3	Neurological Deficit	94
5.4	The Effect of Blood Gas Levels on Neurological Outcome.....	99
5.5	Physical Deficit.....	109
5.6	Comparison of Neurological and Physical Deficits.....	118
5.7	Summary.....	120

CHAPTER 6

Multivariate Analyses of Coronary Artery ByPass Graft Patient Data

6.1	Introduction.....	122
6.2	Critique of Univariate Analysis.....	122
6.3	Logit Models.....	124
6.4	Logistic Regression	125
6.5	Comparing Models.....	134
6.6	Loglinear Models for Ordinal Variables	135
6.7	Multinomial Response Models.....	139
6.8	Ordered Polytomous Models.....	140
6.9	Proportional Odds Model.....	144
6.10	SAS LOGIST Procedure and Applications.....	147
6.11	Application of Ordered Polytomous	

Regression Models to Coronary Artery	
ByPass Graft Patient Data.....	153
6.12 Discussion of both Models.....	164
6.13 Summary.....	165

CHAPTER 7

A Model for Continuous Quality Improvement

7.1 Introduction.....	167
7.2 TQM in Industry	168
7.3 TQM in Health Care.....	170
7.4 Principles for Design and Use of Models.....	177
7.5 The Continuous Quality Improvement Model.....	182
7.6 Summary.....	211

CHAPTER 8

Conclusions.....	212
------------------	-----

BIBLIOGRAPHY	221
--------------------	-----

APPENDICES

Appendix A	SPSSX Command File for pH.....	A1
Appendix B	SPSSX Command File for Partial Oxygen	B1
Appendix C	SPSSX Command File for Partial Carbon Dioxide	C1
Appendix D	SPSSX Command File for Haemoglobin.....	D1

Appendix E	SPSSX File for Base Excess.....E1
Appendix F	SAS LOGIST Commands for NSUM with all VariablesF1
Appendix G	List File for SAS LOGIST Procedure for NSUM with all Variables.....G1-G3
Appendix H	SAS LOGIT Commands for NSUM with Selected Variables.....H1
Appendix I	List File for SAS LOGIST Procedure on NSUM with Selected Variables.....I1-I5
Appendix J	SAS LOGIST Commands for PSUM with all Variables.....J1
Appendix K	List File for SAS LOGIST Procedure on PSUM for Selected Variables.....K1-K3
Appendix L	SAS LOGIST Commands for PSUM with all VariablesL1
Appendix M	List file for SAS LOGIST Procedure for PSUM with Selected Variables.....M1-M6
Appendix N	Cardiopulmonary Bypass Patient Form at Study Hospital.....N1-N2

LIST OF TABLES

Table 2.1	Core Concepts and Domains of Health-Related Quality of Life.....	18
Table 4.1	Categories for Responses to Neurological Damage...	70
Table 4.2	Categories for Responses to Physical Outcome.....	71
Table 5.1	Comparison of Selected Variables from CABG Follow-Up Studies.....	84
Table 5.2	Distribution of Previous Operations.....	87
Table 5.3	Illness Prior to Surgery.....	87
Table 5.4	Distribution of Body Mass Index for Patients.....	89
Table 5.5	Work Status of Patients One Year after Surgery by Age.....	90
Table 5.6	Work Status of Patients 1 Year after Surgery by Sex.	91
Table 5.7	Geographic Distribution of Respondents.....	92
Table 5.8	Distribution of NSUM Scores.....	95
Table 5.9	Analyses of the Effect of Selected Variables on NSUM.....	97
Table 5.10	Distribution of PSUM Scores	110
Table 5.11	Analyses of the Effect of Selected Variables on PSUM.....	112
Table 5.12	Respondents Description of Physical State	116
Table 6.1	Cancer Remission & Labelling Index.....	128
Table 6.2	Empirical Logits for Table 6.1 and Predicted Number of Remissions.....	128
Table 6.3	Death Penalty Verdict by Defendant's Race and Victim's Race.....	131

Table 6.4	Residual Degrees of Freedom for Loglinear Models for Three-Way Tables.....	133
Table 6.5	Final Model Estimated Parameters and Estimated Standard Errors for the Dumping Syndrome Data.....	149
Table 6.6	Ordered Drinking Categories from the Study by Ashby and Others	150
Table 6.7	Distribution of NSUM Categories.....	154
Table 6.8	Description of Explanatory Variables.....	155
Table 6.9	Pearson Correlation Matrix for Explanatory and Response Variables	156
Table 6.10	Maximum Likelihood Estimates for response NSUM with a Five Category Ordered Logistic Model and Five Explanatory Variables.....	158
Table 6.11	Goodness-of-fit for Five Category Model - Observed and Expected Frequencies for Respondents in each NSUM Category for each Logit Interval.....	159
Table 6.12	Distribution of PSUM Categories.....	161
Table 6.13	Maximum Likelihood Estimates for Response PSUM with a Five Category Ordered Logistic Model and Six Explanatory Variables	162
Table 6.14	Goodness-of-fit for Five Category Model - Observed and Expected Frequencies for Respondents in each PSUM Category for each Logit Interval.....	163

LIST OF FIGURES

Figure 1.1	Clinical Information System for Surgery in Acute Hospital.....	2
Figure 1.2	Outline of Continuous Quality Improvement Model to be Developed.....	6
Figure 2.1	The Architecture of Expert Systems.....	38
Figure 2.2	Simple Decision Tree Model.....	42
Figure 2.3	Outcome Management Process.....	46
Figure 3.1	Cardiopulmonary Bypass.....	57
Figure 3.2	Heart-lung Machine Configuration at Study Hospital	58
	Figure 3.2A	59
	Figure 3.2B.....	60
Figure 3.3	Cardiopulmonary Bypass Circuit at Study Hospital.....	61
Figure 3.4	Cross-sectional Photographs of Oxygenator.....	63
Figure 3.5	Set-up of Equipment in the Study Hospital.....	63
Figure 5.1	Distribution of Respondents by Age.....	85
Figure 5.2	Relative Proportions of Males and Females by Age Groups.....	86
Figure 5.3	Frequency of Coronary Artery Grafts (1989 to 1992)	88
Figure 5.4	Proportion of Respondents with NSUM=0 for each Previous Illness	98
Figure 5.5	pH by Time by Mean NSUM.....	101
Figure 5.6	PO ₂ by Time by Mean NSUM	102
Figure 5.7	PCO ₂ by Time by Mean NSUM.....	103

Figure 5.8	Arterial Pressure by Time by Mean NSUM	104
Figure 5.9	Haemoglobin by Time by Mean NSUM	105
Figure 5.10	Base Excess (BE) by Time by Mean NSUM	106
Figure 5.11	Proportion of CABG Patients with NSUM=0 by Year of Surgery.....	109
Figure 5.12	Proportion of Respondents with PSUM=0 by Previous Operation	113
Figure 5.13	Proportion of Respondents with PSUM=0 by Illness Category.....	114
Figure 5.14	Proportion of Respondents Free of Angina by Year..	117
Figure 5.15	Proportion of Respondents with PSUM=0 by Year ..	118
Figure 7.1	The Shewart Cycle	169
Figure 7.2	Cycle for Improving Performance.....	172
Figure 7.3	Model of the 'Serial V' Concept	174
Figure 7.4	Clinical Value Compass	175
Figure 7.5	Factors Contributing to Patient's Health Outcome ..	178
Figure 7.6	Model Development as an Iterative Process.....	180
Figure 7.7	Outline of Model to be developed.....	182
Figure 7.8	Suggested Format for Follow-Up Interview Instrument	194
Figure 7.9	pH by Time by Mean NSUM.....	196
Figure 7.10	Model for Data Analyses	200
Figure 7.11	Example of Use of Model	201
Figure 7.12	Report Format A for Peer Review	206
Figure 7.13	Report Format B for Peer Review	207
Figure 7.14	Exception Report Sample Format A.....	208
Figure 7.15	Exception Report Sample Format B.....	209

ACKNOWLEDGMENTS

The author acknowledges the work of the perfusionist at the Australian teaching hospital in Sydney and for his permission both to use the data that he has collected over the 11 years and to describe the procedures that he has implemented to improve patient outcome. I would like to endorse his continuing dedication to the improvement of his patient's quality of life and thank him for our lively discussions on these important issues. The hospital's research and ethics committee has given permission for the patient data to be used for research and to be published. One of the conditions of approval is that confidentiality is maintained by not identifying the hospital.

I thank Dr Evelyn Hovenga (Central Queensland University) and Dr Peter Cooke (Department of Statistics, University of New South Wales) for their supervision of this research. Professor William Dunsmuir, Head of the Department of Statistics, University of New South Wales provided valuable insights into categorical data analyses. Thanks also to Dr Tessa Ho, School of Health Services Management, University of New South Wales and Dr Judy Branch, anaesthetist at St Vincents Hospital, Sydney for their comments on an earlier draft of Chapter 7. My two colleagues at the University of New South Wales, Professor George Palmer and Mr Jeffrey Braithwaite have alerted me to some important references.

For their support and understanding during the period that this research was undertaken, I thank my husband, John, and children Belinda, Sean, James and Jordan. I dedicate this thesis to my parents, Ron and Joyce Wilmot.

DECLARATION

I hereby declare that this submission is my own work and that, for the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent, has been accepted for the award of any degree or diploma of a university or other institute of higher learning, except where due acknowledgment is made in the text.

Sign: _____

Date: _____

CHAPTER 1

INTRODUCTION

CHAPTER 1 INTRODUCTION

1.1 Problem Statement

Health service providers argue consistently that improved information systems will assist in providing more efficient ways of rationing health services (for example, Classen *et al.*, 1991; Kerr and Jelinek, 1990; Martin, 1990; and Wrigley, 1990). An important question less frequently asked about such information systems is: will improved health information systems make the services more effective? This research aims to answer this question by demonstrating that improved information systems, which are an integral part of a continuous quality improvement model, will result in more effective decision making leading to improved patient outcomes.

There are a number of specific questions to be addressed in the thesis. First, can patients accurately assess their health outcomes? Secondly, does multivariate statistical models play a role in the development of continuous quality improvement models? Thirdly, can patient-assessed outcomes be used in a cycle of continuous quality improvement?

If health information systems are to make a practical contribution to the health system then there is, according to De Lone and McLean (1992), a need to measure the output concisely. As discussed extensively in the literature, there is now an emphasis on 'customer-focus' in health systems in Australia and overseas. The 'customer' in the health system is the patient; so it is argued that the output should be measured in terms of patients' health outcomes.

“If information system research is to make a contribution to the world of practice, a well-defined outcome measure (or measures) is essential. It does little good to measure various independent or input variables...if the dependent or output measure...cannot be measured with a similar degree of accuracy” (De Lone and McLean, 1992, p. 61).

The thesis is that a link between health information systems and better health outcomes can be established, and that such a link can be demonstrated convincingly by careful examination of a specific health information system and patients records, especially with the use of standard univariate and advanced multivariate statistical techniques. The focus of this research is on patients undergoing surgical procedures in acute hospital settings, and the analysis undertaken is on patients who have had coronary artery bypass graft surgery. One problem with the present frameworks for clinical information systems for surgery in acute hospitals is that they are deficient with respect to the inclusion of patient outcomes. The usual model for an information system for patients undergoing surgery is illustrated in Figure 1.1.

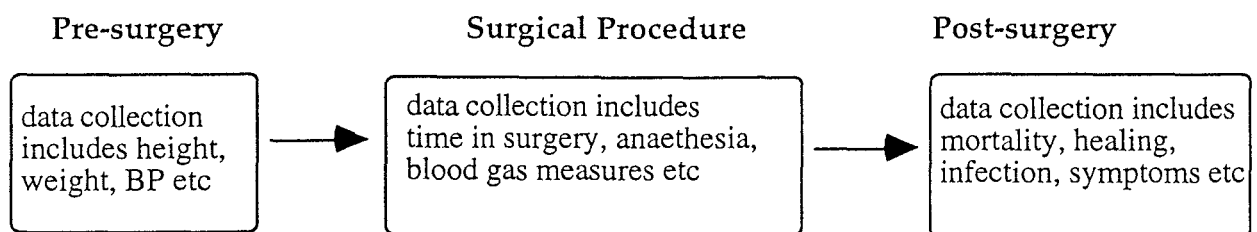


Figure 1.1: Clinical Information System for Surgery in Acute Hospital

In this framework, there are three stages of interaction with the patient: (a) in the first stage, the patient has a number of physical measures taken and recorded pre-surgery which are used to determine anaesthesia and other surgical techniques; (b) the actual surgical procedure takes place in the second

stage, where various blood gas, blood pressure and other physical measures of the patient are recorded during the procedure; and (c) the post-surgery stage, which includes the post-surgical hospital stay and follow-up visit with the consultant. The outcomes of the procedure are generally assessed in terms of mortality, wound healing, infection, other complications (ie co-morbidities) and length of stay. The inclusion of patient outcomes in clinical information system is essential: it is the *patients* who seek the surgical procedure because they have symptoms indicating a particular surgical procedure would be appropriate. It is obvious that a patient would wish to be free of those symptoms following the surgery and not have permanent neurological, anatomical or physiological damage that would limit their quality of life as a result of the surgical intervention. In most cases of major surgery a return to activities of daily living is not possible for some length of time following surgery: these periods of time will vary according to the type of surgery and characteristics of the patient. In all cases, the ideal, long-term, patient-assessed outcome would be a state of health equivalent to that prior to symptoms.

Despite these self-obvious statements, outcome measures are not currently included, routinely, in clinical information systems. The continuous quality improvement model proposed in this thesis therefore extends the conventional approach by including a follow-up with patients after an appropriate period of time has passed since surgery. The timing of the follow-up will be determined by the period of time that it is reasonable for the patient to return to normal activities of daily living for the particular type of surgery. The limitations of these daily activities will depend on the patient's age, co-morbidity, employment and leisure activities. There are, of course, many studies that do follow-up surgical patients when they return to normal

activities (Booth *et al.*, 1991; Lindal, 1990; Newman, 1989; Pinna Pinter *et al.*, 1992; Sarnquist, 1989; Shaw, 1989; Stein *et al.*, 1990; and Wannathemee and Shaper, 1991) but there is no evidence of any systematic framework for patient-assessed outcomes. Indeed, as recently as late last year, Hammermeister (*et al.*, 1995, p. OS5) in describing the rationale for the US Department of Veterans Affairs Cooperative Study, Processes, Structures, and Outcomes in Cardiac Surgery still found relatively few reports linking processes and structures of care to favourable outcomes. Data are not routinely collected and are based only on samples of patients being followed up in a one-off situation. The results of such studies are reported in the relevant literature and surgical practice and protocols may have been changed as a result of the findings of such studies. However, this practice is not explicit from the literature and it does appear that follow-up surveys are not part of routine procedures.

A number of clinical information systems in the United States have extended the conventional model shown in Figure 1.1. Examples of where outcome data are collected routinely and included in information systems, include the Sickness Impact Profile (Bergner *et al.*, 1981), the Outcome Index (Gilbert and Schoolfield, 1991) and the Medical Outcomes Study (Tarlov *et al.*, 1989; Ware and Sherbourne, 1992; and Ware, 1993). However, the weakness with these information systems is that they do not routinely link the patients' outcome when they return to normal activities back to the pre-, intra- and post-surgical clinical measures. Therefore, this thesis addresses this problem, by advocating a practice for surgical procedures in acute hospitals that is demonstrably effective.

1.2 Methodology and Significance

The methodology is an empirical study based on a longitudinal database of coronary artery bypass graft patients which includes their characteristics and health outcomes. Following statistical analyses to identify variables that explain patient-assessed outcome scores, a theoretical model for continuous quality improvement is developed.

In contrast to current practice with clinical information systems, the originality of the model developed in this research is that it continually links the post-discharge outcome measures to the peri-operative clinical measures and then examines these data in an aggregate form to see if any patterns of good or poor outcomes are associated with the clinical measures during surgery. When patterns emerge that have a statistically significant impact on health outcomes, surgical procedures or protocols can be altered to maximise the outcome measure. Figure 1.2 provides a diagrammatic representation of the conceptualisation of the outline of such a model to be further developed in this thesis.

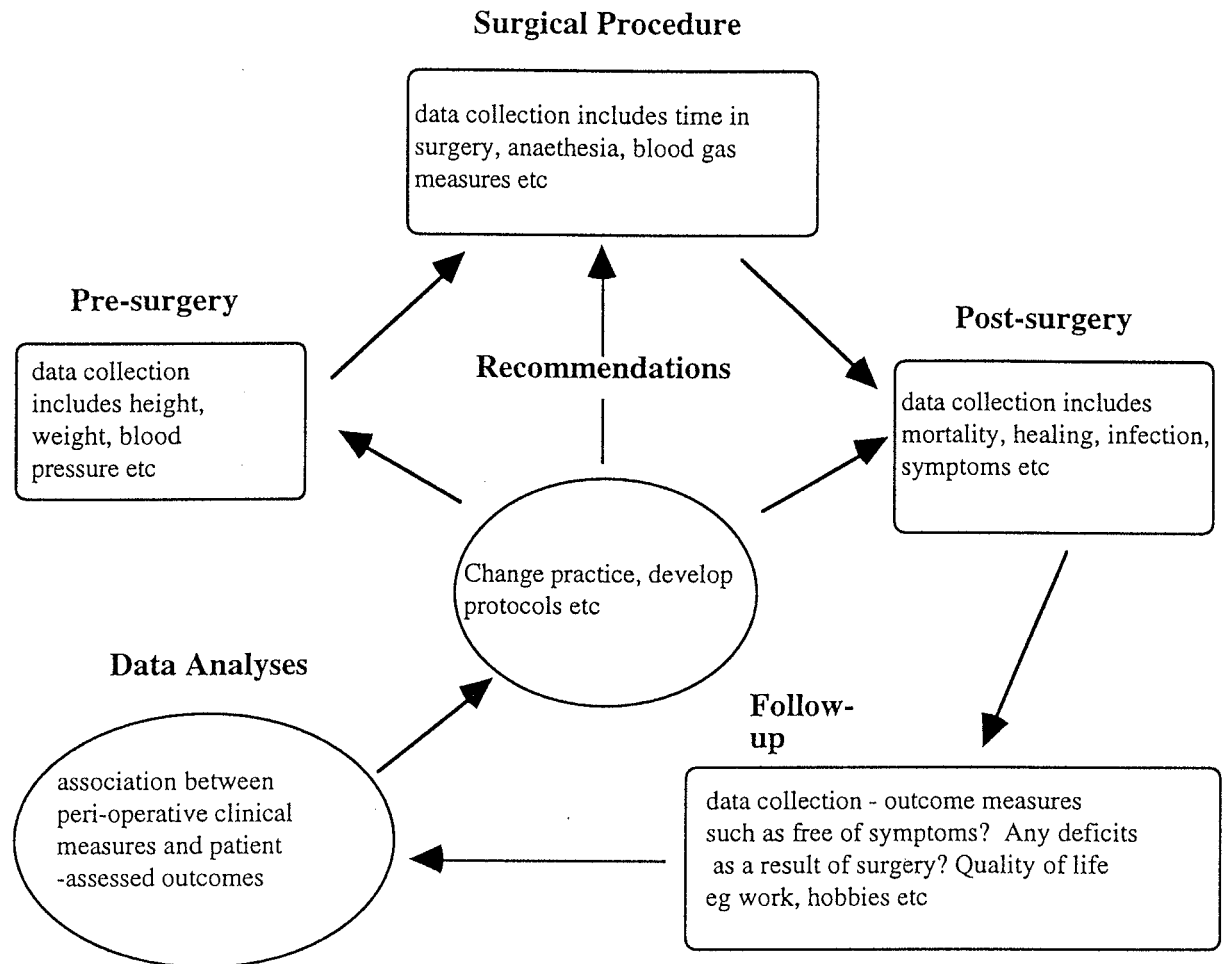


Figure 1.2: Outline of Continuous Quality Improvement Model to be Developed

Such a model was formulated after a wide-ranging literature review in the following areas: health outcomes; quality of life measures; outcomes for coronary artery bypass graft surgery patients; information systems in general and more specifically clinical information systems; total quality management in general and, more specifically, total quality management in health. Research gaps were identified, and these formed the basis of the model proposed in Figure 1.2.

To flesh out the model and to test the hypothesis that patient-assessed outcomes can be used in a cycle of quality improvement, the surgical procedure chosen as a case study was coronary artery bypass graft surgery. This is a high cost and high volume procedure. The surgical process used to illustrate the model in the case study was the perfusion process. To fully understand how this process might be controlled, the perfusion process has been studied and described in general terms with specific perfusion protocols illustrated at the case study hospital.

The follow-up part of the model in Figure 1.2 requires an appropriate measure of quality of life following surgical intervention. Methods used to develop an appropriate instrument to measure the impact of such an intervention on quality of life are described with particular reference to the development of the measure in the case study of coronary artery bypass graft patients.

the purpose of data analyses, in this study, was to test the hypothesis that variables that can be altered during the perfusion process can explain patient-assessed health outcome. data analyses, initially examined factors associated with outcome on a one-to-one basis (univariate analyses). Because outcomes assessed by patients invariably take the form of some sort of Likert scale - that is an ordered scale with more than two categories - research was extended to multivariate analyses. Suitable categorical regression models are examined before statistical software procedures appropriate to ordered categorical polytomous response variables were formulated and applied to the coronary artery bypass graft data.

The information system for the coronary artery bypass graft patients, includes clinical data collected, routinely, during the patient's episode of care, and follow-up data on not only the patient's post-surgery clinical health status but also how they personally perceive that health status. This study describes how those data are collected and recorded. This database has approximately 4000 patients' episodes of care documented for the period 1983 to 1993 and is continuously being updated.

This research describes how the results of the analyses of these data informed recommendations for changes to practice for all stages of pre-surgery, surgery and post-surgery. Although the development of recommendations for this surgical procedure focuses on the perfusion process, the methodological techniques employed in this thesis can be applied more generally to all areas of surgery.

This is the most comprehensive and extensive information system on coronary artery graft patients in Australia and, possibly, in the world. No others are documented in the published literature. The analyses of these data, using multivariate statistical techniques, form the empirical section of the research. The following three significant contributions to research in health care may be attributed to the findings arising from this thesis:

- First, the use of statistical techniques will demonstrate that the link between patient-assessed outcomes, other clinical-based outcome measures and the surgical procedure was not just due to chance. Therefore, the outcomes, both qualitative and quantitative, can be

attributed to the process. Attribution of outcomes to the peri-operative process can only be achieved if there exists a comprehensive database pre-intervention, during the intervention and post-intervention. Whereas the literature contains a wealth of hospital in-patient data it does not show that these data are routinely linked to post-discharge measures of patient health status.

- Second, to enable the attribution of outcome to the peri-operative process, there is a need to validate the use of patient-assessed outcomes measures, which are invariably based on ordinal variables in more than one category. Hence, statistical techniques are developed to handle data that are ordered and categorical. Although the patient-assessed outcome measures have been specifically designed for open-heart surgery, the statistical techniques developed will be generalisable to other procedures because they relate to the ability to carry out various life tasks. The literature review reveals that most clinicians either trust only clinical measures of outcome, or that they recognise the importance of self-assessed measures of health status, but maintain there are not valid tools for measurement. The validation of self-assessed measures, therefore, provides an important contribution to the measurement of outcome.
- Third, the process used by the perfusionist for continuous quality improvement (CQI) reported in this thesis is generalisable for most procedures. An outcome management process, where specific post-discharge measures of outcome are used to continually develop protocols for the peri-operative process, has been formulated and this represents a

significant contribution in areas of both health outcomes and of health information systems.

To summarise the significance of the research, it demonstrates that the maintenance of comprehensive clinical information systems can positively influence patient health outcomes. Also, the study defines and provides suitable measures of patient health outcomes for a given case study (a health information system for coronary artery bypass graft surgery). The research provides examples of databases from hospital patients with comparative measures of patient outcomes. Surrogates for outcome are defined and measured. In the past, patient outcomes in acute episodes of care have largely been measured objectively; not by considering and measuring individual patient perceptions of their health status. As there is also only limited international literature which has demonstrated the measurement of health outcomes in terms of patients' perspective (ie subjectively) of their health and functional status post-discharge, this research contributes to the international literature. It provides a benchmark study within the Australian health system.

1.3 Organisation of the Thesis

Chapter 2 is a comprehensive literature review describing and measuring: general outcome measures; outcome measures specific to surgery; and outcome measures specific to coronary artery graft patients. It further reviews both literature on health information systems and literature where information systems are linked to health outcomes. Gaps in the literature are

identified which have led to the conceptualisation of the model of continuous quality improvement shown in Figure 1.2.

Chapter 3 provides an overview of the perfusion process used in the case study hospital. As this thesis demonstrates how changes to the perfusion process can be used to improve patient outcomes, it is important to demonstrate an understanding of this procedure in terms of controlling blood gases, blood pressures and temperatures.

The development and piloting of the instrument to measure patient-assessed outcomes at the study hospital is described in Chapter 4. The method of data collection, validation and data analyses by the perfusionist is then presented. Finally, the method developed in this research study for transferring, transforming and analysing the coronary artery bypass graft data is given at the end of the chapter.

A summary of characteristics of approximately 4 000 patients in the database is presented in descriptive form at the beginning of Chapter 5. This chapter then identifies factors associated with the presence or absence of neurological deficit. A similar analysis is given for the presence or absence of physical deficit. It is finally demonstrated how the use of these descriptive statistics have been used to improve patient outcomes at the study hospital.

Chapter 6 presents the detailed statistical analyses that identifies characteristics of the patients that explain self-assessed outcomes. As the self-assessed outcomes are classified in five ordered categories, regression techniques appropriate to ordered polytomous categorical responses are

identified, described and used in their analyses. Essentially, the technique used in Chapter 5 to improve outcomes is validated in Chapter 6.

The final substantive chapter represents a synthesis of all the areas covered in the research. A detailed continuous quality improvement model that can be applied to other surgical procedures taking place in acute hospitals is proposed in Chapter 7. This is presented with particular reference to the procedure undertaken at the study hospital. Finally, Chapter 8 provides conclusions to this thesis, and suggests areas for further research.

CHAPTER 2

LITERATURE REVIEW

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

Prior research in the area related to this thesis has been organised into five subsections. Firstly, a brief review of literature on measuring patient outcomes will be presented. Secondly, a review of literature examining health outcomes for the proposed area of study - coronary artery bypass graft surgery (CABG) will be covered. Thirdly, a summary of relevant health information systems literature will be given. Fourthly, literature that examines health information systems, and their effects on health outcomes, will be discussed. Finally, a brief review of literature on neurological damage as a result of perfusion is discussed. The sub-sections were chosen to initially provide an overview of literature in the areas of health outcomes and health information systems and then followed by a more detailed review in areas specific to the thesis topic. The literature search methods were based on keyword searches on information system, health and medical CD-Rom databases; and searches on specific journal. From this review the gaps in the literature will be identified based on a critical assessment of the first four major subsections.

It should be noted here that relevant statistical theory literature is discussed in Chapter 6. Also, the discussion of literature relating to developing continuous quality improvement models will be covered in Chapter 7. It is more logical that these literature reviews be included in these chapters so as to provide the results within the context of the relevant theory for the respective chapters.

2.2 Defining and Measuring Outcomes

It is important to have a clear definition of patient or health outcomes. Linder (1991, p.21) describes outcome measurement as "results". These measurements range variously from death and disease to emotional health and patient satisfaction. Fries and Spitz (1990, p.25) also propose that outcome measures are multi-dimensional and state:

"Patient health outcome usually refers to a final health status measurement after the passage of time and the application of treatment. In the future, patient outcome will be increasingly described by a cumulative series of health status measurements".

Wilkin (et al., 1993) define outcomes in public health as achievement or failure to achieve specified goals. Outcomes can therefore be positive or negative.

In their study, patient or health outcomes embrace both the emotional and physical health status of the patient following treatment. Patient outcomes come in many forms - physiological terms, physical terms, mental or psychological terms and social terms (Batalden, et al., 1994). It is important to note that the outcome measure is relative to the pre-treatment health status of the patient and the expectation of well-being, given age, gender and other socioeconomic factors. This was reinforced in a study of outcomes for

acutely-ill patients where mortality rates were adjusted for pre-existing risk factors.

To expand and explain this problem, consider an example of elderly patients, following open heart surgery. They would have far more limited expectations of their physical activities than, say, a 35-year old following the same surgery. This important point leads on to the associated measure - "quality of life".

There are also a number of measures of quality of life following treatment that have taken the prior health status of the patient into account. Both Schipper (*et al.*, 1990) and Donabedian (1981), discuss the importance of a standard by which to measure outcome. Donabedian noted the need to be flexible when defining standards and not to adopt an "all or nothing standard". The standard suggested by Schipper (Calman's Gap) takes the patient's expectation of quality of life as a standard, and then measures the gap between patient expectations and achievements. The smaller the gap the greater the quality of life.

An example measuring such a gap is given by Sager (*et al.*, 1996). Their study was of 1 279 patients who had been hospitalised for acute medical illness. A pre-admission measure of functioning formed the baseline for functional measures at discharge using Activities of Daily Living and at three months using Activities of Daily Living and Instrumental Activities of Daily Living. They concluded that there was a high incidence of functional decline in the cohort three months after hospitalisation and that research should be conducted to find out what was influencing this decline. In other words,

what part of the process of care was changing the outcome and could they be controlled?

A method frequently used to assess quality of life are quality of life years (QALYs). QALYs are used as a way of describing the social value of health care benefits. They combine the value of life and life expectancy into a single numerical index and are mainly used in relation to the allocation of scarce health care resources. Carr-Hill and Morris (1991) state that one year of current life in perfect health is counted as one QALY. Lower quality of health, "q" ($q < 1$) is counted as q QALYs and future years of life are reduced at a rate of "r". The classification system for QALYs was based on valuations of 29 health states by Rosser and Kind (1978).

QALYs have been criticised, for example, by Nord (1992), because of their emphasis on health status rather than health improvements. Nord believes this approach has three major problems: that quality of life measured in numbers carries little meaning to people; there is some stigma attached to giving less value to a life with a disability than one without a disability; and the health status index emphasises the size of the health improvement without reference to the health status start or end point. Nord is concerned with the measurement of quality of life in years rather than quality of life in people. Thus, he proposed to introduce a new procedure called the saved young life equivalent (SAVE). A SAVE measures social value in that 1 SAVE is equivalent to saving the life of a young person and returning them to full health. One SAVE is the maximum possible benefit with all other interventions being measured against a SAVE. Nord argues a SAVE is a social value for health gains and that it is easy to understand this standard -

that is, the value of saving a young life. These are, in fact, measures for rationing health services.

Another measure of quality of life, Q-TWIST, is given by Schwartz (*et al.*, 1995) for multiple sclerosis patients. Q-TWIST is an acronym for Quality-adjusted Time Without Symptoms and Toxicities. Traditionally the emphasis for quality of life measurement for multiple sclerosis patients is ambulation. However, the authors maintain that patients believe this is not the major contributor to quality of life. Q-TWIST uses patients' experiences and specifically measures the trade-offs between negative and positive aspects of their treatment. The authors state:

"when clinical trials rely on observers' assessments of the patients experience, ignoring the patients' point of view, ie, the ultimate usefulness of the tested therapy may not be highlighted" (Schwartz *et al.*, 1995, p. 755).

One measure of quality of life using a multi dimensional conceptualisation is proposed by Patrick and Erikson (1993, p. 77) and this is summarised in Table 2.1. Patrick and Erikson define five categories of concepts and domains of health - opportunity; disadvantage; health perceptions; functional status; impairment; duration of life and death. The views of the authors reinforce the concept that health outcome can only be measured within the constraints of the patient's environment and their perception of their health.

Table 2.1: Core Concepts and Domains of Health-Related Quality of Life

CONCEPTS AND DOMAINS	DEFINITIONS/INDICATORS
OPPORTUNITY	
Social or cultural disadvantage	- Disadvantage because of health; stigma; societal reaction
Resilience	- Capacity for health; ability to withstand stress; physiological reserves
HEALTH PERCEPTIONS	
General health perceptions	- Self-rating of health; health concern/worry
Satisfaction with health	- Satisfaction with physical, psychological, social function
FUNCTIONAL STATUS	
Social Function	
Limitations in usual roles	- Acute or chronic limitations in usual social roles (major activities) of child, student, worker
Integration	- Participation in the community
Contact	- Interaction with others
Intimacy and sexual function	- Perceived feelings of closeness; sexual activity and/or problems
Psychological Function	-
Affective	Psychological attitudes and behaviours, including distress and well-being
Cognitive	- Alertness; disorientation; problems in reasoning
Physical Function	
Activity restrictions	- Acute or chronic reduction in physical activity, mobility, self-care, sleep, communication
Fitness	- Performance of activity with vigour and without excessive fatigue
IMPAIRMENT	
Symptoms/subjective complaints	- Reports of physical and psychological symptoms, sensations, pain, health problems or feelings not directly observable
Signs	- Physical examination; observable evidence of defect or abnormality
Self-reported disease	- Patient listing of medical conditions or impairments
Physiological measures	- Laboratory data, records, and their clinical interpretation
Tissue alterations	- Pathological evidence
Diagnoses	- Clinical judgments after "all the evidence"
DEATH AND DURATION OF LIFE	- Mortality; survival; years of life lost

The absence of much literature pre-1990 raises the question as to why there has been a recent focus on measuring health outcomes? Ellwood (1988) maintains that the trend began, in fact, in the United States as early as 1969 when the President proclaimed that health was in crisis because increasing costs for health care were becoming unsustainable. With patients challenging medical decisions and increasing costs there was a concern with the efficacy of expensive medical procedures. A need arose to contain costs whilst demonstrating a maintenance of quality of care that is a demonstration of both medical efficiency and effectiveness (Chute, 1992).

In 1988, the Health Care Financing Administration, and the recently established Agency for Health Care Policy and Research, began a program which aimed at gauging the effectiveness of medical intervention and developing guidelines for medical practice through the assessment of patient outcomes (Epstein, 1990). Part of this program was the introduction of Programmed Outcome Research Teams (PORTs) targeting specific disease entities and introducing the concept of managed care. In fact Nissen (1993) noted that the US state of Minnesota introduced legislation, in 1992, which would change how health care providers conduct business. The health care delivery system became accountable to the public by exposing and comparing treatment outcomes.

Guadagnoli and McNeil (1994) define the stakeholders in outcomes research as payers and providers. All three have played a role in the design of outcomes research. Payers wish to reduce costs based on effective results. Hospitals and health care organisations use effectiveness research to produce

protocols and guidelines. Providers give effectiveness information to patients for them to make informed treatment decisions and patients respond to the effectiveness data to make actual treatment decisions.

Outcome measurement has been introduced to provide a solution to the spiralling health costs in the USA, as emphasised by Linder (1991, p. 23):

"A cynical description of today's state of affairs is that punitive witch-hunting regulators vainly attempt to inspect an entrenched clan of professionals who protect, but do not discipline, each other despite delivering inadequate or inappropriate services to customers who cannot tell what they are getting for ever-increasing prices. In theory, outcome measurement promises a remedy for this ineffective system" .

Outcomes research has provoked some resistance from the medical profession. This resistance is summarised in a statement by Gulliford (1992, p. 325) where he states:

"unequivocal measurement of health care outcomes can only be made in clinical trials".

A systematic review of the effectiveness of health care, which is based, in most instances on the results of clinical trials, is given by the Cochrane Collaboration (Bero and Rennie, 1995; Silverman, 1995). Archie Cochrane emphasised the importance of care rather than cure - a philosophy based on

his experiences as a doctor in a German prisoner of war camp. The Cochrane Database of Systematic Reviews has as its objective, the provision of information that is needed by physicians to make clinical decisions. The Collaboration is an international group of individuals "committed to preparing, maintaining, and disseminating systematic review of the effects of health care"(Bero and Rennie, 1995, p. 1935). However, the reviews are limited to specifically defined methods that ensure the exclusion of bias - that is, they are only based on clinical trials not on routinely collected patient data. Another term often used for this type of procedure is evidence-based medicine.

Gage (1993) noted that randomised clinical trials are acceptable for outcomes research but research based on actual practice is seen as a threat to professional autonomy. It is also noted by Gage that patient care is not just 'a series of doctor-patient encounters subject to purely external constraints but a coordination of services with a need to share data'. Epstein (1990) has suggested three factors that have reinforced such an emphasis on measuring outcome: cost-containment and the concern that initiatives to control increasing medical services might negatively affect quality of care; there was a renewed sense of competition in the health system, and that competition on the basis of price was inadequate; and the documentation of variations in practice across different geographic areas demonstrated a need for defined outcome measures for comparison.

A study that examined regionalisation, cost and outcome (measured as mortality) for pancreaticoduodenectomies was reported by Gordon (*et al.*, 1995) in Maryland, USA. The results demonstrated an inverse relation between

mortality and volume and between cost and volume. The study concluded that the major explanation for better outcomes and lower costs was the more standardised approach of an expert team of staff at the high volume institutes.

Other studies (Anderson and Lomas, 1989; Wennberg *et al.*, 1980; Wennberg and Gittelsohn, 1982; Keller *et al.*, 1990; and Renwick and Sax, 1991) have noted variations in both utilisation and cost of medical and surgical procedures between different areas. In a study of geographical differences in utilisation rates of a number of common surgical procedures in the United States, Wennberg (*et al.*, 1980, p. 278) noted:

"The significance of the different levels of population exposure to the procedures cannot be fully assessed because of incomplete evidence on outcomes... We argue that the cost implications of professional uncertainty justify a substantial investment to clarify the advantages and disadvantages of the various approaches to common medical problems".

It is acknowledged that classifying patients in terms of diagnosis-related groups (DRGs) is a way of measuring outputs not outcomes. However, as DRGs classify patients' episodes of care to be not only resource homogeneous but also to be clinically coherent, patient outcomes disaggregated by DRGs, may be measured and compared. The decrease in lengths of stay accompanying the implementation of DRG reimbursement under the Medicare prospective payment system (PPS) prompted Stearns (1991) to examine clinical outcomes of New Jersey patients admitted with congestive

heart failure and prompted Mirin and Namerow (1991) to develop a more reliable measure of outcome for the increasing cost of mental health care.

Three ways of measuring patient health outcomes are suggested: objectively using clinical measures; subjectively using patients' perceptions; and a combination of both subjective and objective measures. Examples from the literature of the use of each measurement technique are described in the next section.

Whilst acknowledging patient satisfaction as an important indicator of outcome, Linder (1991) focused on clinical outcome in her study of 31 hospitals. Outcomes were measured according to disease severity. A frequently used measure of outcome, which is the most basic and supposedly objective, is mortality rate. Gilbert and Schoolfield (1992) describe a clinical outcome measure based on five simple clinical components which assess mortality risk - the system outcome score (SOS). They propose that $\frac{\text{SOS}}{\text{mortality rate}}$ should equal a constant unless the quality of care changes.

Luft (1980) outlined problems in comparing outcomes in terms of mortality rate. He found when taking outcomes in the aggregate and comparing mortality rates for a given procedure between hospitals that the rate was inverse to the volume of cases for that hospital and procedure. Was the outcome affected by the physicians experience or some regional factor? A less rigid measure described by Williams (*et al.*, 1992, p 1776) relates outcome measures to structures and processes:

"If variations in outcomes can be linked to variations in the provision of services, providers and managers can acquire a rationale and bench marks improving the quality of care".

Brook and Appel (1973) describe this process as what the physician does on behalf of the patient, such as diagnostic investigations, therapeutic interventions, source of medical care and the patient's compliance.

Designing appropriate outcome measures is obviously very difficult. Ellwood (1988, p. 1554) believes that part of the problem associated with health outcomes is that there are data only on those individuals who have sought treatment:

"At the outset, outcomes measurement lacks a denominator. It is only based on observations made on people who happen to use medical care. It will not always contain enough information about the natural history of the untreated condition."

Further difficulties were emphasised by Von Korff (et al., 1992) when stating that the determinants of health behaviours and health outcomes are multi-level. They can be both individual and environmental - that is aggregate variables. This difficulty was also acknowledged by Williams (et al., 1992) when they state that clinical epidemiologists have been sceptical about the inferences drawn by "proxy" or "surrogate" measures of outcome that rely less on biochemical, radiographic and physiological tests and more on measures of long-term health and health status. Formal trials, such as

double-blind placebo controlled trials, try to maintain equivalent structure and process whereas the outcome movement is about making services work routinely as well as they do in trials.

Fletcher (*et al.*, 1992) categorised instruments to measure health outcomes into four areas: generic, such as the Nottingham Health Profile; disease specific; dimension specific; and items specific to study. Other examples of protocols measuring health status include the Sickness Impact Profile (Bergner *et al.*, 1981). This is a behavioural-based health status measure. It is claimed it is efficient across various demographic and cultural groups.

Subjective measures of patient outcomes cannot be discussed without reference to patient satisfaction. In fact, Weisman and Koch (1989) maintain that patient satisfaction is an outcome of care. In a patient satisfaction study at Brigham and Women's Hospital, Boston, USA, Cleary (*et al.*, 1989) noted that perceived health was a strong predictor of overall patient satisfaction. Similarly, Pascoe (1983, p. 189) stated that degree of satisfaction, along a continuum from "totally satisfied" to "totally dissatisfied", is a dependent measure of the structure, process and outcome of service.

Notwithstanding this, Scott and Smith (1994) caution about the use of patient satisfaction surveys as outcome measures as they only measure part of the outcome of care. They state:

"The relationship between satisfaction with the process of care and satisfaction with health status is complex... Furthermore, if patient satisfaction is used as an outcome

measure in a competitive market then it is necessary to evaluate the behaviour of both purchasers and providers to assess the way they use satisfaction surveys relative to the way they use measures of health status and health-related quality of life" (Scott and Smith, 1994, p. 358).

They conclude that patient satisfaction measures only part of an outcome measure and that outcomes should be maximised according to what the patient values in these measures. In other words, a patient can be seen to have a high satisfaction score but the items measured in the score may not be of value to the patient. Thus, they are not adequate measures of outcome. Scott and Smith (1994) further state that a hospital with the highest satisfaction may also be the hospital with the worst outcomes.

Aharony and Strasser (1993) criticise subjective measures because they lack a standardised approach to patient satisfaction and that they are not clear in the determination of patient satisfaction. They acknowledge the importance of patient satisfaction in evaluation and planning but state there is little evidence that demonstrate the use of patient satisfaction monitors effects the way health care is delivered. In the implementation of any survey, it is important to demonstrate how much the opinion of the respondent is valued and how their contribution, through participation in the survey, will be used to affect change. Aharony and Strasser also question the validity of patient satisfaction surveys because of the lack of expert technical knowledge in patients and the accuracy of their statements, given their physical and emotional states.

They further cite the problem of patients' inability to differentiate between a good bedside manner and technical quality. However, they admit there can be a placebo effect on patients with better medical interactions. Although research has not found a simple correlation between patient satisfaction and improved outcome, Aharony and Strasser conceded a satisfied patient cooperates more with treatment. This is reiterated by Pascoe (1983, p. 189):

"Regarding outcome, a satisfied patient is seen as participating more carefully and accurately in his or her treatment and therefore achieving a better clinical outcome than a dissatisfied patient".

Similarly, Mahler and Kulik (1990) have found that active patient participation results in better health outcomes. Patient satisfaction has been demonstrated to be greater amongst older people (Pascoe, 1983; and Carmel, 1985). This supports the earlier assertion that expectation of outcomes vary with age. Pascoe also found that female patients expressed greater levels of satisfaction than male patients.

Ware (1993) describes the search for reliable measures of the way patients perceive their health. He discusses recent advances in standardised self-reporting measures of functioning and well-being and their use in monitoring outcome. One example given by Ware is the American short form (SF36) which has separate norms for gender, age and occupation. Fries and Spitz in Spilker (1990, p. 26) also define outcome in terms of a patient perspective:

"Patients desire to be alive as long as possible; to function normally; to be free of pain and other physical, psychological, or social symptoms; to be free of iatrogenic problems from the treatment regime; and to remain solvent. These five dimensions (death, disability, discomfort, drug side effects and dollar costs) define patient outcome".

Fries and Spitz maintain that these five dimensions, not necessarily in the same order, emerge consistently in multiple surveys of patients. Outcome is similarly defined by Brook and Appel (1973) as a patient's response to care with respect to mortality, symptoms, ability to both work and perform daily activities and physiological measures. They set up a study to look at both explicit (group consensus) and implicit (individual reviewers) judgements of quality of care. They found that

"it is ironic that the outcome indicators of vital importance to the patient ie his activity and symptom level had insignificant correlations with either the implicit or explicit process judgements".

Measures using both patient-reported and clinical outcome are described by Tarlov (*et al.*, 1989) in their paper on the Medical Outcome Study undertaken in the USA. This was an observational study designed to assist understanding of how the health care system affects outcomes. The two major aims were to relate variations in patients' outcomes to differences in systems and to design tools to monitor patient outcomes. The study used

patient-assessed outcomes, clinician-assessed outcomes and objective measures adjusted for casemix.

Mirin and Namerow (1991) recognised the difficulty of developing an instrument to measure psychiatric patient outcomes and also produced a tool to measure clinical symptomatology and patients' social, interpersonal and occupational adjustments. Wilkin (et al., 1993) provide a summary of non-clinical instruments measuring both need and outcome for primary health care. The instruments described include disease-specific outcome measures along with outcome measures of mental health, social support and patient satisfaction.

In the Australian context, Harvey (1991), in the National Health Strategy Paper No. 8, also discusses variations in surgical procedures between and within states of Australia. He also states that up to 10 percent of all hospital admissions are inappropriate, prompting a need for measurement of health outcomes in Australia and lamenting the dearth of Australian outcome studies compared to North America and Europe.

The importance of having Australia-specific measures of medical effectiveness, in terms of measured health outcomes, has emerged only relatively recently as emphasised by two studies. Firstly, the provision of approximately \$250 000 by the NSW Health Department for 17 demonstration projects under the NSW Health Outcomes Program; and secondly, by the launching of the programs targeting health outcomes by the NSW Minister for Health in June 1993. Details of the 17 projects are given in the NSW Public Health Bulletin (Vol. 4 No. 5, 1993) where the NSW Health

Department officer responsible for the outcome programs outlined other departmental initiatives related to health outcomes. These include the setting up of health outcome councils in each area and region of the NSW Health Department, and that targeted outcome measures would be included in the performance agreements for all Chief Executive Officers of NSW public hospitals.

In August, 1994 the NSW Health Department organised a two-day conference on health outcomes. Seven hundred delegates attended the conference, including Dr John Ware from Boston USA who was responsible for the piloting of the SF-36 form as part of the Medical Outcomes Study world-wide, as noted by (Tarlov *et al.*, 1989). Tarlov is the president of the Medical Outcome Trust which puts out a bi-monthly bulletin that provides a forum for developments in health outcome measures throughout the USA and Europe such as the Patient Outcomes Research Teams in the USA (Tarlov, 1996).

2.3 Health Outcomes for Cardiopulmonary Bypass

Cardiopulmonary bypass is defined and discussed in detail in the next chapter. This section addresses outcomes for cardiopulmonary bypass, and more specifically, the most common procedure using cardiopulmonary bypass - coronary artery bypass grafting (CABG). However, cardiopulmonary bypass is also used for heart and lung transplants and valve procedures. The results of a study examining general surgical complications as an outcome measure were reported by Spotnitz (*et al.*, 1995). This study identified, through multivariate techniques, the factors that were associated with general

surgical complications. The major predictor of poor outcomes or complications was the use of intra-aortic balloon pumps.

The major procedure using cardiopulmonary bypass - coronary artery bypass grafting (CABG) - is a surgical treatment for the relief of angina. This procedure of saphenous vein autografting of diseased arteries was introduced in 1967 by Favalaro (1968). Both Pinna Pintor (*et al.*, 1992) and Booth (*et al.*, 1991) suggested that success following this type of surgery has traditionally been measured in terms of reduced mortality, relief from symptoms, activity limitations, work status, levels of re-hospitalisation and medication use. A major study that used inpatient mortality for coronary artery bypass graft surgery as an outcome measure was reported by O'Connor (*et al.*, 1996). A further measure of outcome - left ventricular ejection fraction - was cited by Booth (*et al.*, 1991). Griffith (*et al.*, 1995) emphasised the need to adjust outcome for risk in a study of coronary artery bypass graft patients in Pennsylvania, USA. The analysis of hospital mortality data for these patients showed that the results were inappropriate for distribution to consumers and that consumers needed to be informed about the variation in risk for given pre-operative conditions. They concluded that surgeons need to have a greater understanding of models that evaluate outcome.

Sjö land (*et al.*, 1996) conducted a study of improvement in quality of life for 2365 coronary artery graft bypass patients who had been hospitalised in 15 Western Sweden hospitals. Quality of life was assessed before (n=1396) and two years after (n=1745) using the Nottingham Health Profile, the Psychological Well-Being Index and the Physical Activity score. Their study showed that the greatest improvement in quality of life after surgery occurred

for women, those with the most impaired exercise capacity pre-operatively and those with the most severe angina pectoris pre-operatively.

Traditional measures of outcome do not include a measure of psychological outcome. Notwithstanding this, Lindal (1990) stated that 54% of patients following CABG suffered from depression. A similar proportion, 61% of patients suffering neurological complications following this type of surgery, has been cited by Shaw (et al., 1985).

Newman (et al., 1990) studied 66 patients' neurological deficits 8 days, 8 weeks and 12 months post-coronary artery graft surgery. The results show that 29% of patients had neurological deficits 12 months post surgery. This proportion was approximately the same as the proportion with deficits at 8 weeks indicating that neurological deficits are relatively stable from 8 weeks post surgery. There was, however, a marked drop in deficits from 8 days to 8 weeks post surgery. The major explanations for these deficits were embolic and/or related to perfusion.

They found that post-operative neurological complaints were not related to complications in surgery. Pinna Pintor (et al., 1992, p. 78) stated that "psychological and behavioural factors, both as predictors and post-operative course and outcome measures have seldom been investigated". Keller (1991) also identified this lack of information both pre- and post-surgery. In the study by Pinna Pintor (et al., 1992) the incidence of neuropsychological dysfunctions after CABG was 24% a year after surgery. They found that negative psycho-social prognosis was not linked to the medical prognosis and therefore concluded post-operative outcomes depended more on pre-

operative psychological and behavioural factors than on physical improvement. These findings have been confirmed in three other studies (Willner, 1990; Stein *et al.*, 1990; Redeker, 1992). In another study by Caine (*et al.*, 1991), it was found that the interventions influencing outcomes after CABG were waiting time, rehabilitation information and the quality of information for the patient and their family.

In this country, the Australian Health Ministers Advisory Council (Australia, Department of Human Services and Health, 1994) have noted that there is a need for the collection of national data on outcomes following cardiovascular disease.

2.4 Health Information Systems

This subsection will cover health information system literature which has a somewhat loose link with patient outcomes. (A fuller discussion of the literature where health outcomes have been strongly linked with information systems can be found in the next subsection). In fact, most literature on health information systems, and their success, has focused on cost reduction, not on patient management for outcome (Classen *et al.*, 1991; Kerr and Jelinek, 1990; Martin, 1990; and Wrigley, 1990).

Health information systems can be discussed according to the following five categorisations made by Martin (1990): core systems (ie. hospital admissions, discharges and transfers), the master patient index, order entry and results reporting); business and financial management systems, communications and networks; medical support systems (ie. decision support); departmental

management systems (eg. laboratory, radiology, pharmacy, nursing and surgical); and; medical documentation systems (ie. the hard copy medical record and their evaluation such as quality assurance, utilisation review, infection control and discharge plans).

There is a clear need for these separate departmental, or organisational, information systems to be fully integrated with a focus on the customer - the patient. Wrigley (1990, p. 161) noted:

"it must be recognised that information systems should be built to support and monitor the basic product of the industry... the basic product is medical care not administrative functions (eg. billing, accounts receivable)".

Current literature on health information systems with a customer focus tends to be limited to medical decision support systems and medical documentation systems. Marsden and Pingry (1993, p. 183) describe decision support systems as a tool to "increase the efficiency and effectiveness of the search for structures to unstructured problems". One of the problems of measuring the success of decision support system implementation is that they are not mandatory systems but voluntary (Alavi and Joachimsthaler, 1992). Failure occurs in implementation when they are not used or underutilised.

In a study of outcome-oriented computer decision support systems (DSS), Lau (et al., 1991, p. 367) summed up the major problem in using artificial intelligence to support medical decision making:

"At present the deployment of expert level knowledge against realistic volumes of data remains a computational bottle-neck in AI applications".

In a similar vein, Chute (1992) acknowledges that the problems in maximising outcomes relate to the tedious nature of detailing conditions , interventions and outcomes. It is time-consuming, expensive and needs consistent methods. He states:

"The process of outcome analysis constitutes the determination of sequential correlations amongst all these elements in the hope of discovering optimal pathways that maximise outcome and minimise cost" (Chute, 1992, p. 135).

This is reinforced by Wong (et al., 1994) where they state that medical knowledge doubles every eight years but in spite of this diagnostic errors increase. They also state that few diagnostic decision support systems interface with a comprehensive database. One example they give which does do this is the ILIAD system which interfaces with the HELP information system discussed later in this chapter. In a review of decision support systems (DSSs), Johnston (et al., 1994) found strong evidence that they can improve physician performance but that additional well-designed studies are needed to assess the effect on patient outcomes.

It is important to define a medical decision. Gage (1993) states that a 'medical decision' is a random experiment that links patient measurements to treatments to quantify treatments. Each possible treatment must be labelled and probabilities assigned. These probabilities should add to one. Focussing initially on medical decision support systems, Sonneberg and Eckman (1991, p. S15) linked medical decision making and outcomes as follows:

"The Patient Outcome Assessment Research Program (POARP) and more recently the Medical Treatment Effectiveness Program (MEDTEP) of the Agency for Health Care Policy and Research have prominently featured decision analysis as a component of the overall effort to utilise health care resources more appropriately."

Two major methods used in medical decision support systems are probabilistic decision support and artificial intelligence. Nykänen (et al., 1991) state that these systems are designed to support, and, at times, take the place of human expertise. Blum (1986, p. 302) describes types of decision making tools:

"It is useful to separate the clinical decision making tools into those that are strongly formalised - ie, those that use algorithmic methods or statistical pattern classification techniques-and those that rely on symbolic reasoning approaches - ie artificial intelligence."

Artificial intelligence is "the subfield of computer science that is concerned with symbolic reasoning and problem solving" (Turban, 1989,p. 679). More specifically, expert systems, are computer systems that provide high level of performance which emulates the expertise of individuals with extensive education and training. Medical decision support systems, which employ artificial intelligence, are examples of expert systems. Sprague and Watson (1989) describe decision support systems as interactive systems that help decision makers utilise data and models to solve unstructured problems. It is emphasised by both Sprague and Watson (1989) and Turban (1988) that there is extensive user involvement in decision support systems. The architecture of expert systems is presented in Figure 2.1.

The figure below shows the two-way link between the knowledge base and the inference engine. The inference engine is linked in two-way exchanges with special interfaces, to the expert, to the explanation subsystem and to the user.

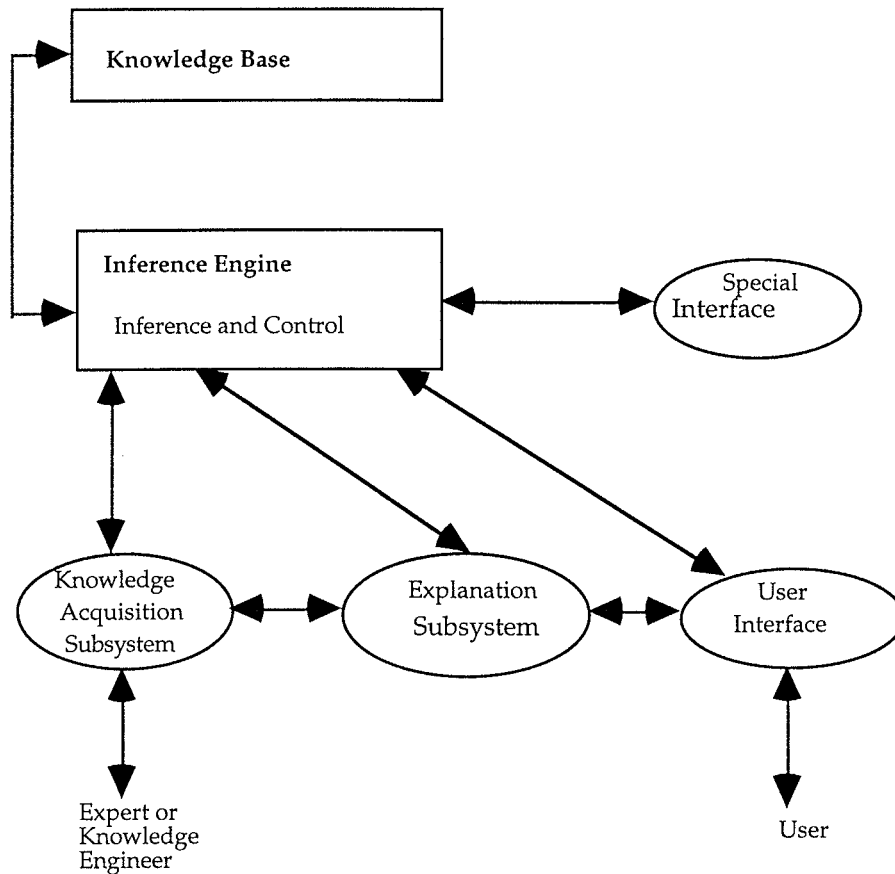


Figure 2.1: The Architecture of Expert Systems

(Source: Based on Harmon and Sawyer, 1990, Figure 2.6, p. 20)

There are two types of knowledge in expert systems - declarative and procedural. Declarative knowledge is data with some logical or empirical relationship between terms. Declarative knowledge is found in the knowledge base. Procedural knowledge is a program with some well-defined procedure or algorithm. Procedural knowledge is located in the inference engine. Harmon and Sawyer (1990) state that there are two advantages in separating knowledge and inference. First, the system is more flexible and responsive and can be readily updated. Secondly, a more complex program

can be produced if the expert's knowledge and their decision processes can be captured separately.

Owens and Sox (1990) describe an algorithmic method using Bayesian probability theory to assist in medical decision making. Bayes theorem is a quantitative method using pre-test probabilities which are adjusted to take account of new information gained from, in the case of medicine, diagnostic tests. The use, in medicine, of both subjective (or heuristic) and objective probability assessment is discussed by Owens and Sox. They also give an example of the Bayesian probabilities in the calculations of QALYs (see Section 2.1) for knee-replacement surgery.

The Hospital Evaluation through Logical Processing (HELP) system, developed at the University of Utah's teaching hospital - the Latter Day Saints (LDS) hospital - would be the best known use of Bayesian probabilities in medical decision making. The system was designed to meet not only clinical needs but also teaching and research needs. Although there are numerous examples in the literature of the HELP system being employed to assist in medical decision making (for example, Pryor, 1994), this review will be limited to a discussion of those specifically relating to patient outcome.

The use of the HELP system in developing a model to evaluate pregnancy is covered in Sager (et al., 1991). The patient charge was used as a crude measure of outcome. Again using the HELP system, Evans (et al., 1986, p. 1007) developed a computerised infectious disease monitor which automatically provides four types of patient 'alerts' relating to hospital-acquired infection and the appropriate use of antibiotics:

"(1) computerisation of microbiology test results combined with other computerised information sources, eg pharmacy, radiology; (2) a knowledge base created by infectious disease physicians; (3) 'automatic' medical decision making capabilities activated by patient test results; and (4) timely reporting of these computerised medical decisions to infectious disease personnel".

Surveillance personnel using these computerised methods found more hospital-acquired infections in around a third of the time than they would have achieved using traditional methods. The system also identified: 37 patients not receiving appropriate antibiotics; 31 patients who could receive cheaper antibiotics; and 142 patients receiving cephalosporin prophylaxis for excessive periods of time.

Another use of Bayes' rule is calculating expected utilities of medical and reproductive options (Pauker and Pauker, 1987). The process is described as enhancing decision making in clinical decision making. The use of another statistical algorithm to reduce a dataset of AIDS patients' classifications to formulate a single prognostic score is described by Stitt (*et al.*, 1991). The model correctly predicted 82 percent of all cases (deaths, survival) at discharge. De Neef (1987, p. 313) showed the importance of quality assurance by presenting the potential for a number of adverse patient outcomes should the provider fail to follow patient care protocols "by linking the process and outcome of patient care explicitly, quantitatively and graphically, decision

trees communicate the importance of following appropriate standards of patient care."

Other studies which have used algorithmic techniques in patient management are as follows: Kuperman (*et al.*, 1991) used a systems-analytic approach to total quality management; Kresel (*et al.*, 1987) showed that low cost drugs do not always provide cost-effective treatment using decision theory; Akers (1991) used algorithmic methods as a tool to improve patient management; Sigurgeirsson (1991) used confidence intervals based on the Poisson probability distribution to predict cancer development; Hatcher (*et al.*, 1986) evaluated the potential of the use of a decision model in triaging patients; Corder and Ellwein (1984) assessed treatments for symptomatic Hodgkin's lymphoma with decision analysis; and Bernelot Moens and van der Korst (1991) used Bayes' rule to diagnose rheumatic disorders.

Shortliffe (1991) claims that classical probability theory has been abandoned since the 1970s, despite an abundance of recent literature to the contrary. He proposed that this method has been discarded because of major limitations, which include the fact that classical probability theory does not employ problem-solving skills in the same way as they are used by humans. However, Lehmann and Shortliffe (1991) combine both artificial intelligence and statistical techniques in the development of a prototype system, THOMAS (named after Reverend Thomas Bayes), which integrates probabilistic, subjective, methodological and domain knowledge. They outline the use of THOMAS in assessing literature on randomised clinical trials as initially presenting the decision problem to the machine. Next, the physician tells the machine of methodological reservations with the clinical

research paper. The machine, with this information, constructs a knowledge base in the form of a statistical model. Following this, the machine updates the models parameters probabilistically using both data from the study and the physician's prior beliefs. Finally, the machine recommends a treatment based on maximised utility.

Figure 2.2 is a model of a decision tree using Bayesian probabilities based on Hagen (1992). He states that the development of the tree is based on characteristics of the patient and the disease. As an example, the patient might be a 59 year old woman with coronary artery disease, say triple vessel, and normal ejection fraction. This model is very simple, with only two possible strategies and four outcomes, but provides a framework for development of more sophisticated decision trees. Strategy 1 might be medical intervention, such as drug therapy, whilst Strategy 2 might be a surgical intervention. The outcomes might be death and years of survival.

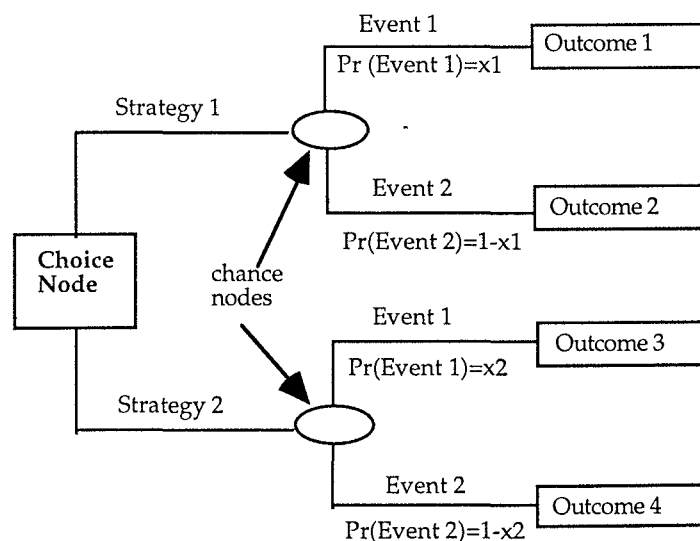


Figure 2.2: Simple Decision Tree Model

(Source: Hagen *et al.*, 1992, p. 349)

Another system using both expert knowledge and statistical reasoning are causal probability networks (CPNs). Andreassen (*et al.*, 1991) demonstrated the use of CPNs in a simple causal network based on two diseases - influenza and throat infection - with two symptoms - sore throat and fever. In support of the simplicity of this type of system, Andreassen (*et al.*, 1991, p. 1) state:

"The way in which knowledge is acquired and represented in CPNs makes it easy to express 'deep knowledge' for example in the form of physiological models and the facilities for learning make it possible to make a smooth transition from expert opinion to statistics based on empirical data."

A comprehensive review of artificial intelligence systems in medicine has been written by Perry (1990). The review includes relatively simple systems such as on-line medical textbooks that provide faster access and automatic updates which are not available with conventional textbooks. She also gives examples of rule-based systems, causal models and hypothesis-based systems and outlines some problems with knowledge-based systems. Two forms of AI - knowledge-based systems and expert systems - are distinguished. Knowledge-based systems require a large body of knowledge (what to know) whilst expert systems are built to rival human experts (what to do). However, Perry (1990, p. 274) defines knowledge-based systems as follows:

"In all cases, knowledge-based systems are considered to be distinct from programs based primarily on

mathematical models, statistical techniques or pattern matching, although some knowledge-based systems may incorporate statistical components along with knowledge bases in an effort to account for uncertainty".

Spiehler (1989) used expert systems to determine response and time of death in morphine-related deaths. An expert system is identified as a branch of AI which attempts to create computer programs which emulate decision-making and diagnosis by experts (Spiehler, 1989, p. 1104). Spiehler states that forensic scientists cannot simply rely on formulae and numeric calculations for their work but they also need interpretation by qualified and experienced human judges. He therefore states that expert systems are most apt for this type of science. The three different types of expert systems tested by Spiehler in the study were successful in determining time and response 70-90 percent of the time.

Safran (et al., 1996) gave an example of a study where the use of knowledge-based electronic patient records was tested to see if its use had any effect on clinicians' adherence to practice guidelines. The study looked at clinicians' response times for electronic alerts and reminders that had been triggered for HIV patients compared to the response time of control groups. The electronic alert in the intervention groups was a computer message that appeared each time the clinician logged on to his or her computer. An electronic reminder only appeared when the clinician looked at the patient's medical record. messages and alerts were not delivered electronically to the control group.

The response times to electronic alerts and messages were significantly less than those of the control group.

An example of computerised medical documentation systems used to improve patient care was given by Hendrickson (*et al.*, 1991). They describe the role of nurses as "care integrators" and care givers. As care integrators, nurses organise patient sessions with both physicians and allied health professionals such as physiotherapists, social workers and dietitians. They demonstrate how computerised systems can enhance care planning, discharge planning, patient monitoring and other patient care. Because cancer patients often present with a number of co-morbidities, the computerised system is described as being even more beneficial for this type of patient.

2.5 Health Information Systems Directly Linked to Health Outcomes

Recognising there should be strong links between the development of health information systems and outcomes, Shortliffe (1991) noted the importance of the outcome management process as illustrated in Figure 2.3. The figure shows that outcome is not just an output of the system but is fed back data to further refine the analysis and thus decision making. In support of the notion that outcomes should be routinely collected and fed back into the information system, Shortliffe states:

"The notion that we need to measure the decisions that clinicians make about management, to measure the clinical outcomes that are achieved, and to build databases that summarise experience... This point is

especially important because the organisations interested in this topic see outcome measurement and management not as a short-term research issue, but as the way care should ultimately routinely be provided" (Shortliffe, 1991,p. S10).

Shortliffe further suggests that providers could use computer-based records to automatically extract outcome data at the time of care.

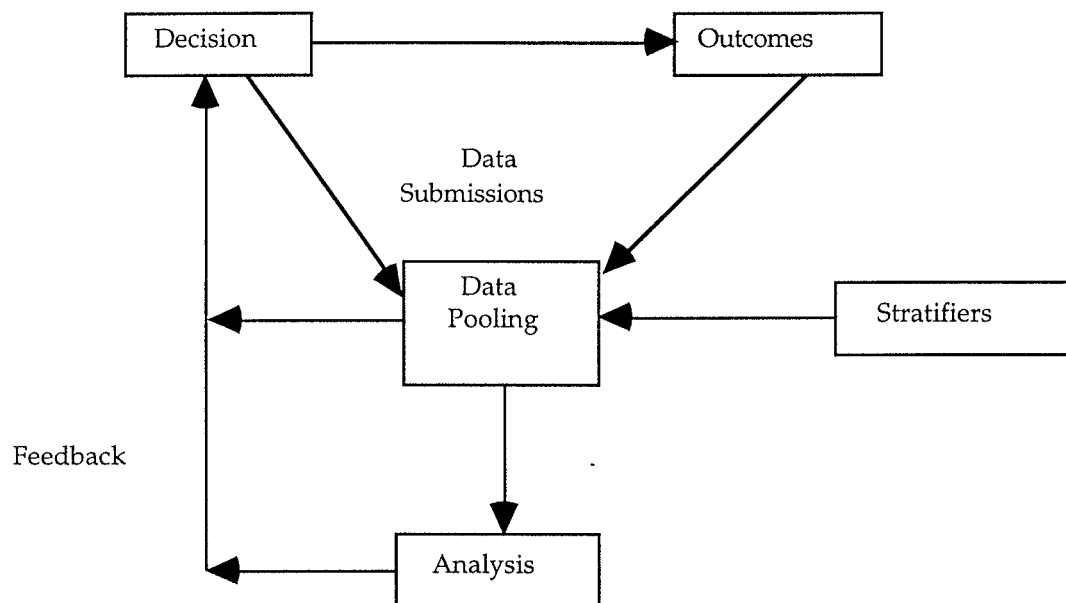


Figure 2.3: Outcome Management Process
(Source: based on Shortliffe, 1991, Figure 12, p. s10)

The importance of the link between information systems and outcomes was also emphasised by Shanon (1992, p. 62) in his description of expert system technology:

"The power of such a system is that it can scan an enormous amount of information quickly for a specified set of conditions pre-defined by physicians, nurses, pharmacists, lab technicians. . . the clinical professional, who has more time to evaluate pertinent data and spends less time scanning irrelevant data. Most important, a system that is designed from the ground up to utilise and support expert system technology can assist in improving quality by preventing errors before they occur, avoiding costly and unnecessary procedures and measuring the effectiveness of improvements to such procedure."

In linking outcomes to cost consideration Ellwood (1988, p. 1551) sets out the reasons that health care organisations are using information systems:

"In an effort to wed greater understanding of medical intervention to financial systems, health care organisations are installing information systems to record what the organisation is doing for the patient and how much does it cost".

Furthermore, Ellwood gives as an example the Harvard Community Health Plan System which integrates medical records with quality improvement techniques developed outside the health industry. At the time he considered

this to be the most advanced health quality measurement system in the United States.

Elfstrom (*et al.*, 1996) demonstrated with a database of 636 patients who had undergone intrainguinal bypass that the outcomes of patients who had been excluded from medical audit were worse than patients who were included in the audit. They concluded that all clinical databases should be complete. One way to do this, they suggested, was to automatically link the clinical database to the administrative database. It is further stated by Garnick (*et al.*, 1994) that administrative databases can only be useful in measuring quality if careful investigation is given to medical care utilisation, patient characteristics, provider characteristics and health plans. They make suggestions as to the inclusions to administrative databases so that they do measure quality of health care.

The need to provide better information systems to improve health outcomes in terms of quality of life has been recognised in a recent publication on health informatics in Australia (Hovenga, 1996). Hovenga recognises that the data requirements for the measurement of patients' quality of life is dependent on information besides that recorded in medical records. Therefore, she states there are major implications for health information systems in Australia in terms of measurement of quality of life post-intervention. This literature review of studies linking health information systems to health outcomes does not cite any published examples of Australian research in the area.

However, the author is aware of a number of researchers working in the area of medical decision support. A major contributor to this field is Associate Professor Paul Compton, Head of the School of Computer Science and Engineering at the University of New South Wales, who has developed a rule-based expert systems to assist in decision making for chemical pathology. Details on this system, the PEIRS system, are published in a paper by Edwards (*et al.*, 1993). Other researchers in this area, to the best of the author's knowledge, have not published. Notwithstanding this, the author, in a review of the literature for this research, has been unable to find published evidence of a direct link between patient health outcomes and information systems based on Australian case studies despite extensive searches through literature databases.

2.6 Damage Relating to Perfusion

Having established there is no comparable Australian research on the relationship between the process of care through health information systems and patient outcomes, research does provides empirical evidence of the use of health information systems in improving patient health outcomes by continuously monitoring outcomes, relating them to process and adjusting the process. This empirical evidence is available from the analysis of pre-existing data collected by a perfusionist at a Sydney teaching hospital. The perfusionist has, for the past eleven years, collected both clinical and patient-assessed outcome measures for patients undergoing open-heart surgery at this hospital. His research has concentrated on the process of coronary artery bypass surgery. Although these data have yet to be analysed statistically, the

perfusionist has been able to improve outcomes using continuous quality improvement techniques.

The perfusionist has concentrated on the neurological outcome following this type of surgery and produced observable improvements in outcome over time by relating the outcome to the perfusion process. Shaw (et al., 1985) stated that there were no long-term follow-up of patients for neurological complications for patients following this type of surgery. The author has been unable to find in the literature any published data on follow-up since this was asserted in 1985.

Neurological complications following open-heart surgery have had wide publicity in the media following the court case of the Australian entrepreneur, Mr Alan Bond who was unable to remember certain events following such surgery (The Sydney Morning Herald, May 24, 1994). The link between the perfusion process and neurological damage is discussed in the following statement by Shaw (et al., 1985, p. 1386) :

"The extracorporeal circulation process remains an imperfect replacement for normal functions of the heart and lung".

These complications have also been acknowledged in the scientific literature and are summarised in an editorial by Gilman (1990), who cites complications such as stroke, seizure, hypoxic encephalopathy, delirium, visual fields disorder, focal motor and sensory disorders. Also noted are psychiatric disorders such as insomnia, anorexia, depression, hostility, disorientation,

hallucinations and delusions. A number of studies have also linked brain injury to the perfusion process (Sarnquist, 1989; Govier and Reeves, 1989; Murkin, 1989; Shaw, 1989; Newman, 1989; and Venn, 1989).

2.7 Summary of Gaps in the Literature

The literature review has identified a number of gaps which this thesis aims to redress. Firstly, although the importance of patient input into the measurement of health outcome is recognised, the majority of the literature contends that there are no appropriate tools to carry out such a measurement. This thesis aims to provide evidence, through rigorously defining and testing outcome measures, that patient-assessed outcomes for coronary artery bypass surgery are valid tools.

Secondly, another gap, highlighted by Berwick (1996, p.877) in a recent editorial in JAMA, was the reliance on randomised clinical trials for evidence of medical effectiveness rather than acknowledging the "value of real-time science". By real-time science Berwick means time series data or data that are collected routinely over time and examined for patterns rather than setting up specific trials. The literature review has demonstrated that studies that follow-up patients' outcomes after a medical intervention have been conducted on an ad-hoc basis. However, these studies have not been part of routine data collection included as a component in their respective clinical information systems.

More specifically, there have been no publications of Australian research where patient outcomes are routinely collected as part of an information

system. Therefore another aim of this thesis is to provide evidence of an Australian clinical information system that includes routinely collected patient outcomes when they return to normal activities of daily living.

Furthermore, there has been no documentation in the public arena of information systems (that include patient-assessed outcomes) that have been successfully employed to improve outcome by relating the outcomes, on a continuous basis, to the process involved in the episode of care. This research aims, to demonstrate that patient-assessed health outcomes for surgical patients can be improved by the use of continuous quality improvement models where outcomes are related to the surgical process.

Any patient assessment measure of their health status, invariably, uses ordered categorical responses with more than two categories. These are described as ordered polytomous variables. However, outcome measures in the literature have used, mainly, binary outcome such as surviving or not surviving or infection or no infection. As well, documented analysis of outcomes and their association with other explanatory factors has been limited to univariate analysis. There is no published documentation of ordered polytomous health outcome measures used as a response variable and related to explanatory factors using multivariate regression techniques. The sparse literature on this multivariate regression technique is limited, in most cases, to theoretical rather than practical applications of the technique. Thus, data sets used in these analyses, with the exception of the study by Ashby (*et al.*, 1986), have been small in size with few explanatory variables. This research aims to produce, using appropriate statistical computer software, models in terms of explanatory variables for polytomous ordered

response variables (measuring patient-assessed outcome). This will provide a significant contribution to future studies of patient-assessed outcomes.

The database to be used for the development of the continuous quality improvement model and for the statistical analysis relates to coronary artery bypass graft patients. The literature review in this chapter is limited specifically to outcomes for coronary artery bypass graft surgery. The process variables to be controlled in the continuous quality improvement model are those that are collected whilst the patient is being perfused. In order to gain a better understanding of the principles of this process, the next chapter outlines the principles of perfusion.

CHAPTER 3

THE PERFUSION PROCESS

CHAPTER 3 THE PERFUSION PROCESS

3.1 Introduction

The purpose of this chapter is to provide an understanding of the perfusion process used for coronary artery bypass graft patients. The measures recorded during this process are central to the development of the continuous quality improvement model that will be discussed in Chapter 7.

This chapter, first, provides a definition for the term 'perfusion'. Secondly, a brief historical review of the developments of the procedure is discussed. In the third section, the architecture of the perfusion process specific to the study hospital is illustrated. The next section provides a more detailed description of oxygenators - equipment central to this process. Also the method for recording data during perfusion is discussed in this fourth section. The fifth section discusses acid-base management during perfusion because of the relevance of pH to other blood gas measures. A summary concludes this chapter.

3.2 Definition of Perfusion

Perfusion is 'the passage of fluid through a tissue, especially the passage of blood through the lung tissue to pick up oxygen from the air in the alveoli' (Oxford Reference Concise Medical Dictionary, 1992, p. 520). In this thesis, perfusion relates to the passage of blood through a membrane oxygenator when patients are on cardiopulmonary bypass (CPB). CPB is a technique where the circulation of blood by the heart and the functions of the lungs are taken over, for a short period of time - from between 5 minutes and 5 hours - by a device which is

external to the body (extracorporeal). The device oxygenates the blood and returns it to the arterial system of the patient.

3.3 Background to the Perfusion Procedure

CPB by extracorporeal circulation (ECC) allows emptying of blood from the heart so surgeons may carry out repairs without rushing. A milestone for this procedure was the successful closure of an atrial secundum defect in a five-year old girl by Dr F. J. Lewis in 1952 (Lillehei, 1993). This surgery was significant because it was conducted under direct vision. An important finding was that patients could be perfused at very low cardiac output (8 to 14ml/kg body weight/minute). These low flow levels allowed improved vision because the previous greater blood loss obscured the defects in the heart.

There were three major obstacles to successful open-heart surgery: finding an effective method for safely emptying the heart of blood for a reasonable amount of time (ECC); development of surgical methods for unfamiliar lesions; and providing post-operative care at a higher level of sophistication than had historically been available.

3.4 Description of the Process at the Study Hospital

The study hospital is a major teaching hospital in Sydney, Australia. A longitudinal database for coronary artery bypass graft surgery patients has been developed and maintained by the hospital perfusionist. Data from this database have been used by the perfusionist, employing graphical techniques, to identify

associations between variables that can be controlled by the perfusionist and health outcomes. these results, at the time of writing, have not been published.

This chapter describes, specifically, the perfusion process at the Australian teaching hospital where the data have been collected (Chapter 4) and analysed (Chapters 5 and 6). Firstly, a diagrammatic representation of the general perfusion process is provided in Figure 3.1.

PLATE XIII CONDUCT OF CARDIOPULMONARY BYPASS

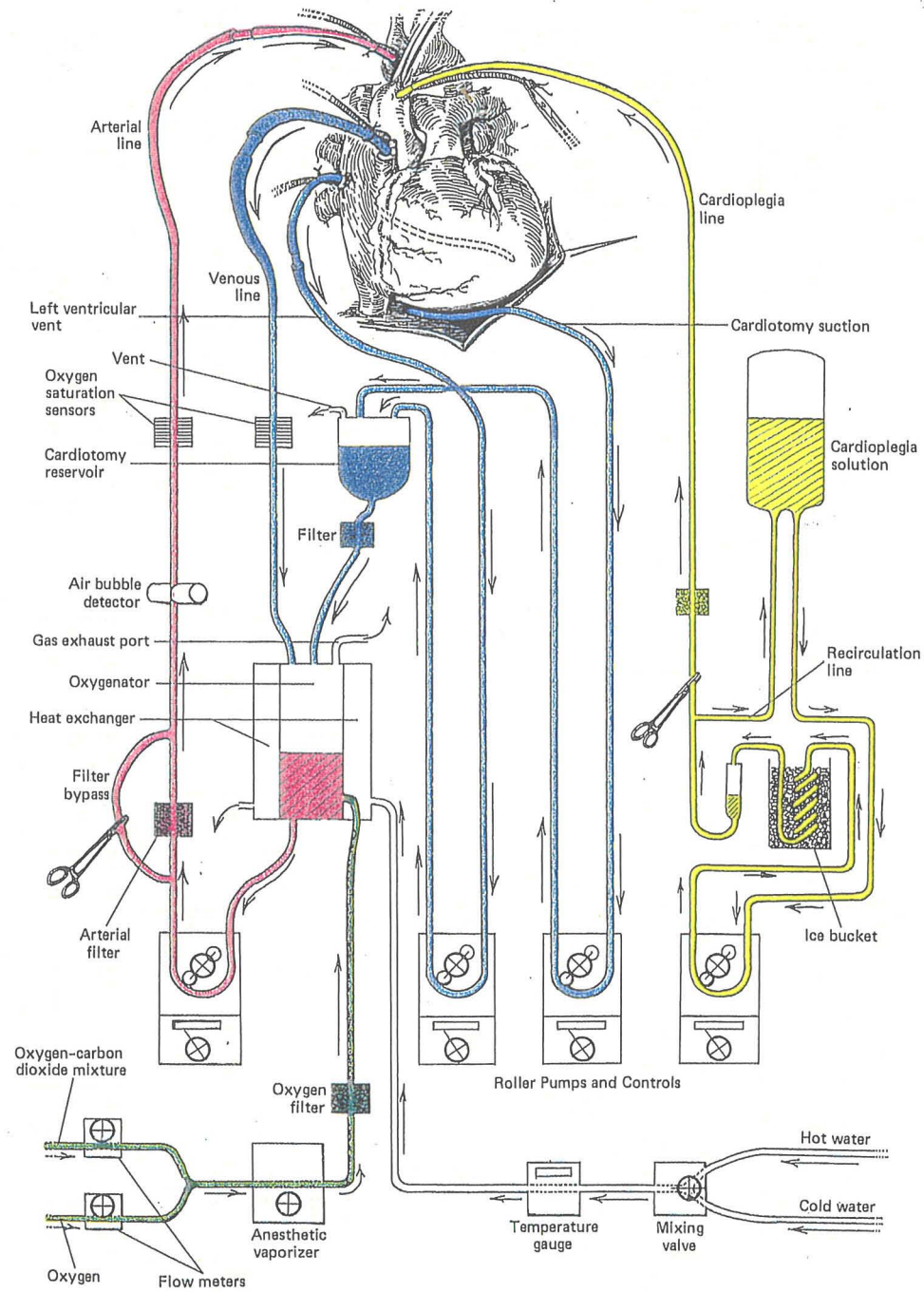


Figure 3.1 Cardiopulmonary Bypass

(Source: Based on Dillard and Miller, 1993, p.19)

Figure 3.2 Illustrates the specific system used at this hospital.

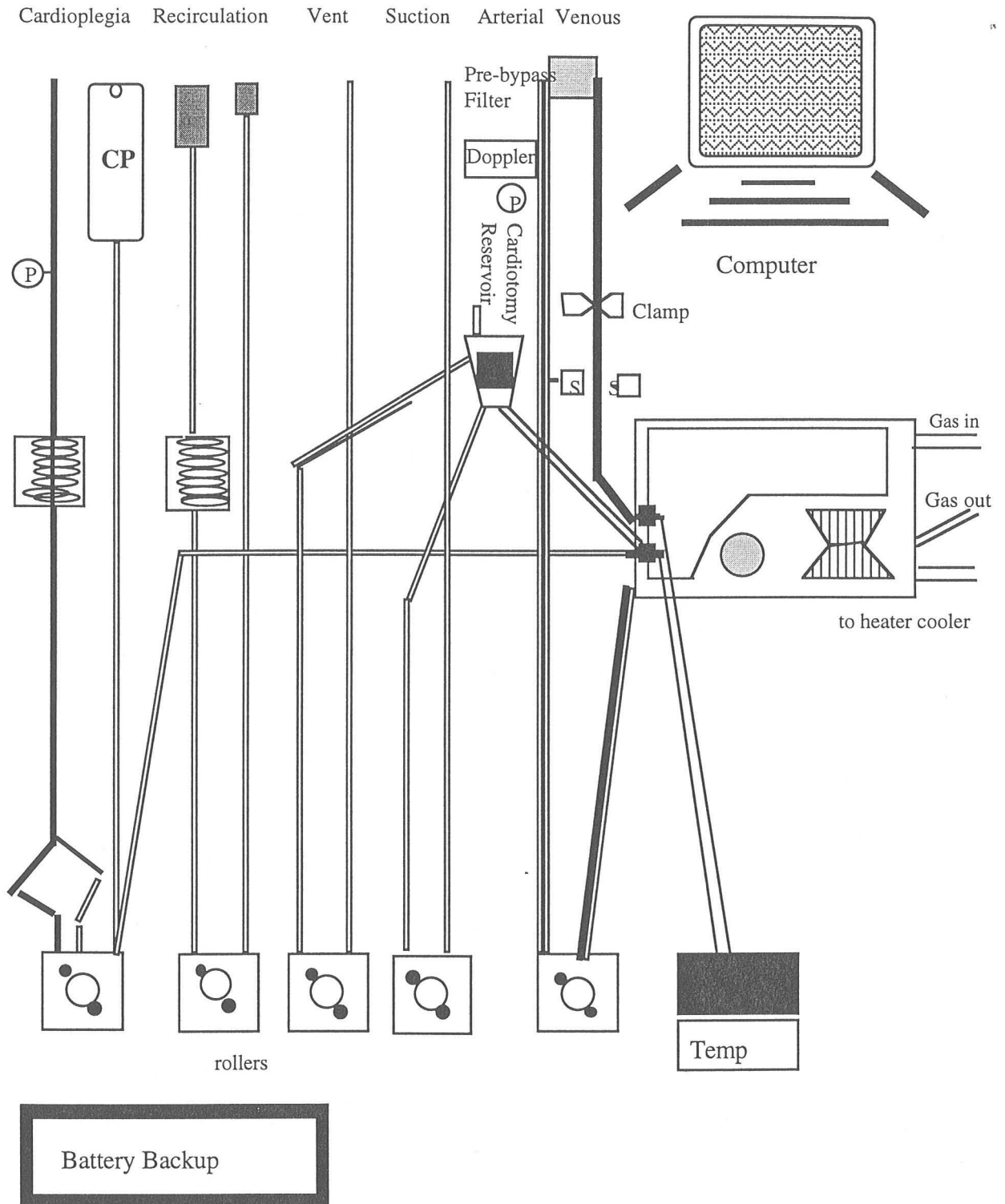


Figure 3.2: Heart-lung Machine Configuration at Study Hospital

Cross-clamping takes place to produce a flat line on the electrocardiograph (ECG) which lasts for 30 minutes and is repeated every 20 minutes. The recirculation coil provides topical cooling. The blood is sucked back by the pump after passing through the esky and the icy saline slush circulates around the heart for topical cooling and myocardial protection. Recirculation is not always used at the study hospital. Its use at the discretion of the surgeon. Figure 3.2A illustrates the recirculation of the process.

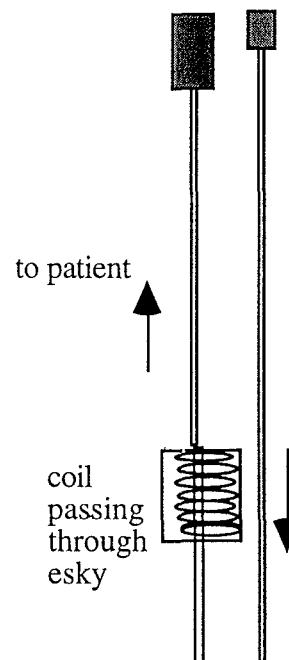


Figure 3.2A:

A high potassium solution of oxygenated blood cardioplegia mixes in a ratio 1:4 with blood through the cooling coil. This arrests and protects the heart by cooling to reduce metabolic demand.

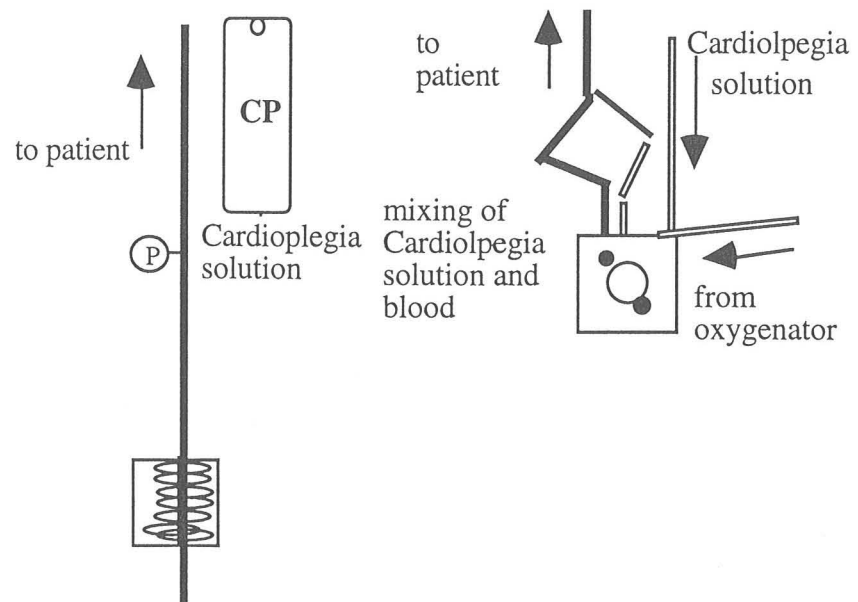


Figure 3.2B:

The vent-de-air decompresses the heart to provide a bloodless field. This is used in open heart surgery but not for coronary artery graft surgery.

The suction sucks lost blood from the chest cavity. Both the vent and the suction lines pass through the cardiectomy reservoir which initially handles excess flow and is filtered out. It then passes through the same circuit again and then passes through the oxygenator where it is reoxygenated.

A simple diagram of the set up of the cardiopulmonary bypass circuit at the study hospital is given in Figure 3.3.

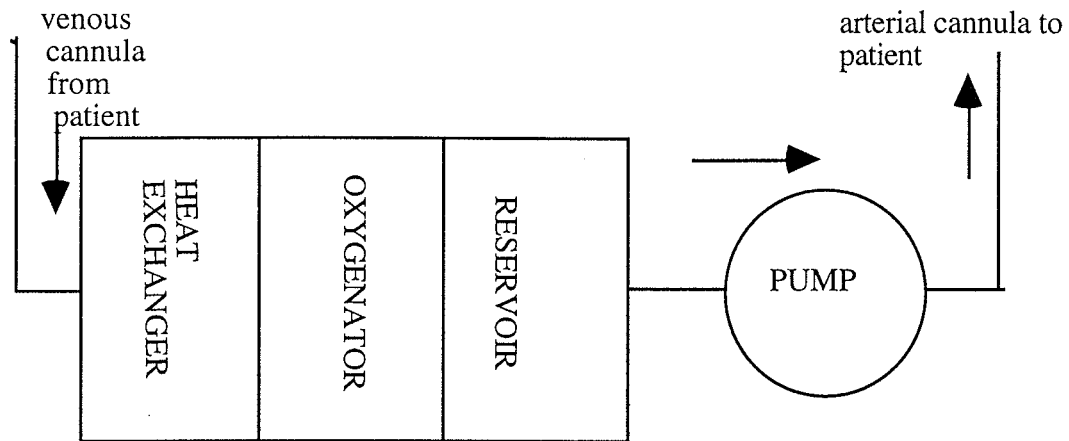


Figure 3.3: Cardiopulmonary Bypass Circuit at Study Hospital

3.5 The Oxygenator

The 'oxygenator' is the part of the perfusion equipment that takes over the patient's lungs during bypass. It is, in fact, a gas exchanger transferring O_2 , CO_2 , anaesthetics and other gases into and out of circulation. The oxygenators do not perform endocrine functions such as removal of toxic wastes and production of hormones. Gas exchange from gas to liquid and vice-versa is driven by diffusion according to partial pressure differences for the particular gas. The partial pressures of all gases dissolved in the blood adds up to 760mmHg at sea level. There are three methods that improve gas diffusion: "increasing drive gradient"; "increasing dwell time" and decreasing the diffusion path. Because increasing the drive gradient and dwell time have some limitations, the major way to improve gas diffusion is to decrease the diffusion path by putting membranes as close as possible together without causing problems with pressure drop across the oxygenator.

There are four obstacles that oxygenators have in emulating lungs. Red blood cells pass through pulmonary capillaries one at a time providing a short distance for O₂ diffusion. There is difference in partial pressures between the gas and blood phases. The smaller surface area in artificial lungs over which gas exchange occurs compared to the lung. Finally, artificial lungs cannot achieve the gas exchange of normal lungs. These problems are alleviated by reducing patient metabolic requirements through hypothermia, muscle paralyses and anaesthesia.

Oxygenators are of two types - bubble and membrane. Bubble oxygenators are divided into two sections. Small bubbles form when venous blood enters the first mixing chamber and fresh gas flows into the blood through a screen. Gas exchange occurs. The mix then enters the second chamber to be defoamed. High (et al., 1993) stated that there was a belief, prior to 1955, that bubble oxygenators should not be used because of problems with air embolisms. However, in 1955, De Wall and Lillihei used a disposable bubble oxygenator. Bypass took place at normal temperatures with flows of 25-30ml/kg of body weight with no apparent neurological or physical deficits. By 1976, 90 percent of all bypass operations involved bubble oxygenators (Bartlett and Harken, 1976). The trend, however, in the 1990's has been towards membrane oxygenators.

Membrane lungs attempt to achieve separation between blood and gas in a manner analogous to the natural lung (High et al., 1993, p. 38). Gas transfer occurs by diffusion of gas through membrane material. A membrane oxygenator is used at the study hospital. A cross-sectional photograph of the membrane oxygenator is provided in Figure 3.4.

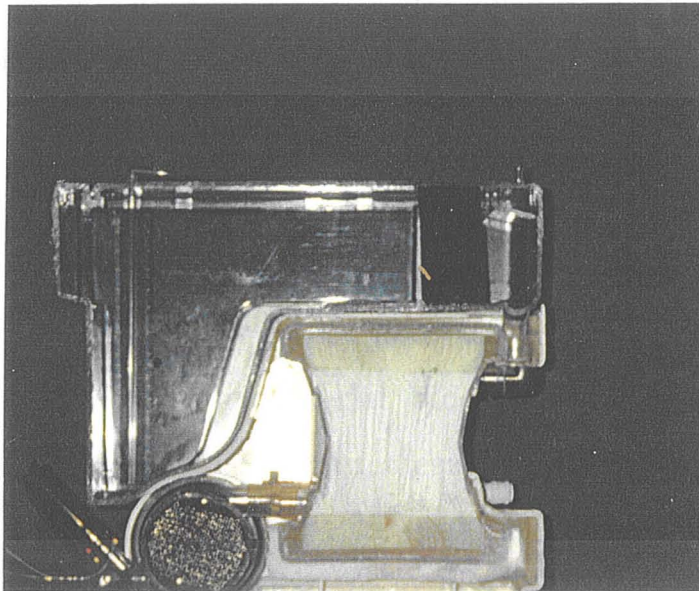


Figure 3.4: Cross-sectional Photographs of Oxygenator

Figure 3.5 provides a comparison of the actual set-up of equipment at this hospital with the diagrammatic representation in Figure 3.2.

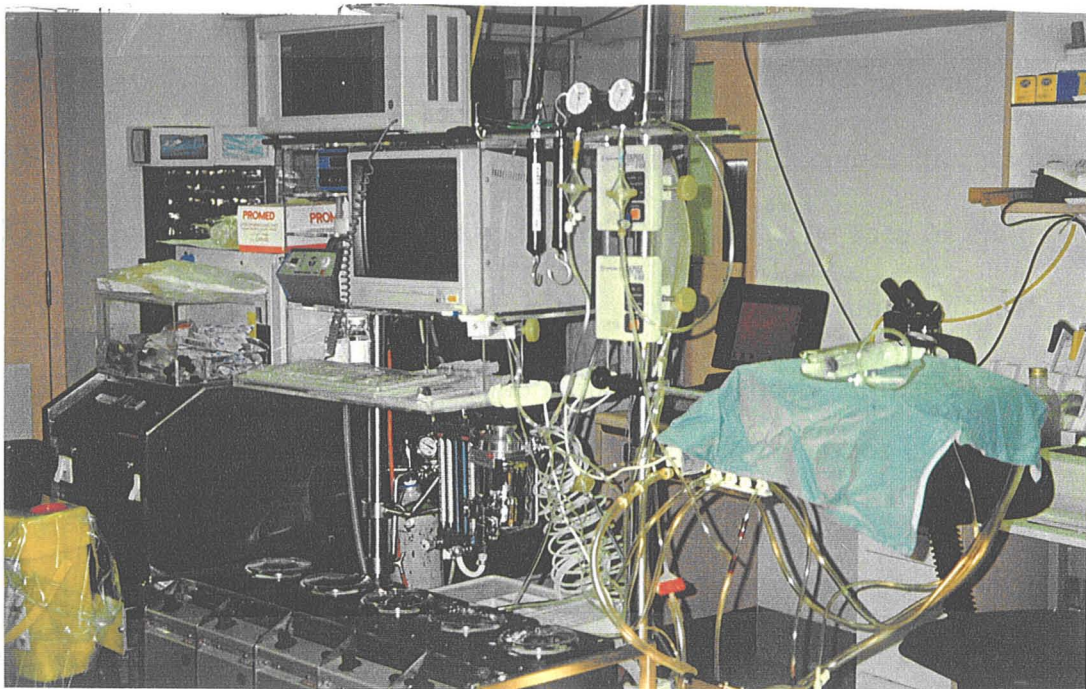


Figure 3.5: Photograph of Set-up of Equipment in the Study Hospital

Whilst the patient is on bypass, the perfusionist is prompted by a beep every five minutes to record various measures including arterial blood pressure, arterial line pressure, temperature, CO₂, O₂, blood flow, anaesthetic gas and central venous pressure on the computer. Also, samples of blood are taken pre-bypass, 10 minutes on bypass and then every 20 minutes (ie. 30 minutes, 50 minutes and so on). Measures of arterial blood gas, clotting time, electrolytes and haemoglobin are taken from these samples and entered on the patient's electronic record. These spontaneous measures of blood gases contrast to the comment made by Lillehei (1993) that, in the early years of bypass surgery (1950's), blood gases were not available clinically. Then, the perfusionist would have had no idea of physiological problems given that even emergency plasma electrolytes took four to five hours to process.

3.6 Acid-Base Management

Two acid-base management strategies are discussed in this section - alpha-stat and pH-stat. A pH of 7.4 and partial pressure of CO₂ of 40 is believed appropriate at 37°C. However, if a sample of blood is cooled CO₂ becomes more soluble and PCO₂ decreases to maintain a constant CO₂ content. The rate of change in pH given a unit change in temperature, or the gradient, is constant at $-\frac{dpH}{dT} = -0.015$ (ie $\frac{dpH}{dT} = -0.015$). This gradient in water is also constant and quite similar at -0.017 . If the CO₂ stores are to be kept constant during temperature changes then pH can be altering. This process is called alpha-stat. Therefore, if the alpha-stat method is employed in bypass surgery then for every drop in temperature of 1°C there is a commensurate increase in pH of 0.015mmHg.

By contrast, the term pH-stat refers to pH being maintained at 7.4 over varying temperatures. Therefore, during blood cooling, CO₂ must be added to maintain a PCO₂ of 40 and a pH of 7.4. The most commonly used method is alpha-stat.

3.7 Summary

A short summary concludes this chapter. Continuous quality improvement models for surgery are cyclical with the central part of the model being the surgical process. In the case of coronary artery bypass graft surgery, one of the central processes is perfusion. This chapter, therefore, has described the process in great detail covering the equipment used, the data collected in this process and the ways the process can be altered and controlled. The measures collected during this process will be used as explanatory variables for patient outcomes in Chapters 5 and 6. They also form the process variables to be controlled in the continuous quality improvement model that will be developed in Chapter 7. The next chapter details exactly how the perfusionist in this study follows these strategies and carries out acid-base management. It also describes the management of other controllable variables by looking at associations between patient's outcomes and these variables.

CHAPTER 4

METHOD

CHAPTER 4 METHOD

4.1 Introduction

The process of perfusion that has been linked in a continuous cycle of improvement in health outcomes for CABG patients has been discussed in Chapter 3, Section 3.4. This chapter describes how those data are used for quality improvement in the case study hospital.

The perfusionist has not published any results based on these data, nor described in conference forums his approach in using these data for establishing whether better health outcomes have been achieved.

It is important to note that the instrument in this study was designed by the perfusionist in 1983 and raw data were already collected. A number of variables in the raw data were in textual format with different texts, on occasions, being used for same category. Examples include the use of 'CAB', 'CABG' and 'graft' for type of surgery. The author, therefore, has made a number of transformations on and aggregation of categories in the data in preparation for analysis. Because the research findings are based on this instrument, section 4.2 is a retrospective critique of the development of that instrument.

This critical appraisal of the instrument draws on the theoretical work by Wilkin (*et al.*, 1993).¹ The purpose of this appraisal is to establish the suitability of the database for subsequent analyses.

Section 4.2 documents how the instrument to measure patient-assessed health outcomes at the study hospital was developed. The methods of data collection and data entry of both outcome and process variables are presented in the next section. Section 4.3 details the procedure followed by the perfusionist to analyse the outcome and process data. Section 4.6 provides the method of univariate and multivariate data analyses which are reported in detail in later chapters - Chapter 5 (for the univariate case) and Chapter 6 (for the multivariate case).

4.2 Development of an Instrument to Measure Patient-Assessed Outcome for CABG Patients

¹ The author has been involved in the design, implementation and reporting on more than 50 surveys of attitudinal and social surveys. The surveys include: a study of people with Multiple Sclerosis for Multiple Sclerosis Society of NSW; surveys of lead levels and Hib vaccine uptake in child care centres for the Eastern Sydney Area Public Health Unit; patient attitudes to road traffic noise and vehicular pollution for the Royal Prince Alfred Hospital, Sydney; a survey of alcohol consumption patterns in young sportsmen for the NSW Department of Health; an attitudinal survey of owners and executives to changes in corporate structure and future cooperation in commercial ventures of private Catholic hospitals for the Australian Catholic Health Care Association; and a study of community attitudes towards NSW Health for the NSW Department of Health. Also, the author has been involved in three research projects for the Commonwealth Department of Human Services and Health. One project is the Public Health Education and Workforce Study, the second project examines casemix knowledge and attitudes nationally and the third examines the impact of child care arrangements on childhood illnesses and workplace absenteeism. The design of the survey instruments and analyses are reported in the published literature (Black, 1994; Black and Cameron, 1993; Chapman, Smith and Black, 1988; Chen and Black, 1993; Cowie, Black and Ferson, 1995; Fisher, Black and Ferson, 1995; Furber and Black, 1995; Hines and Black, 1993; Lawson and Black, 1993a; b; Lawson and Black, 1995; Lawson, *et al.*, 1992; Rigby, O'Connor and Black, 1994; Rotem A. *et al.*, 1995; Degeling P. *et al.*, 1995).

The review of literature on health outcomes in Chapter 2 has shown how contentious the issue of the validity of patient-assessed health outcomes is amongst health professionals. Much of the debate about this measurement has been summarised by Testa and Nackely (1994). Perhaps the most important comment on the development of instruments that measure outcomes in terms of patient-assessed quality of life measures was made by a group of statisticians, Cox (*et al.*, 1992), in emphasising the importance of keeping the instrument simple. Clearly, the development of any instrument requires careful documentation. The description of the instrument development for CABG below is based on the six guidelines for methods of measurement proposed by Wilkin (*et al.*, 1993); (a) purpose of measure; (b) level of measure; (c) construction of measures and scaling; (d) variability of scores; (e) reliability; and (f) validity.

(a) Purpose of measure

Any measure requires a rigorous definition. The measure in this study is quality of life. Quality of life, in this context, is assessed by the patient both pre- and post- CABG surgery. It is measured in terms of the patient being free of symptoms, free of pain, free of neurological and physical deficits, as a result of the procedures. Quality of life means the patient is capable of working and/or carrying out activities of daily living with respect to their expectations prior to surgery. Quality of life also means that patients suffer no financial burden which can be attributed to the procedure.

In general, there are three purposes of measures of health outcome in terms of quality of life: discrimination; evaluation; and prediction. All three are equally applicable to the purpose of the measure for CABG patients.

The major purpose of the patient-assessed health outcome in this study is to provide a measure of the size of longitudinal change in individual patients' assessment of their health status that can be attributed to the perfusion process. Therefore it is an evaluation tool. Its purpose is also to discriminate between different levels of outcome and to predict outcomes given certain measures of intra-operative variables.

(b) Levels of measure

The pre-operative interview is unstructured in consideration of the circumstances of the patient. The interviews are sympathetically administered providing reassurance, and are, therefore, adapted to individual patient's requirements. In general, patients are informed of their likely post-operative experiences and of the quality enhancement program. They are told that they will be followed-up one year post-surgery as part of this program. Patients are asked about their current work status and activities of daily living and are asked what their expectations of these activities will be one year post- surgery.

The one year follow-up interview is a semi-structured interview with a mixture of nominal and ordinal scale measures. The response categories used in determining both neurological and physical outcome are unprompted but follow structured guidelines. In the follow-up interview, categories of neurological damage are recorded based on the patients response to "Any noticeable changes in memory, ability to concentrate, any mood changes, depression or irritability?" Further responses are recorded based on their response to "Any changes to your social life, do you get on well with

people like before the operation?" The 35 categories for neurological responses are listed in Table 4.1.

Table 4.1: Categories for Responses to Neurological Damage

Code	Category	Code	Category
11	memory	34	personality change
12	concentration	35	aggressive
13	can't remember names	36	restless/impatient
14	slow thinking	41	dreams
15	vague	42	nightmares
16	loses track of conversation	43	disturbed sleep
17	trouble reading	44	can't sleep
18	not as sharp	51	emotional
19	comprehension down	52	moody/uptight
21	confused	53	depressed
22	disorientated	54	anxiety disorder
23	psychosis	55	out-of-body experience
25	paranoid	61	fear of crowds
26	impotent/poor no sex	62	fear of traffic
31	bad nerves	63	noise intolerance
32	irritable/temperamental	64	claustrophobia
33	intolerant	65	low confidence
		66	other phobia

Similarly, the categories of physical damage are recorded based on the patients response to 'no more angina type chest pains ever on exercise?', 'do you need to take any treatment for angina? 'No heart attacks other frights during the year?' Further, they are asked 'Any limitations on what you can and want to do?' 'Any abnormal breathlessness on exercise?' 'Can you walk as much as you wish?' 'Have you changed any sports, such as golf?' 'Can you do your hobbies as you could before, such as reading, watching TV, going out?' The 39 categories for physical responses are listed in Table 4.2.

Table 4.2: Categories for Responses to Physical Outcome

Code	Category	Code	Category
11	pain free	33	sore legs
12	improved (ie pain)	34	oedema
13	not improved	35	infection
14	worse after operation	37	sore arms
15	sorry they had operation	38	impotent
16	fit & well/happy	39	PVD
17	never been better	41	tires easily
18	not well	43	cannot walk
19	angina returned	44	fitter than before
21	breathless	45	low activity/potters around
22	pulmonary oedema	46	other illness
23	AMI after operation	47	dizzy/balance problems
24	bronchitis	48	arrythmia
25	weight problem	49	no strength/weak
26	poor eyesight	50	hernia
27	can't identify problem	51	keloid
28	heart failure	52	ulcers of legs
29	stroke post-operation	53	ulcer
31	sore chest	54	scar
32	varicose veins		

For each of the neurological categories mentioned, respondents were asked to score the symptom on a scale from '0' to '5'. Up to 3 categories of response were allowed. On this scale the following scores are defined: '0' = absence or denial of symptom even if prompted; '1' = a just noticeable or occasional symptom; '2' = a daily noticeable, consistent but still tolerable complaint; '3' = severe enough to be of daily annoyance and interfering with functional living; '4' = markedly interfering with enjoyment of life; '5' = disabling. An identical scale is used for physical defects.

(c) Construction of Measures and Scaling

The measure for neurological outcome is defined as 'NSUM' and is calculated by taking the symptoms mentioned by the patient (up to 3 are taken) and adding the scores for each of the symptoms. For example, if a patient mentioned they were 'confused' and score a '2' for this symptom they may have also mentioned they 'can't remember names' and scored '1'. In this case their NSUM would be $2 + 1 = 3$. NSUM could take values over the range '0' to '15'.

PSUM is the measure of physical outcome and is calculated in exactly the same way as NSUM. That is, the scores for up to 3 physical symptoms are aggregated. PSUM also can take values over the range '0' to '15'.

(d) Variability of Scores

Willner (et al., 1993) state that a good instrument will be 'normally' distributed. However, they acknowledge scores that concentrate on dysfunction (such as NSUM and PSUM) will have a number of respondents scoring '0' - that is no dysfunction at all. As the instrument was designed to measure dysfunction, it will not be normally distributed. Methods used for analysing data that are not distributed normally are discussed in 4.6 of this chapter.

(e) Reliability

A reliable instrument is one that minimises random error. Reliability of a particular measure is the degree to which the same results are given in

repeated implementations of the measure. Reliability is measured in three ways: 'test-retest reliability'; 'inter-rated reliability'; and 'internal consistency'.

Test-retest reliability is calculated by applying the measure to the same population at different points in time under the same circumstances. As the measures of outcome in this instrument have been specifically designed to measure change over time, there are real difficulties in assessing the same population at a different point in time. By using the test-retest method of reliability assessment in this study, only the responsiveness of the instrument would be measured and not the random error.

Inter-rated reliability refers to consistency between users of the instruments. This is tested by different interviewers administering the instrument to the same population. This again is not practical because of the temporal nature of measure, that is if the same person is interviewed by different people at different times this may affect the consistency of their responses.

Internal consistency is where measures do not depend on repeated implementation of the instrument. They assess the extent to which individual items are correlated with each other. Multi dimensional measures such as PSUM and NSUM are not expected to have a high internal consistency. However, internal consistency was calculated for the individual measures that make up the measures of NSUM and PSUM. There was some internal consistency measured between the rating of the neurological symptoms. The Pearson's Correlation Coefficient¹ between the first

¹ Pearson's Correlation Coefficient for variables x and y = $\frac{\text{cov}(x,y)}{\sqrt{S(x-\bar{x})^2 S(y-\bar{y})^2}}$

neurological symptoms scaled deficit and the second neurological symptoms scaled deficit was found to be 0.67. The Pearson's Correlation Coefficient was 0.59 between the second neurological symptom scale and the third neurological symptom .

All physical symptoms scores had correlations less than 0.3 with each other and also correlations less than 0.3 with scores for neurological symptoms. Thus there was some internal consistency for neurological symptom scores but little correlation for physical scores.

(f) Validity

A measure of validity is a measure of non-random or systematic error. There are three types of validity: 'content', 'criterion', and 'construct validity'. *Content Validity* relates to whether the components of the measure developed are appropriate for the item they are trying to measure. The demonstration of content validity for measures of health status is recognised as being difficult (Wilkin *et al.*, 1993). One method used for demonstration of content validity is to show that the items used to make up the measure are commonly mentioned as measuring the desired effect by qualified judges. The literature review reveals firm agreement that patients following a medical intervention wish to be free of pain, free of physical, psychological and social symptoms and wish to have no financial difficulties relating to the procedure (Patrick and Erickson, 1993; Fries and Spitz, 1990; and Brooks and Appel, 1973). Such measures have all been included in the study instrument to ensure content validity.

Criterion validity is when a measure is validated against the 'gold standard' for such measures. In this case, where patients' quality of life is being measured, it is impossible to have a 'gold standard', such as a functional test, with which to compare the developed measure. As the only way quality of life can be measured is by asking the patient and their family, there is no such standard. Therefore, criterion validity is not relevant in the formulation of this instrument.

The term *construct validity* questions whether the results obtained give rise to the expected hypotheses derived from the theoretical constructs on which the measure has been based. This is used when a 'gold standard' is not available. This involves, in the case of this survey instrument, comparing the neurological deficit score with established psychometric tests and the physical deficit against activities of daily living. The difficulty with such comparisons is that a person may score well on a psychometric test but still perceive that they have some neurological deficit. Similarly, in terms of physical functionality tests, a person may score poorly but within their expectations perceive no problems with physical functioning. This distinction between a quantitative and qualitative response is well known to psychologists. Discussion of this distinction, and the difficulty of measuring perceived utilities, can be found in Coombs (et al., 1970) and Luce (et al., 1968).

In summary, in critically appraising the development of this instrument in terms of the guidelines of Wilkin (et al., 1993), the purpose of the measure has been clearly defined. The levels of measure and construction of measures and scaling have been clearly defined. In terms of the variability of the scores, they are not expected to be normally distributed because of the

acknowledged problem of measuring dysfunction. Reliability of the instrument has only been partially addressed because of the difficulty of testing and re-testing the same population with such a survey. Similarly, validity has only been partially addressed because of the noted problem of providing a 'gold standard' for quality of life measures assessed by the patient. Overall, given the known difficulties of measuring quality of life, the survey instrument satisfies most of these criteria.

4.3 Pilot of the Instrument

The instrument so designed was presented to the study hospital's other perfusionist for comments on the instrument in 1983. The scoring systems for the outcome measures were reviewed by staff at the National Health and Medical Research Councils' (NH&MRC) Clinical Trials Centre at Sydney University. The instrument was piloted on a group of patients in 1983 and then further modified. This review of the instrument took place after the modifications.

4.4 Administration of the Instrument and Data Entry

The pre-operative interview is a personal interview conducted by the perfusionist with the inpatient at the bedside. These data are recorded on cards. The one-year follow-up interviews were telephone interviews. At least three attempts were been made at different times to contact each CABG patient. All effort was given to tracing the patient if they had moved, and only when absolutely necessary a spouse, or a close relative may have been interviewed. There has been only one interviewer for all interviews. The

interviewer was blind to both the details on the patient's medical record and those details recorded in the pre-operative interview. All details are recorded on cards with the patient's identification number and date of surgery.

There are four data sources that make up the perfusionist database: the pre-operative interview; the post-operative interview; the patient's medical record; and the data recorded during the perfusion process.

The pre-operative interview and the post-operative interview details are recorded on cards. The data on these cards are all entered manually to the perfusionist's personal computer (PC) Network separate ACCESS III files. Prior to the introduction of ACCESS software in 1991, EXCEL was used for databases.

The relevant data from the patient's medical records are also entered manually on ACCESS III files on the perfusionist's PC network. The data from the pre- and post-operative interviews and the medical records are entered in batches every 6 months.

The patient's details recorded during perfusion are entered directly to the one of the perfusionist's PC at the time of surgery. The PC forms part of the setup of equipment with the heart-lung machine and has already been described in Chapter 3 (Figure 3.1). These data are again entered in ACCESS III databases.

4.5 Perfusionist Data Analyses

The purpose of perfusion analysis is to see whether there was a relationship between the patient-assessed outcomes recorded one year after surgery and the peri-operative variables. That is, the aim is to see whether peri-operative variables were predictors of outcome. The variables chosen for analyses by the perfusionist are as follows, as detailed in Chapter 3: date of the operation; patients' age; arterial blood pressure whilst on cardiopulmonary bypass; length of time on bypass; and intra-operative blood gas analyses including haemoglobin levels.

Arterial blood gas readings, whilst on cardiopulmonary bypass, were entered into the heart lung machine computer at five-minute intervals and arterial blood gas and haemoglobin levels were recorded pre-bypass, 10 minutes on bypass and then every 20 minutes that the patient remained on bypass.

The NSUM and PSUM scores were related to the peri-operative variables by grouping the NSUM and PSUM scores for set levels or ranges of peri-operative variables and taking the average of the NSUM or PSUM score for that subgroup. As an example, all patients whose pH was >7.32 and ≤ 7.34 after 10 minutes bypass would be grouped separately and the mean PSUM and NSUM for that group would be put on a chart. This was repeated for all times on bypass and for all ranges of pH values. The peri-operative variables were plotted on the x-axis, the different times on bypass were plotted as the y-axis and the mean NSUM or PSUM were plotted on the z-axis. As both NSUM and PSUM measured outcome deficits and were plotted on the z-axis, the

optimal outcome values would occur when valleys occur in the three-dimensional plot. Returning to the example of pH, for a given time on bypass, the range of pH values that minimise NSUM or PSUM would be the proposed level that pH should be maintained for subsequent patients. This should be repeated for all points on the three-dimensional chart. Further discussion of these analyses for a number of specific peri-operative variables is given in Chapter 5.

4.6 Data Transfer and Analyses

The data were transferred from one of the perfusionist's PC using Mac Link Plus/PC software as MAC Excel 4 files to an external hard drive to allow analyses for this research project. Data analyses for this thesis was carried out in two phases: initial univariate statistical analyses; and more complex multivariate analyses.

The next chapter presents the results of the univariate analyses. Analyses were carried out using SPSSX for Mac and graphics were produced using Cricket graph. The analyses include descriptive results using frequency tables and graphics. Inferential results were limited to chi square and t-tests testing for relationships between the outcome variables NSUM and PSUM and from the following peri-operative process and explanatory variables - body mass index, cross clamp time, time on bypass, age, sex, marital status, retirement status, angina, previous operations, symptoms prior to surgery, year of surgery, socio-economic status, pH, partial oxygen, partial carbon dioxide, haemoglobin and base excess.

Explanatory variables include patients' socio-economic variables, pre-operative variables and relevant patient data from the medical record. These analyses treated PSUM and NSUM as dichotomous variables. The subgroups for NSUM were patients with NSUM equal to 0 and patients with NSUM greater than 0. PSUM was divided into subgroups in the same way.

Although these results provide details on groups where there is no deficit, and where there is some deficit further multivariate analyses treating NSUM and PSUM as dependent logistic variables was not considered as this type of disaggregation of patients would consider a patient who scored a NSUM of '1' equivalent to a patient with an NSUM score of '15' as having the same neurological or physical deficit. Therefore, the results of analyses where NSUM and PSUM are treated as multinomial response variables are presented in Chapter 6. The categories selected for disaggregation for NSUM and PSUM were chosen so that they were coherent in terms of deficit whilst ensuring there were sufficient response levels in the category to allow analyses. The categories for NSUM and PSUM were the same, that is '0', '1-2', '3-4', '5-6' and 7 or more.

Multinomial analyses were carried out using the SAS procedure 'LOGIST' on the University of New South Wales Vax mainframe computer.

4.7 Summary

This chapter has, initially, described the methods used by the perfusionist in developing the instruments for the pre- and post-operative phases of data collection. Details of the peri-operative data collection have already been

described in the previous chapter. Secondly, the appropriateness of the measures in the instrument have been discussed in terms of purpose, level of measure, construction, scaling, variability of scores, reliability and validity. The perfusionist's collection, entry and analysis methods have then been discussed. Finally, data transfer and the method for data analyses proposed in this study in terms of both univariate and multivariate procedures has been addressed.

Chapter 5 examines statistical associations between both patient characteristics and data collected during perfusion with health outcomes. Outcomes in the next chapter are assessed in terms of the presence or absence of neurological and/or physical deficits.

CHAPTER 5

UNIVARIATE ANALYSES OF CORONARY ARTERY BYPASS GRAFT PATIENT DATA

CHAPTER 5 UNIVARIATE ANALYSES OF CORONARY ARTERY BYPASS GRAFT PATIENT DATA

5.1 Introduction

This chapter initially describes the characteristics of the CABG patients in this study in terms of their demographic characteristics, symptoms and comorbidities. Section 5.3 presents the distribution of NSUM (as defined in Chapter 4) scores and an analyses of demographic factors, symptoms and comorbidities that might have associations with NSUM. This is followed, in Section 5.4, by an analysis of the effect of various blood gas measures recorded during bypass on NSUM. Section 5.5 gives the distribution of PSUM scores and an analyses of demographic factors, symptoms and comorbidities that might have associations with PSUM. The final section summarises and compares results for PSUM and NSUM.

5.2 Description of Respondents

The study hospital is a large teaching hospital in Sydney, Australia. From July 1993 to June 1994, there were 4428 discharges for coronary artery bypass graft procedures (AN-DRGs 287 to 291) in New South Wales. In the same period, there were 759 discharges for the same procedures at the study hospital¹. Thus the study hospital accounts for 17% of the discharges in the state for these procedures.

¹ Number of discharges, casemix and length of stay characteristics for all AN-DRGs (Version 3) New South Wales acute public hospitals 1993/94 untrimmed - all discharges included (Professor George Palmer, Director, the Centre for Hospital Management and Information Systems Research, UNSW, pers. comm.).

In the period 1984 to 1992, there were 6466 patients who had CABG surgery at the study hospital. Follow-up interviews were successfully completed with 3979 respondents one year later - a response rate of almost 62%. Of those patients who were not followed-up, 1838 patients (28.5%) could not be located; 194 (3%) had moved and had left no forwarding address; 78 (1.2%) patients had the wrong phone number recorded on their card; 180 (2.8%) had no telephone; 40 (0.6%) had either poor or no English; 112 (1.7%) had died in hospital; and 45 (0.7%) were overseas.

Comparing the 112 patients who died in hospital with the 3979 patients who were successfully followed-up, it was found that those who died in hospital had significantly greater mean age, mean time on bypass and mean cross clamp time. The mean age for those who died was 64.2 years compared to 61.2 years for the follow-up group ($t=3.52$, $df=4047$, $p<.01$). The mean cross clamp time for those who died in hospital was 59.8 minutes compared to 48.0 minutes for the study respondents ($t=5.98$, $df=4089$, $p<.01$). Mean bypass time for those who died was 130.1 minutes compared to 88.2 minutes for the follow-up sample ($t=12.79$, $df=4089$, $p<.01$). There was no statistically significant difference by gender.

Around 80 percent of the patients who were followed up were male. The mean age was 61.2 years and standard deviation of 8.8 years. Table 5.1 presents, for comparison, a summary of mean age, sex distribution and mortality rate in hospital for other follow-up studies of CABG patients. The demographics from this research show similar results for gender distribution, mean age and mortality rate to a number of other studies mentioned in the

table. It may be noted that the sample size in this research is the second largest from an international review of literature on follow-up studies of CABG patients. Only the study by Sergeant (*et al.*, 1991) has followed up more than the 3979 respondents analysed in this study.

Table 5.1: Comparison of Selected Variables from CABG Follow-Up Studies

Study	Size	% Male	Mean Age(s.d.)	Inpatient Mortality(%)
Black, (1995)	3979	80%	61.2 (8.8)	112 (3%)
Blumenthal (<i>et al.</i> , 1991)	20	80%	58.9 (*)	* (*)
Carrier (<i>et al.</i> , 1993)	206	73%	59.0 (7.6)	9 (4%)
Louagie (<i>et al.</i> , 1995)	474	68%	65.0 (*)	32 (7%)
McGee (<i>et al.</i> , 1993)	112	84%	55.2 (*)	1 (1%)
Newman (<i>et al.</i> , 1989)	62	92%	55.1 (*)	* (*)
Pinna Pintor (<i>et al.</i> , 1992)	626	86%	61.0 (8)	* (*)
Redeker, (1992)	129	79%	62.0 (9.7)	1 (1%)
Sergeant (<i>et al.</i> , 1991)	5880	86%	55.8 (8)	78 (1%)
Shaw (<i>et al.</i> , 1985)	312	89%	53.4 (7.4)	4 (1%)

* indicates not stated

Figure 5.1 shows that the largest cohort of patients (900) in this study was in the 65 to 69 years range that is post-retirement age. More than 6 in 10 were aged 60 years or more. A small number (20) of patients aged 80 years or more were considered suitable for surgery. There were 48 patients aged less than 40 years.

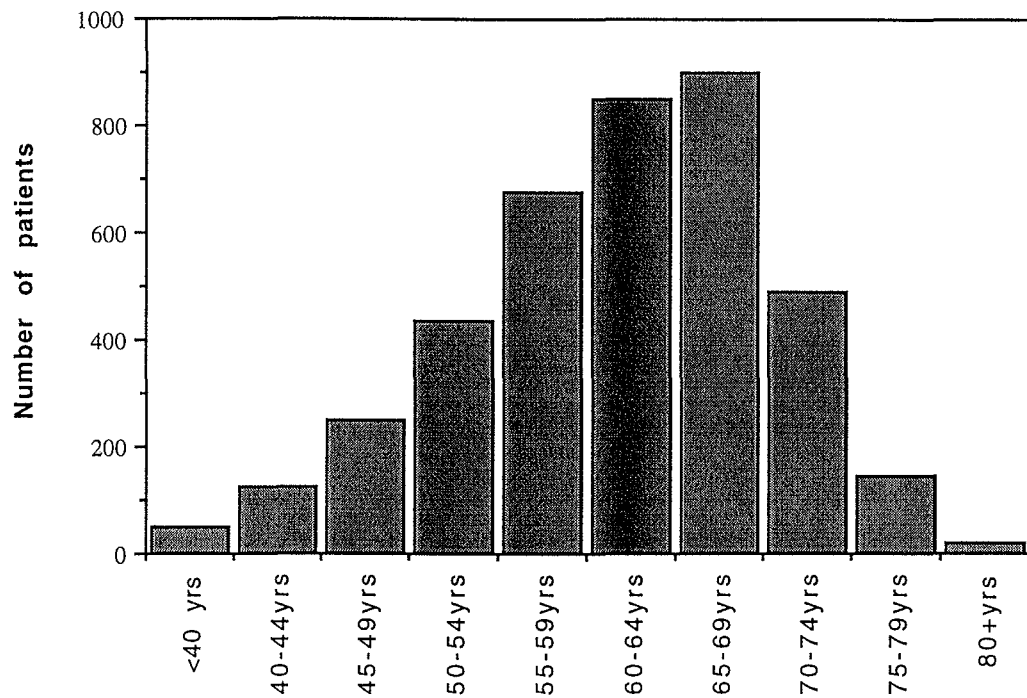


Figure 5.1: Distribution of Respondents by Age

The proportion of males and females in each age cohort, illustrated in Figure 5.2, shows that proportionally more females than males have undergone CABG surgery in the older age groups with 50% of females aged 65 years or more compared 35% of males aged 65 years or more.

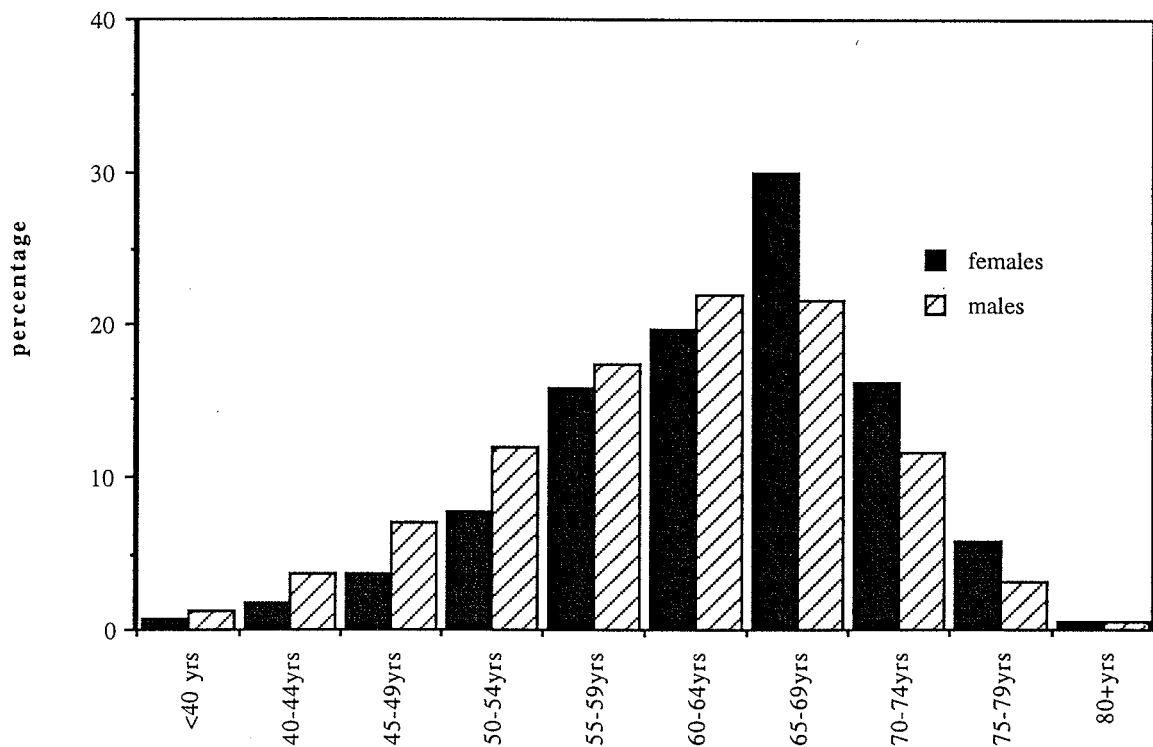


Figure 5.2: Relative Proportions of Males and Females by Age Groups

Of the patients who were followed up, half had undergone previous operations. The most frequently mentioned type of surgery was appendectomy (around 15%). The next most mentioned surgery, at almost 10%, was those who had undergone previous open heart surgery, including CABG. Around 5% of the women had hysterectomies. Only 1.3% of the men had previously had trans-urinary resection. Table 5.2 details the results for previous surgery classified into seven operations.

Table 5.2: Distribution of Previous Operations

Previous Operations	Frequency	Percentage (%)
Cholecystectomy	215	5.4
TURP*	39	1.0
Vascular surgery	44	1.1
Appendectomy	580	14.6
CAG/heart operation	388	9.8
Hernia operation	51	1.3
Hysterectomy	40	1.0
No surgery	1989	50.0
Other	633	15.9
Total	3979	100.0

* transurethral resection procedure

Patients were asked about their symptoms prior to surgery. Table 5.3 presents details of the previous major symptom. Around 3 in 10 had a myocardial infarction. A similar proportion had angina. Less frequently mentioned were hypertension (around 1 in 10) and shortness of breath.

Table 5.3: Illness Prior to Surgery

Previous Illness	Frequency	Proportion(%)
Myocardial infarction	1177	29.6
Shortness of breath	231	5.8
Stroke	37	.9
Angina	1187	29.8
Hypertension	388	9.8
Not defined	573	14.4
Other	86	9.7
Total	3979	100.0

The number of grafts done during surgery was only recorded from 1989 onwards. Of the 1400 patients for whom this information was recorded, more than one third (36.4%) had three grafts. Around a quarter had four grafts with 1 in 5 having two grafts. Only one patient had seven grafts. Figure 5.3 illustrates the distribution of grafts for 1989 to 1992.

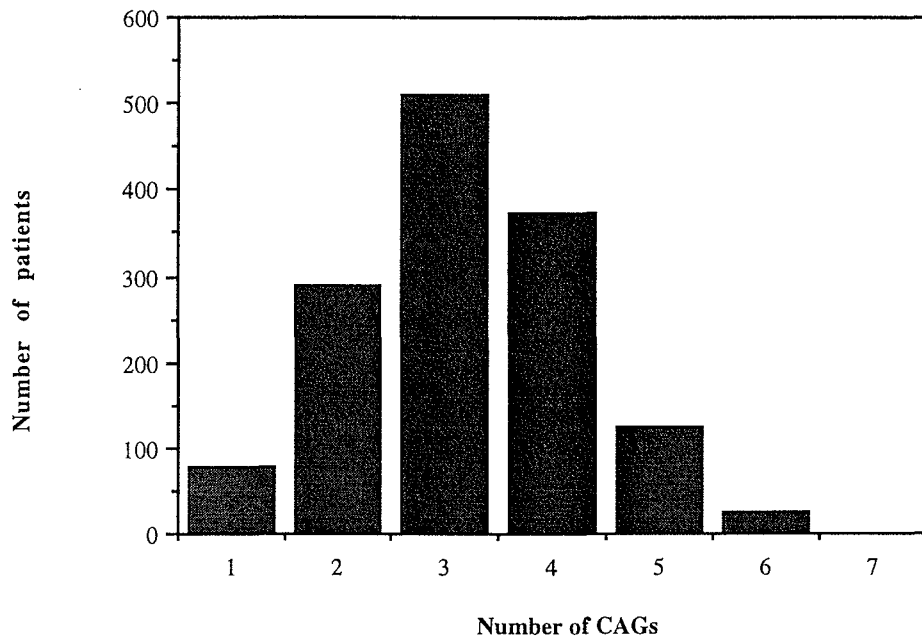


Figure 5.3: Frequency of Coronary Artery Grafts (1989 to 1992)

Patients' weights and heights were recorded on admission to hospital. From these data, their body mass indices (BMI) have been calculated. Table 5.4 gives the distribution of these indices where,

$$\text{BMI} = \frac{\text{weight(in kgs)}}{\text{height}^2(\text{m}^2)}.$$

Table 5.4: Distribution of Body Mass Index for Patients

BMI	Frequency	Proportion (%)
<20kg/m ²	73	1.9%
20 to 24kg/m ²	1169	30.2%
25 to 29kg/m ²	2044	52.8%
30 to 39kg/m ²	570	14.7%
40+kg/m ²	15	0.4%

Garrow (1981) defines BMI scores of 20 to 24 kg/m² as 'desirable' and 30 to 39kg/m² as 'clinically obese'. As seen from Table 5.4, around two thirds of patients (68%) had a BMI above 'desirable' and approximately 15% were 'clinically obese'.

Patients were asked in the follow-up interview about their work status one year after surgery. In terms of not being able to return to work, less than half of a percent (17 respondents) had been retired by their employer since surgery; less than 3% (95 respondents) had retired due to illness; 8 respondents could not do any work; and less than 1% (30 respondents) were not yet back at work.

These results were further examined by classifying patients into those aged 65 years (retirement age) or more and those aged less than 65 years. Table 5.5 shows that more than half of the patients aged less than 65 years had returned to work and around one quarter had retired prior to surgery. Less than 10% had not returned to work or retired because of factors likely to be related to their health status (ie. 'retired by employee', 'retired due to ill health', 'can't do any work' and 'not yet back at work'). Despite the fact that the respondents

were aged 65 years or more at the time of their surgery, more than 13% were back at work one year after surgery.

Work status in the follow-up interview disaggregated by gender is given in Table 5.6. Nearly 13% of females had returned to work compared to around 43% of the males. This is not surprising given the ages of the females in this study, they were probably not in the workforce.

Work Status	aged <65 yrs	(% <65 yrs)	aged 65+ yrs	(% 65+ yrs)
back at work	1117	(52.15%)	178	(13.10%)
retired prior to surgery	544	(25.40%)	966	(71.08%)
retired by choice on age	64	(2.99%)	6	(0.44%)
retired by employer	17	(0.79%)	0	(0.00%)
retired due to ill health	89	(4.15%)	6	(0.44%)
home duties	218	(10.18%)	199	14.64%)
unemployed	58	(2.71%)	1	(0.07%)
can't do any work	5	(0.23%)	3	(0.22%)
not yet back at work	30	(1.40%)	0	(0.00%)
total	2142		1359	

Table 5.5: Work Status of Patients One Year after Surgery by Age

Table 5.6: Work Status of Patients 1 Year after Surgery by Sex

Work Status	female (% fem.)	male (% male)
back at work	85 (12.71%)	1228 (42.82%)
retired prior to surgery	160 (23.92%)	1363 (47.52%)
retired by choice on age	6 (0.90%)	66 (2.30%)
retired by employee	1 (0.15%)	16 (0.56%)
retired due to ill health	4 (0.60%)	92 (3.21%)
home duties	409 (61.14%)	10 (0.35%)
unemployed	0 (0.00%)	59 (2.06%)
can't do any work	4 (0.60%)	4 (0.14%)
not yet back at work	0 (0.00%)	30 (1.05%)
total	669	2868

Macklin (1992) reported that never married/divorced/widowed people in Australia have a higher death rate than their married counterparts. This applied to both males and females. Whether respondents, in this study, were married or not married was recorded for only those cases from 1990 onwards. Of those respondents from 1990 onwards, 78% were now married, 9% were widowed, 6% were separated or divorced and 7% had never married. However, only 61% of female respondents were now married compared to 83% of males. The association between gender and marital status was statistically significant ($\chi^2=62.6$, $df=1$, $p<.01$).

The residential locations of the respondents are presented in Table 5.7. Postcodes of respondents, recorded by the hospital, were aggregated into regions within New South Wales (NSW), other Australian States and territories. Postcodes were not recorded for 166 respondents. The majority of

respondents were from NSW (94%) with the distribution of respondents across the State in a similar pattern to the distribution of population in the State. There were relative high numbers of respondents from Tasmania and The Australian Capital Territory (ACT) compared to other states. Whilst referrals for CABG surgery from the ACT would be expected, referrals from Tasmania are surprising given such procedures are carried out in Melbourne. However, further analyses of respondents from Tasmania revealed that most (80%) had the same surgeon and this anomaly was, therefore, a reflection of the referral patterns for this surgeon. Further disaggregation of the respondents into regions within the metropolitan area of Sydney would have identified the hospital and so breach an undertaking to maintain the confidentiality of the hospital.

Table 5.7: Geographic Distribution of Respondents

Area	Proportion	Number
Sydney Metropolitan	41.9%	1599
Central Coast(NSW)	4.4%	169
Newcastle and Hunter	13.3%	507
North Western NSW	9.6%	365
Mid-North Coast	2.2%	84
Far-North Coast	2.1%	80
Wollongong & district	4.2%	160
Far South Coast(NSW)	2.6%	100
S.W. Sydney & S. Highlands	1.4%	54
Goulburn & District	1.3%	51
Snowy Mountains & MIA	6.5%	247
Penrith & Blue Mountains	2.2%	82
Central West (NSW)	1.2%	46
Far West (NSW)	1.3%	49
ACT	4.1%	157
Victoria	0.0%	3
Queensland	0.0%	9
South Australia	0.0%	3
Western Australia	0.0%	2
Tasmania	1.1%	41

Macklin (1992, p.19) stated that socio-economic area disadvantage (SEI) causes greater threat to those from low SEI compared to those from high SEI. More specifically, the following statements were made:

"low SEI men are 54% more likely to die from ischaemic heart disease and low SEI women 124% more likely to do so;
low SEI men are 102% more likely to die from cerebrovascular disease and low SEI women 69% more likely to do so."

The postcodes of the respondents were further used to place respondents into socio-economic status. These categories of socio-economic status were based on the Socio Economic Index for Areas (SEIFA) classified by the Australian Bureau of Statistics (1993). The five indexes included in SEIFA are:

- Urban Index of Relative Socio-Economic Advantage;
- Rural Index of Relative Socio-Economic Advantage;
- Index of Relative Socio-Economic Advantage;
- Index of Economic Resources; and
- Index of Education and Occupation.

The index of education and occupation was chosen as the measure for this study because of the large proportion of respondents who had retired and thus would not have a measure of economic resources commensurate with their socio-economic status. Since the respondents came from a mix of rural and urban postcodes, neither urban or rural measures of advantage would be appropriate to reflect their status. The Occupation and Education Index has been designed to have a mean of 1000 across all ABS Collector Districts in

Australia with a standard deviation of 100. The apportioning of the index is based on weighting factor multiplied by the proportion of households in the geographic location with selected occupation and education levels. A more detailed description of the weighting procedure is given in the Information Paper for SEIFA (Australian Bureau of Statistics, 1993, p.19). This index was assigned to the respondents according to their postcode in NSW and ACT using the SEIFA software and database on an IBM PC.

In this research, the mean Occupation and Education Index for respondents residing in the ACT was 1149.12 (sd=34.27, n=155) compared with the mean of 1136 from the ACT population from the 1991 Census. The mean index for the respondents in the ACT is significantly higher than the mean for the population in the 1991 Census ($t=4.77, df=154, p<.01$). The mean Occupation and Education Index for respondents residing in NSW was 1017.16 (sd=86.30, n=3490) compared with the mean of 997 from NSW from the 1991 Census. Although the mean for the respondents was higher in NSW compared to the 1991 Census population, the difference was not significant ($t=1.37, df=3489, p=.09$). There was no difference in the mean index for male and female respondents ($t=.65, df=3643, p=.51$).

5.3 Neurological Deficit

Chapter 4 has described how neurological deficit (NSUM) is scored for respondents in the follow up survey. The distribution of NSUM scores is presented in Table 5.8. Almost 7 in 10 CABG patients stated that they had no neurological deficit one year after surgery. Almost 9 out of 10 scored 3 or less on the NSUM scale. Only one patient had the maximum score of 15.

Table 5.8: Distribution of NSUM Scores

NSUM	Frequency	Proportion(%)
0	2768	69.6
1	241	6.1
2	375	9.4
3	112	2.8
4	201	5.1
5	69	1.7
6	117	2.9
7	35	0.9
8	27	0.7
9	20	0.5
10	6	0.2
11	5	0.1
12	2	0.1
15	1	0.0
Total	3979	100

In Table 5.9, respondents are grouped into those with an NSUM score of 0 and those with an NSUM score greater than 0. The table gives the results of tests of significance on the two NSUM categories for 12 factors which may affect neurological outcome. These univariate analyses show that older respondents are less likely to have neurological deficit than younger respondents. Respondents who had retired prior to surgery fare better than those still employed. Longer time on bypass and cross-clamp time have a

significant association with the presence of neurological deficit. The presence of angina at the time of the one year follow-up is also significantly associated with the presence of neurological deficit. Gender and current marital status showed no significant association with the presence of neurological deficit. Neither whether respondents had undergone previous operations nor the type of surgery had a significant association with the presence of neurological deficits. The nature of patients' previous illness had a significant association with the presence of neurological deficit.

Table 5.9: Analyses of the Effect of Selected Variables on NSUM

Category	All	NSUM=0	NSUM \geq 0	Significance
Age (mean)	61.2yrs (n=3937)	62.0yrs (n ₁ =2749)	59.5yrs (n ₂ =1188)	t=8.39,df=3935, p<.01
Sex (%male)	79.2% (n=3979)	78.9% (n ₁ =2768)	80.1% (n ₂ =1211)	$\chi^2=0.78$,df=1, p=.38
Time on bypass (mean)	88.2mins (n=3979)	87.3mins (n ₁ =2768)	90.3mins (n ₂ =1211)	t=2.63,df=3977, p=.01
Cross clamp time (mean)	48.0mins (n=3979)	48.1mins (n ₁ =2768)	47.8mins (n ₂ =1211)	t=0.38,df=3977, p=.35
Angina present (%yes)	8.2% (n=3979)	6.3% (n ₁ =2768)	12.5% (n ₂ =1211)	$\chi^2=42.8$,df=1 p<.01
Retired (%yes)	52.2% (n=1246)	54.5% (n ₁ =1051)	39.5% (n ₂ =195)	$\chi^2=14.90$,df=1, p<.01
Now married (%yes)	78.0% (n=1204)	77.4% (n ₁ =919)	81.4% (n ₂ =285)	$\chi^2=1.49$,df=1, p=.22
Previous Op. (% no prev. op)	50.0% (n=3979)	49.7% (n ₁ =2768)	50.5% (n ₂ =1211)	$\chi^2=13.99$,df=8, p=.08
Previous Illness (% AMI)	34.6% (n=3406)	30.7% (n ₁ =2403)	50.5% (n ₂ =1003)	$\chi^2=138.44$,df=6, p<.01
No. of grafts (% 3 grafts)	36.4% (n=1400)	36.1% (n ₁ =1157)	37.4% (n ₂ =243)	$\chi^2=1.26$,df=4, p=.87
BMI (% 20-24\kg/m ²)	30.2% (n=3877)	30.4% (n ₁ =2707)	29.7% (n ₂ =1170)	$\chi^2=.33$,df=3, p=.95
Occ. and Ed. Index (Mean)	1022.77 (n=3645)	1023.56 (n ₁ =2554)	1020.92 (n ₂ =1091)	t=.82,df=3643, p=.41
PSUM (%PSUM=0)	72.6% (n=3979)	75.4% (n ₁ =2768)	66.1% (n ₂ =1211)	$\chi^2=36.2$,df=1, p<.01

Figure 5.4 presents, graphically, the proportion of respondents in each of illness categories with no neurological deficits. Respondents whose symptoms were reported as angina (82%) are most likely to have no neurological deficits and those who experienced stroke (46%) were most likely to have neurological deficits. Other categories ranged from 57% with no deficit for diabetes symptoms to 67% for shortness of breath. Of respondents with acute myocardial infarction and angina symptoms around 62% reported no neurological deficit at the follow-up interview.

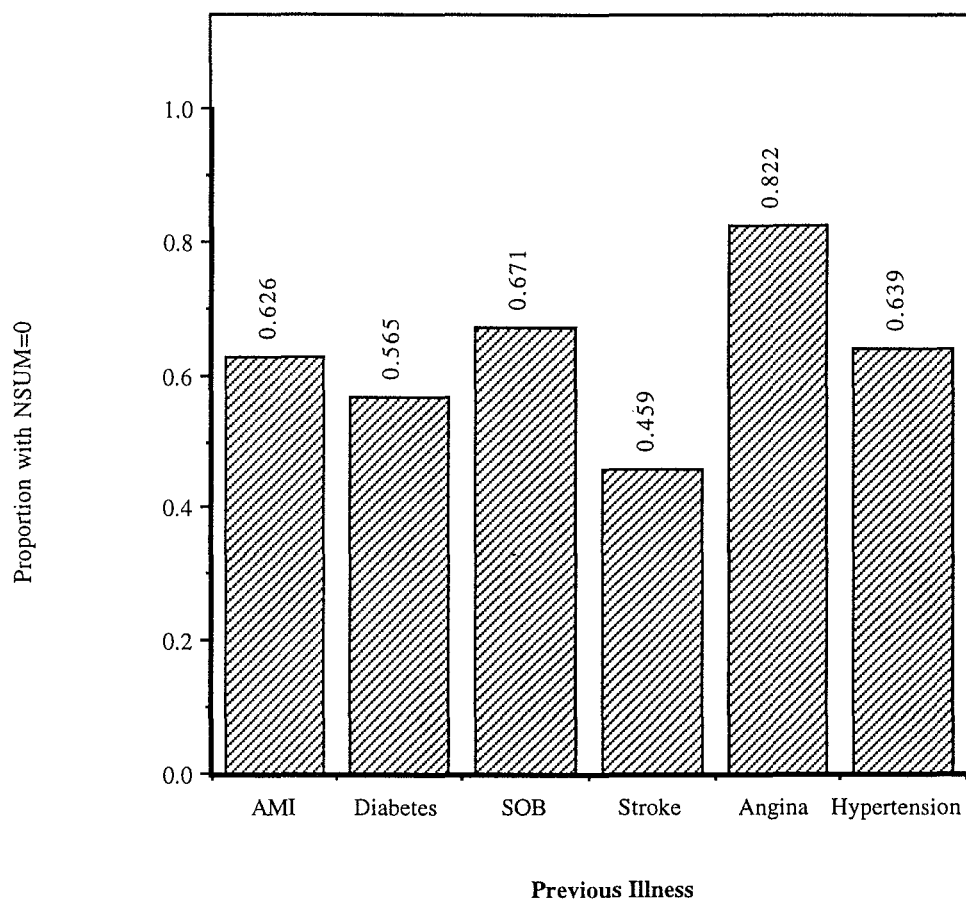


Figure 5.4: Proportion of Respondents with NSUM=0 for each Previous Illness

The number of grafts performed in surgery had no significant association with the presence of neurological deficits. There was a significant association between the presence of neurological deficits and the presence of physical deficits. There was no significant difference in the mean of the Occupation and Education Index for respondents with no neurological deficit and the mean for those with some deficit.

5.4 The Effect of Blood Gas Levels on Neurological Outcome

The effect of blood gas levels, during perfusion, on neurological outcome was analysed by the perfusionist. These analyses found paths for blood gas levels that minimised neurological damage. In addition to the graphical analyses of the perfusionist, analyses were conducted as part of this research to ascertain whether there was a statistically significant difference in neurological outcome for those respondents within the optimum path compared to those outside the optimum path.

Initially the perfusionist examined three dimensional graphs as illustrated in Figures 5.5 to 5.10. The three variables analysed in the three dimensions were:

(a) various blood gas measures including pH, PO₂, PCO₂, arterial blood pressure, haemoglobin (on the x-axis); (b) time on bypass measured at time 0, 10 minutes on bypass and then every 20 minutes until 150 minutes (on the y-axis); and (c) average NSUM, a score for neurological damage (on the z-axis). It can be observed from the three dimensional graphs that the peaks occur when the average NSUM is high. Obviously, the patients would wish to minimise neurological damage and therefore would prefer to be in the

valleys during surgery. More specifically, the process variable in Figure 5.5 is pH. It is clear from the graph, to minimise NSUM, pH should decrease as time on bypass increases. In Figure 5.6, partial oxygen (PO_2) should be minimised for optimal results. Partial carbon dioxide (PCO_2) should slightly increase over time on bypass (Figure 5.7). Figure 5.8 shows that arterial line pressure has little effect on NSUM. There appear to be two distinct valleys for haemoglobin as illustrated in Figure 5.9. In the case of base excess (Figure 5.10), it should be increasing over time to minimise deficit.

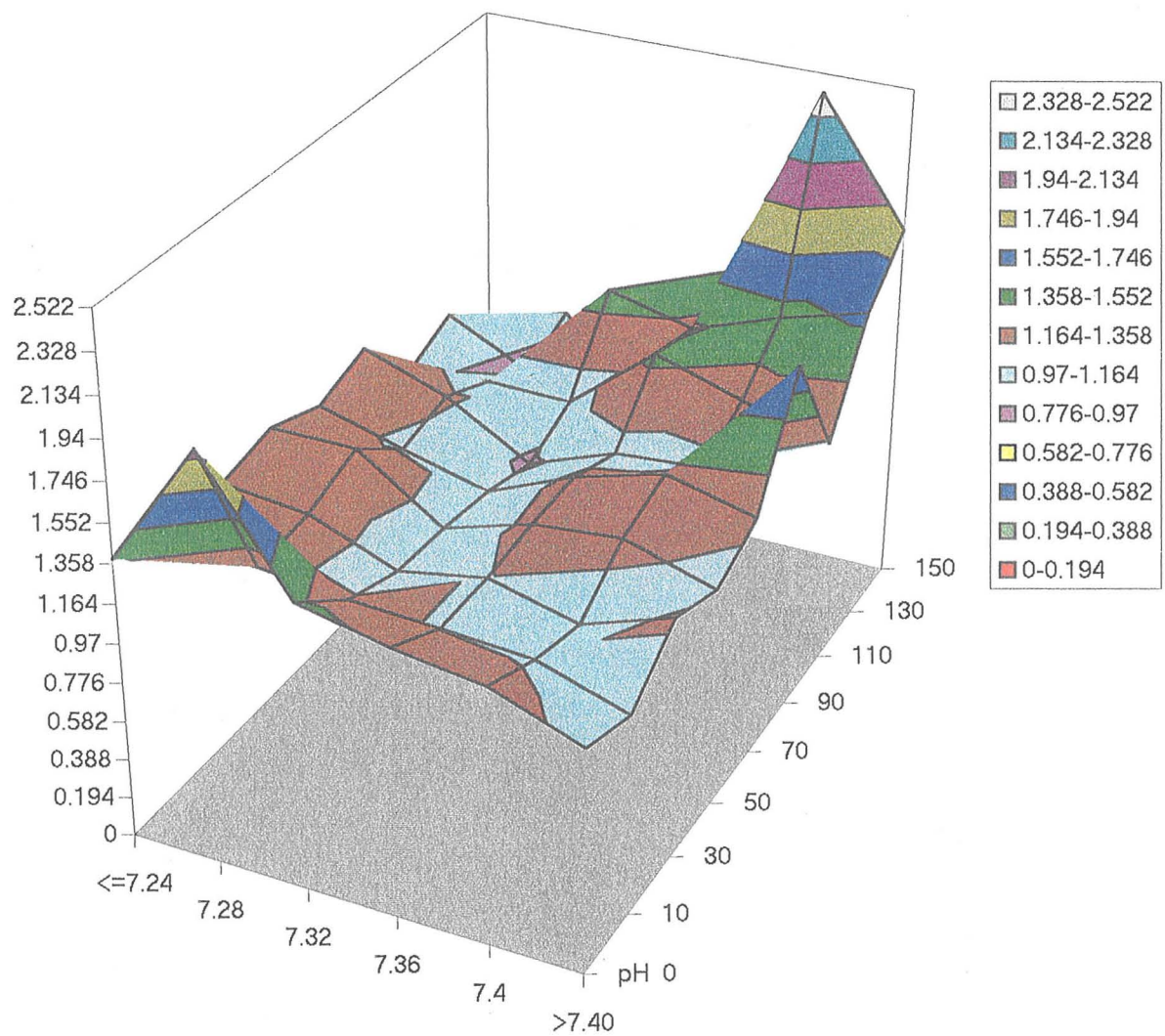


Figure 5.5: pH by Time by Mean NSUM

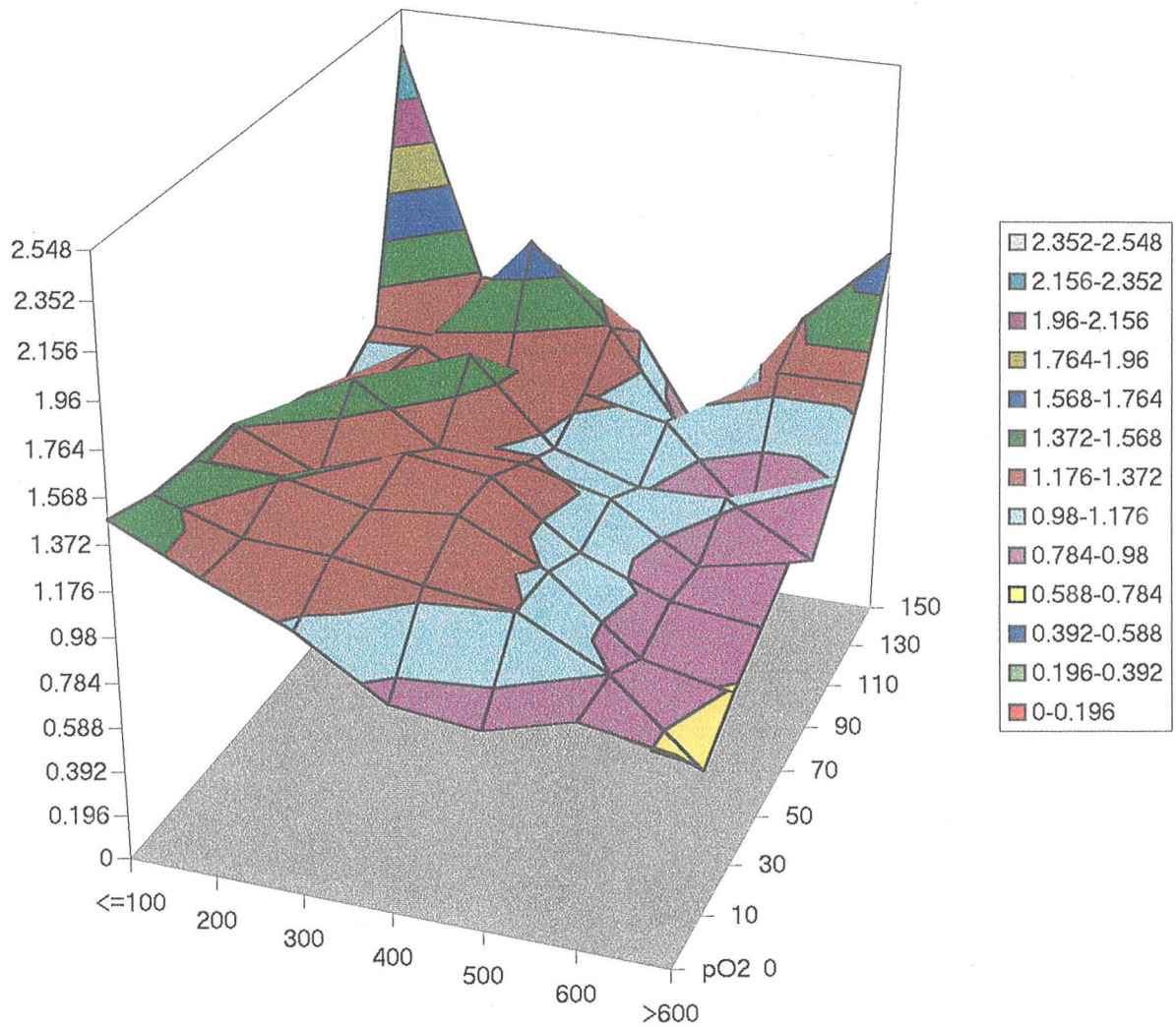


Figure 5.6: PO₂ by Time by Mean NSUM

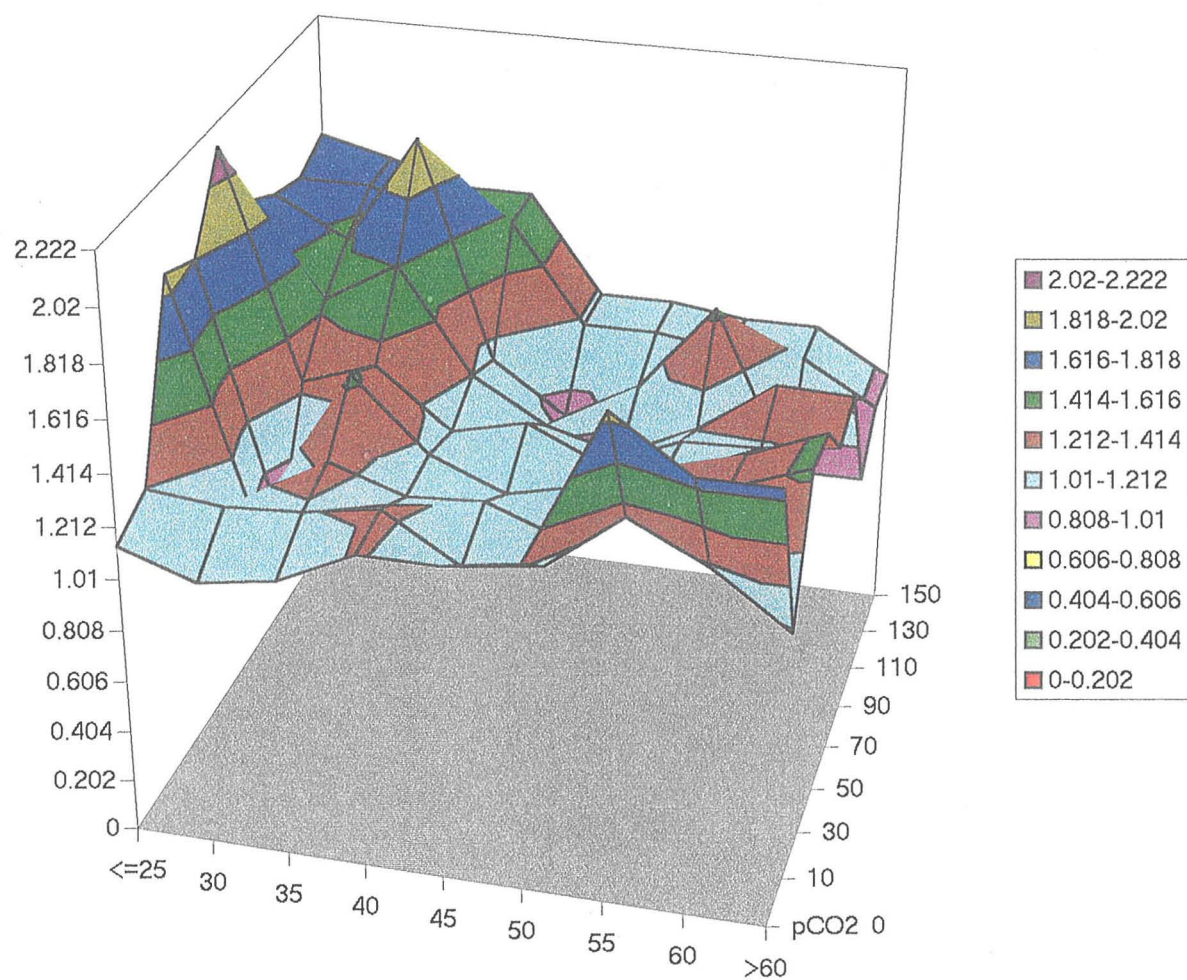


Figure 5.7: PCO₂ by Time by Mean NSUM

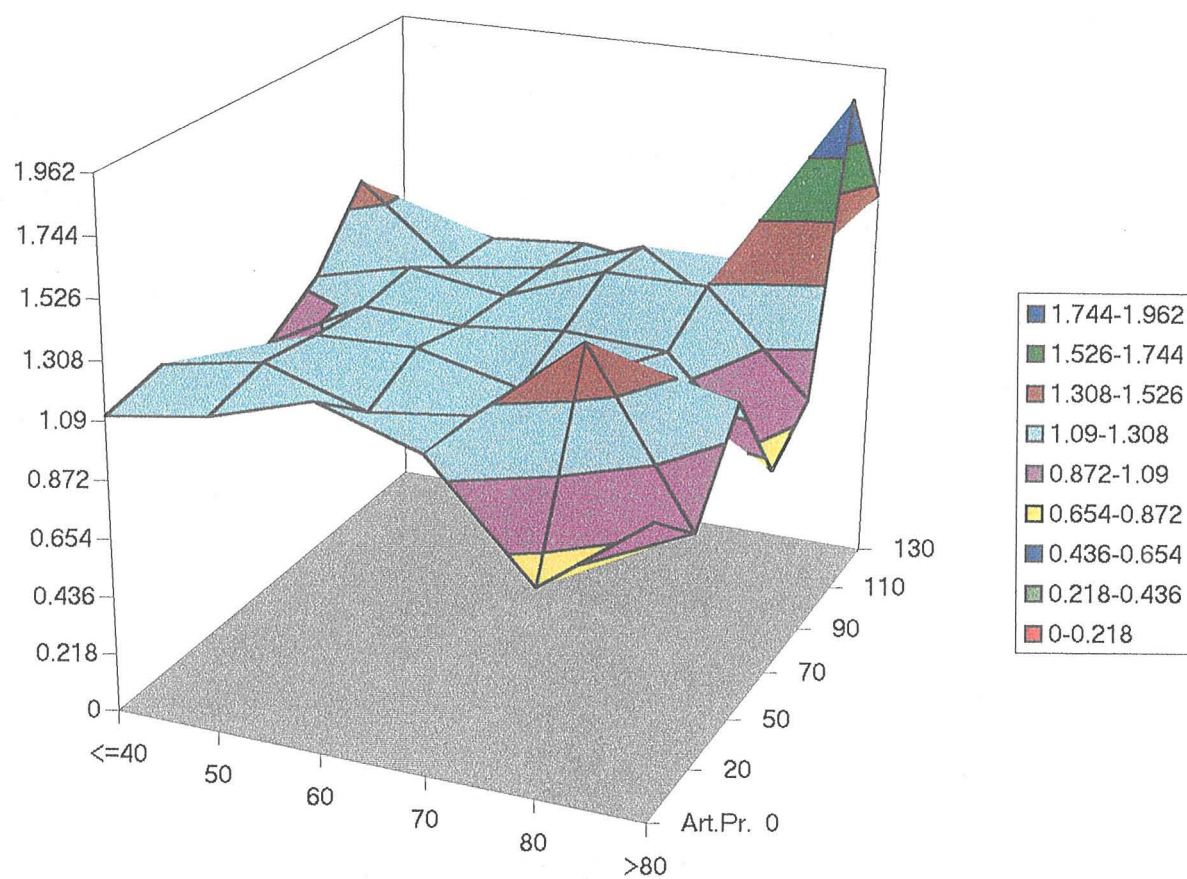


Figure 5.8: Arterial Pressure by Time by Mean NSUM

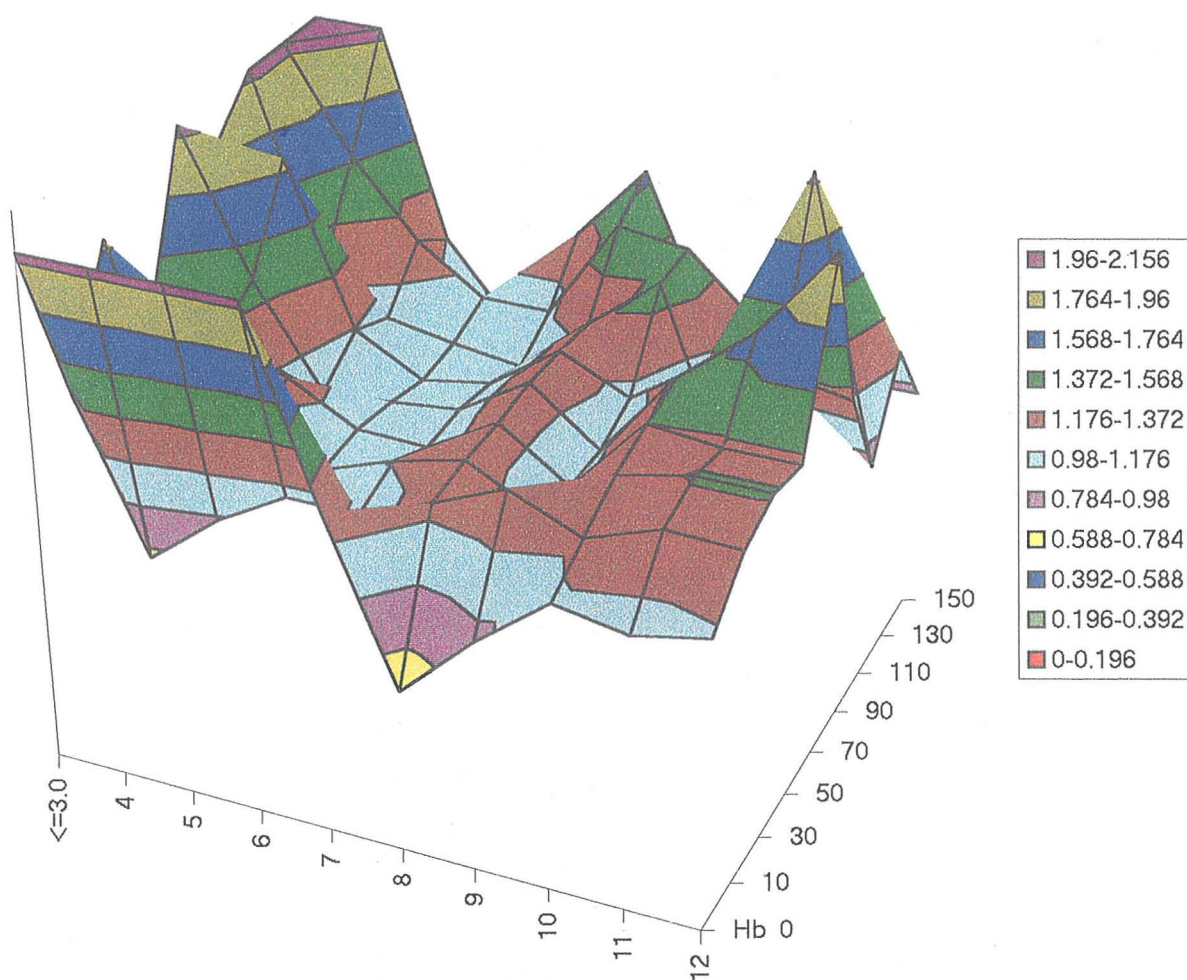


Figure 5.9: Haemoglobin by Time by Mean NSUM

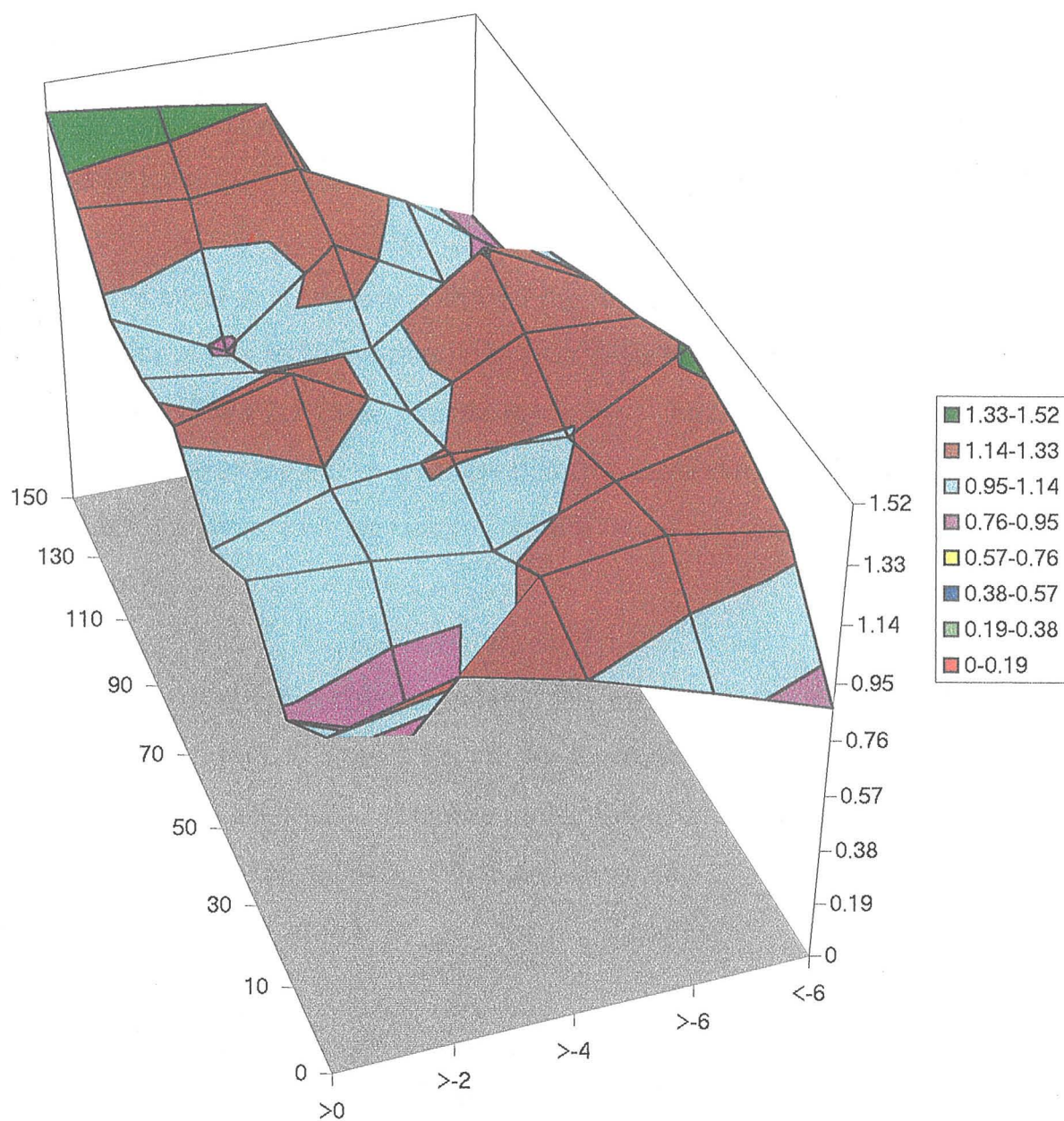


Figure 5.10: Base Excess (BE) by Time by Mean NSUM

In this study, to test whether respondents did, in fact, have better neurological outcomes if the valleys were followed for pH levels over time, they were assigned a code of 1 if they remained in the valley and a code of 2 if they were outside the valley. The SPSSX command program written to allocate these codes is presented in Appendix A. This was done also for pO₂ (Appendix B), pCO₂ (Appendix C), haemoglobin (Appendix D) and base excess (Appendix E). This test was not done for arterial line pressure as the level of this blood gas appears to have little influence on NSUM (see Figure 5.8).

A Student's t-test was used to compare the mean NSUM for respondents in the valley for pH up to 70 minutes on bypass described by Figure 5.5 compared to the mean NSUM for those outside the valley. Although the mean NSUM of 0.86 for those in the valley was less than for those outside the valley (1.02) the difference was not significant at 5% level of significance ($t=1.61$, $df=3829$, $p=.11$).

Applying the same principle of taking the mean NSUM for those in the valley for pO₂ graph (Figure 5.6) and the mean for those outside the valley did result in a significant difference using the Student's t-test. The mean NSUM for those in the valley was 0.45 compared to 1.11 for those outside the valley ($t=7.62$, $df=3876$, $p<.01$).

Using the Student's t-test to examine the difference between mean NSUM in the valley for the pCO₂ graph and the mean outside the valley (Figure 5.7) again produced a lower mean NSUM of 0.94 in the valley compared to 1.05 outside the valley. The difference was not statistically significant ($t=1.78$, $df=3977$, $p=.07$).

As in the previous cases, mean NSUM for respondents in the valley for the haemoglobin graph (Figure 5.9) was less than the mean for those outside the valley - 0.92 compared to 1.04. The difference was not significant using the t-test ($t=1.66, df=3665, p=.10$).

The final test was carried on base excess (BE) where the mean NSUM for respondents in the valley (Figure 5.10) was 0.94 compared to 1.26 for those outside the valley. In this case, the difference for means was found to be significant ($t=3.52, df=3336, p<.01$).

Examination of neurological deficit over time in Figure 5.11 reveals a steady increase in the proportion of people with no neurological deficit from 1985 to 1991 with improvements less marked from 1985 to 1989 and major improvements in this proportion occurring in 1990 and 1991. This improvement in outcomes has corresponded to the implementation of the perfusionist quality improvement process.

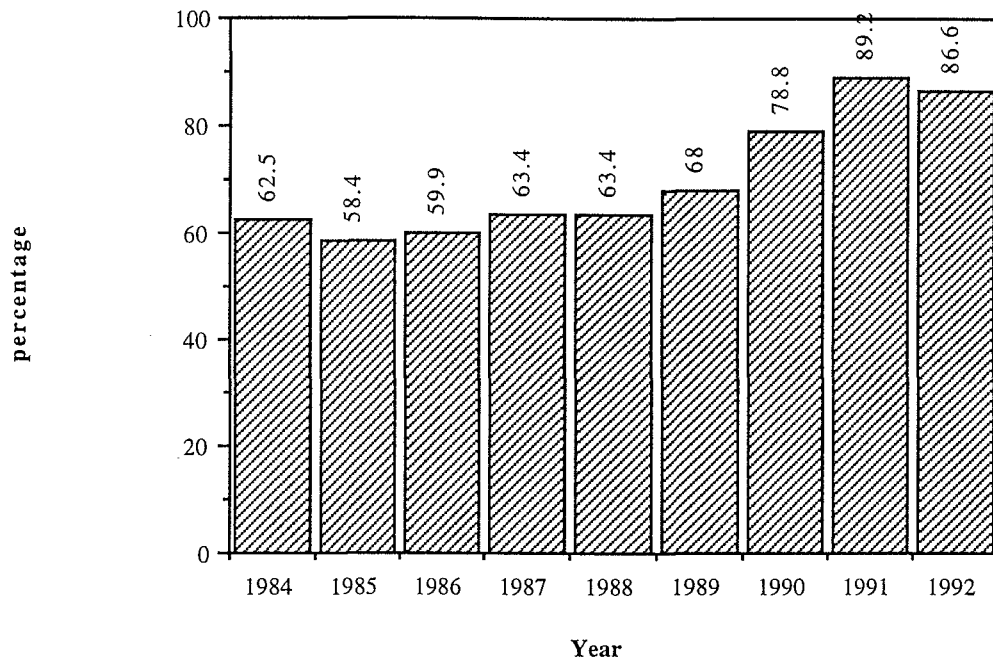


Figure 5.11: Proportion of CABG Patients with NSUM=0 by Year of Surgery

5.5 Physical Deficit

A description of how a physical deficit score (PSUM) is assigned has been given in Chapter 4, Section 4.2. Table 5.10 presents the distribution of PSUM scores for the respondents. More than 70% of respondents reported no physical deficit at the one year follow-up. Most (90%) scored a deficit of 3 or less on the PSUM scale. As was the case for NSUM, only one respondent scored the maximum PSUM of 15.

Table 5.10: Distribution of PSUM Scores

PSUM	Frequency	Proportion(%)
0	2888	72.6
1	217	5.5
2	314	7.9
3	146	3.7
4	165	4.1
5	106	2.7
6	44	1.1
7	34	.9
8	32	.8
9	8	.2
10	15	.4
11	3	.1
12	5	.1
14	1	.0
15	1	.0
Total	3979	100

Table 5.11 presents respondents grouped into categories of no physical deficit (PSUM=0) and some physical deficit (PSUM \geq 0) and examines associations with 11 selected variables. As shown in the table, older respondents were more likely to report physical deficit. Similarly, respondents who had retired were more likely to report some deficit than those who had not retired. Males were less likely to report physical deficit than females with 81% of

males reporting no deficit compared to a proportion of only 76% of females reporting some deficit. The longer respondents were on bypass the more likely they were to report physical deficit. Those with physical deficit spent, on average, 91 minutes on bypass compared to 87% for those with no physical deficit. Cross clamp time had no significant association with physical deficit. There was also a significant association between presence of neurological deficit and the type of previous operation the respondents had undergone. The number of grafts and the respondents' body mass index had no significant association with respondents' reportage of physical deficit. The measure of socio-economic status, the mean of the Occupation and Education Index was significantly higher for those reporting no physical deficit.

Table 5.11: Analyses of the Effect of Selected Variables on PSUM

Category	All	PSUM=0	PSUM>0	Significance
Age (mean)	61.2yrs (n=3979)	60.4yrs (n ₁ =2888)	61.1yrs (n ₂ =1091)	t=1.87,df=3977, p<.01
Sex (%male)	79.2% (n=3979)	80.6% (n ₁ =2888)	75.5% (n ₂ =1091)	$\chi^2=12.60$,df=1, p<.01
Time on bypass (mean)	88.2mins (n=3979)	87.4mins (n ₁ =2888)	90.6mins (n ₂ =1091)	t=2.7,df=3977, p<.01
Cross clamp time (mean)	48.0mins (n=3979)	47.9mins (n ₁ =2888)	48.3mins (n ₂ =1091)	t=0.60,df=3977, p=.55
Retired (%yes)	52.2% (n=1246)	51.0% (n ₁ =945)	55.8% (n ₂ =301)	$\chi^2=14.90$,df=1, p<.01
Married (%yes)	78.0% (n=1204)	79.0% (n ₁ =945)	91.3% (n ₂ =301)	$\chi^2=2.12$,df=1, p=.15
Previous Op. (% no prev. op)	50.0% (n=3979)	51.1% (n ₁ =2888)	47.0% (n ₂ =1091)	$\chi^2=28.55$,df=8, p<.01
Previous Illness (%AMI)	34.6% (n=3406)	32.6% (n ₁ =2471)	39.8% (n ₂ =935)	$\chi^2=18.88$,df=6, p<.01
No. of grafts (% 3 grafts)	36.4% (n=1400)	36.6% (n ₁ =1068)	35.5% (n ₂ =332)	$\chi^2=2.97$,df=4, p=.56
BMI (% 20-24\kg/m ²)	30.2% (n=3877)	30.6% (n ₁ =2821)	29.1% (n ₂ =1056)	$\chi^2=5.40$,df=3, p=.14
Occ. and Ed. Index (Mean)	1022.77 (n=3645)	1025.19 (n ₁ =2655)	1016.28 (n ₂ =990)	t=2.70,df=3643, p=.01

The proportions of respondents reporting no physical deficit for each type of previous surgery are reported in Figure 5.12. Respondents who had undergone previous heart surgery were most likely to report some physical deficit whilst respondents who had hysterectomies were least likely to report deficits. Around 74% of respondents who had no previous operations reported no physical deficit.

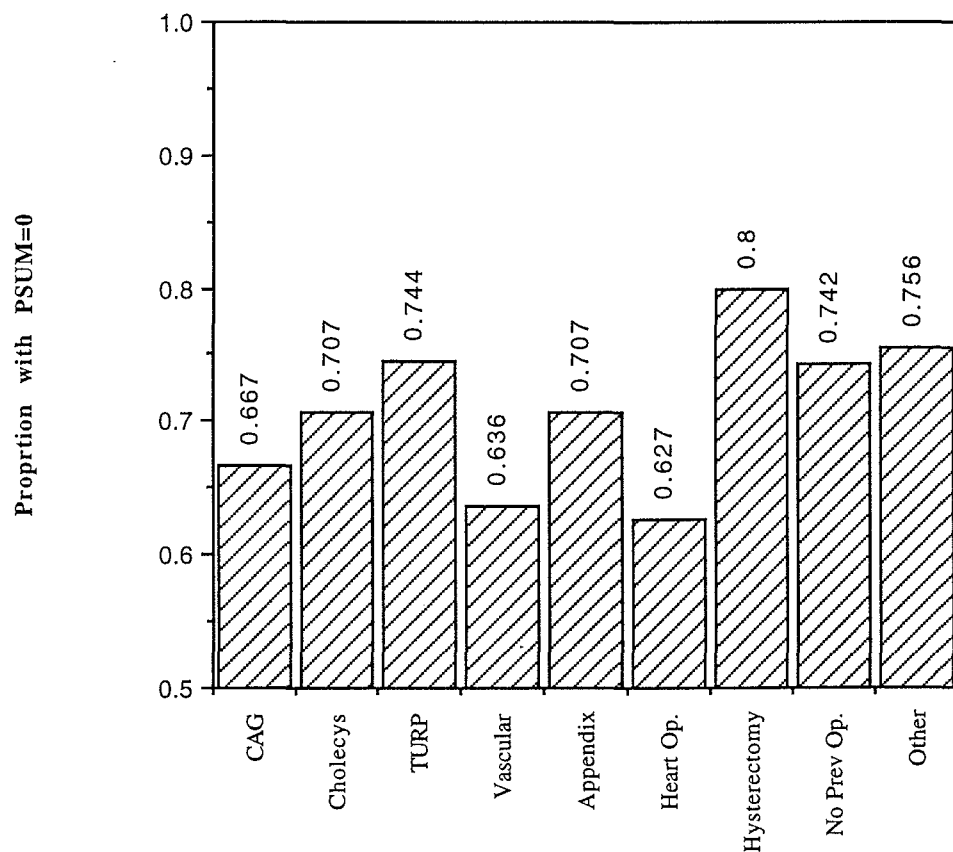


Figure 5.12: Proportion of Respondents with PSUM=0 by Previous Operation

With respect to previous illness or symptoms, respondents whose symptoms were stroke (32%) or acute myocardial infarction (31%) were most likely to report physical deficit compared to those whose symptoms were angina (24%) who were least likely to report a deficit. Figure 5.13 shows similar

proportions (around 73%) of respondents reported no physical deficits if their symptoms were hypertension, diabetes or shortness of breath (SOB).

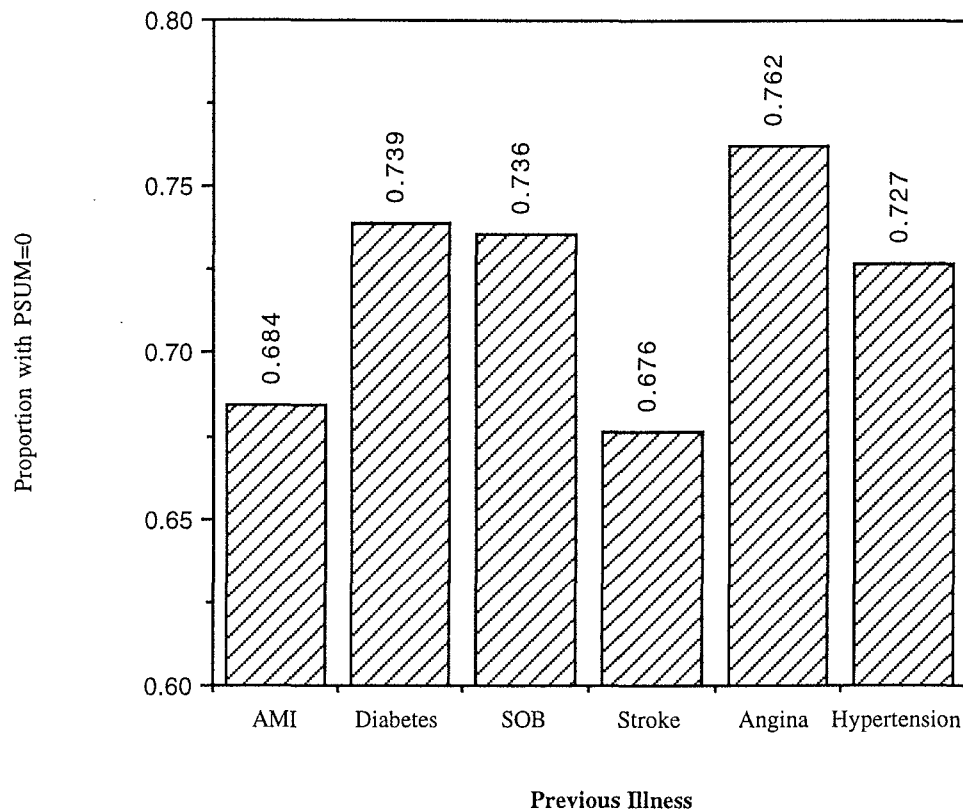


Figure 5.13: Proportion of respondents with PSUM=0 by Illness Category

Table 5.12 summarises the descriptive responses given by respondents of their physical state. Respondents could make more than one comment about their physical state and thus total responses add to more than the number of respondents. Most responses are positive about their physical state - 'pain free', 'improved', 'fit, well and happy' and 'never been better'. The comments made by respondents only provide a rough picture of the types of physical deficits reported but do not provide an accurate picture of levels of deficit.

As an example, only 48 (1.2%) respondents commented that angina had returned, whilst Figure 5.14 shows that a much higher number reported the presence of angina when probed by the follow-up survey questions used to determine PSUM.

Table 5.12: Respondents Description of Physical State

Description	Responses	% Responses
pain free	3402	49.02%
improved(pain)	208	3.00%
not improved	74	1.07%
worse after operation	9	0.13%
sorry had operation	4	0.06%
fit, well & happy	1777	25.61%
never been better	392	5.65%
not well	95	1.37%
angina returned	48	0.69%
breathless	141	2.03%
pulmonary oedema	12	0.17%
AMI since operation	15	0.22%
weight problem	9	0.13%
poor eyesight	23	0.33%
cannot identify problem	12	0.17%
heart failure	23	0.33%
stroke post-op	27	0.39%
sore chest	98	1.41%
varicose veins	11	0.16%
sore legs	55	0.79%
oedema	10	0.14%
infection	25	0.36%
sore arms	19	0.27%
impotent	1	0.01%
pvd	59	0.85%
tires easily	88	1.27%
died later	67	0.97%
cannot walk	3	0.04%
fitter than before	2	0.03%
low activity, potters around	23	0.33%
other illness	140	2.02%
dizzy, balance problems	22	0.32%
arrhythmia	35	0.50%
no strength, weak	2	0.03%
hernia	5	0.07%
keloid	2	0.03%
ulcer	2	0.03%

Figure 5.14 also illustrates the marked improvement in the proportion reporting they are free of angina one year after surgery. The proportion rose from around 90% in 1985 to 1988 to around 94% in 1990 to 1992.

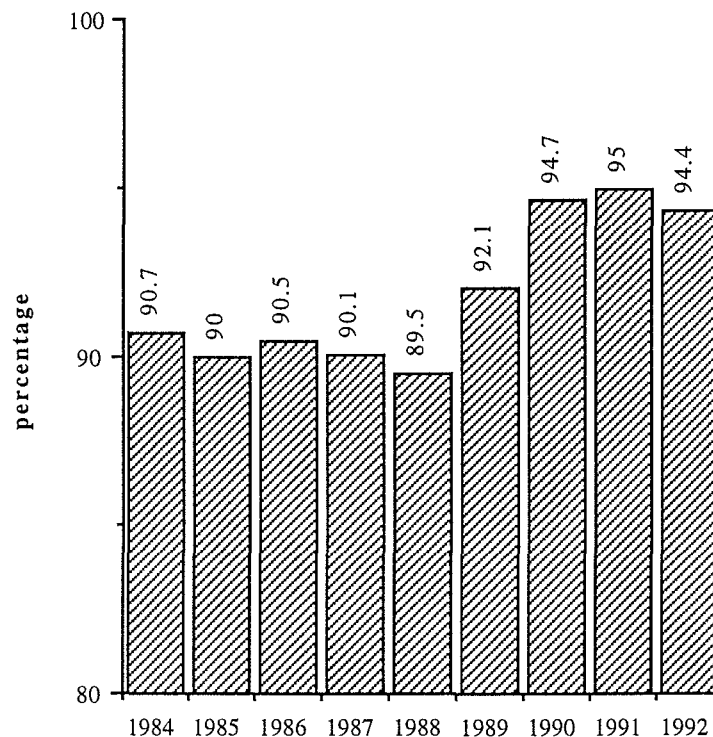


Figure 5.14: Proportion of Respondents Free of Angina by Year

The proportion of respondents who reported no physical deficit for each of the years 1984 to 1992 is presented in Figure 5.15. Although some improvement to 1990 occurred, there was a drop in the proportion in 1991 which was maintained in 1992. The absence of physical deficit has not shown the same improvements over time as has been shown for the absence of neurological deficits (Figure 5.11). One reason for the lack of improvement in physical deficit may be that the perfusionist had not analysed the three-dimensional graphs relating time and blood gases to PSUM as he had done in the case of NSUM, described in Section 5.2.

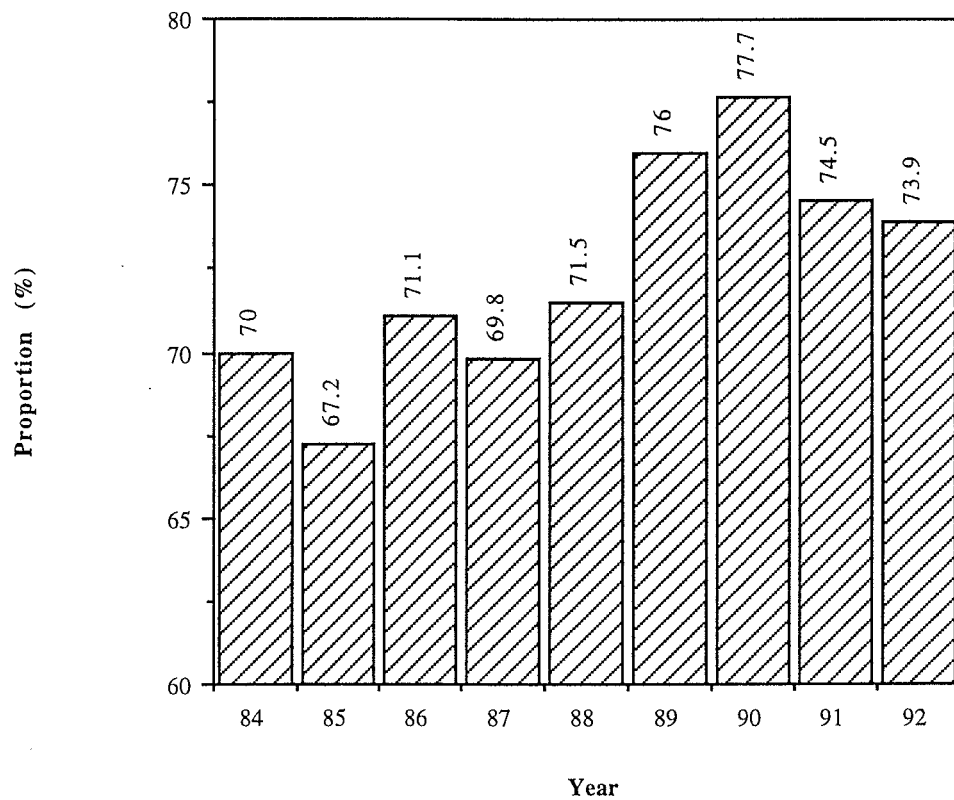


Figure 5.15: Proportion of Respondents with PSUM=0 by Year

5.7 Comparison of Neurological and Physical Deficits

The distributions of NSUM and PSUM, given in Tables 5.8 and 5.10, respectively, show similar patterns with around 70% of respondents reporting no neurological deficit and approximately the same proportion reporting no physical deficit. The maximum score of 15 was recorded by only one respondent for both scores. To measure whether a respondent was likely to report both a neurological and physical deficit, correlation between NSUM

and PSUM was calculated using Pearson's Correlation Coefficient. There was a strong positive correlation of 0.92 ($p < .01$).

Comparing factors that may contribute to NSUM and PSUM, whilst relatively older respondents were less likely to report a neurological deficit, they were more likely to report a physical deficit. Retirees were less likely to report a neurological deficit but more likely to report a physical deficit.

Although gender had no statistically significant effect on the reporting of neurological deficit, women were more likely to report a physical deficit. Mean time on bypass was significantly higher for respondents reporting both neurological and physical deficits.

The type of surgery previous to CABG surgery had no association with reporting of neurological damage but a significant association with the reporting of physical deficit. There was a strong association of previous illness or symptoms with both the reporting of some physical and neurological damage.

The number of grafts and body mass index had no significant association with either neurological or physical deficit. There was no significant difference between the mean Occupation and Education Index for those with neurological damage and the mean for those without neurological damage. However, there was a marked difference between the mean index for respondents with no physical deficit and the mean index for those with physical deficit.

5.8 Summary

A description of 3979 CABG patients followed up was presented at the beginning of this chapter where it is shown that the patients in this study have a similar profile by gender and age to other reported studies. The next section examined analyses associations of various measures with the presence or absence of neurological deficit. Age, length of time on bypass, the presence of angina 12 months post-surgery, whether the patient was retired and the patient's symptoms pre-surgery had significant statistical associations with the presence of neurological deficit. A similar analysis for associations with the presence of physical deficits was presented in the next section. In this case age, gender, length of time on bypass, whether the patient was retired, the nature of previous operations, previous symptoms and socio-economic status were statistically significantly associated with the presence or absence of physical deficit.

The simple univariate analyses in this chapter have either treated NSUM and PSUM as logistic variables (takes only one of two value - '0' and '1') or have taken the mean of NSUM. However, there are some limitations to these types of analyses. Firstly, treatment of NSUM and PSUM as bivariate logistic variables has provided some indication of factors that may be influencing the presence or absence of reported deficit but has not taken into account the range of value taken by NSUM and PSUM (from '1' to '15'). Secondly, the distribution of NSUM is skewed towards '0' with almost 70% of respondents reporting this score. The mean, therefore, does not provide a good measure of centrality. To address some of the limitations identified in this chapter, the

next chapter proposes multivariate models for NSUM and PSUM where each are classified into five separate categories of scores to address some of the limitations .

CHAPTER 6

MULTIVARIATE ANALYSES OF CORONARY ARTERY BYPASS GRAFT PATIENT DATA

CHAPTER 6 MULTIVARIATE ANALYSES OF CORONARY ARTERY BYPASS GRAFT PATIENT DATA

6.1 Introduction

The broad aim of this chapter is to provide multivariate regression models, based on selected explanatory variables, that can be used to predict coronary artery bypass patients' neurological and physical outcomes. Chapter 5 has presented the analyses of the effects of explanatory variables on the outcome measures using univariate analyses.

This chapter first discusses the shortcomings of the univariate analysis undertaken and documented in Chapter 5. This is followed by an examination of the literature on statistical models that have response variables with more than two categories. Thirdly, the SAS Logist Model that is used for analysis is described and the reasons for its selection are given with examples of applications of the procedure. Finally, a 'good fitting model' is presented. Agresti (1990, p.79) defines a good fitting model as one that

"evaluates effects of explanatory variables, describes associations and interaction linkages, and produces improved estimates of response probabilities".

6.2 Critique of Univariate Analysis

The univariate analyses of variables that have a significant statistical association with the presence ($NSUM > 0$) or absence ($NSUM = 0$) of

neurological deficit have been summarised in Chapter 5. Similar analyses have been conducted for the presence ($PSUM > 0$) and absence ($PSUM = 0$) of physical deficit. Both NSUM and PSUM are the response variable. Patients' demographic, environmental and clinical variables are explanatory variables. In the univariate analyses NSUM and PSUM were treated as binary responses. However, these analyses do not account for the size of deficit score for either neurological or for physical deficit. An additional weakness is that there are dangers of looking at large numbers of 2-way tables rather than multi-way tables that describe the full array, as warned by Bishop (*et al.*, 1976). Similarly, Agresti (1990, p.31) has also cautioned on this problem in his statement:

"inadequate and even incorrect conclusions can result from studying the variables only two at a time... association between 2 variables can change drastically, when a third variable is controlled".

A third weakness is that the analysis in Chapter 5 does not account for interactions between explanatory variables. As an example, the results have shown that gender appears to have no effect on the presence or absence of neurological deficit. There may, however, be an interaction between age and gender which may have an effect on neurological outcome.

Therefore, to solve such problems the analysis in this chapter uses ordered polytomous logistic regression. The term *ordered* is used because higher scores for neurological and physical deficit indicate greater damage and lower scores indicate less damage. The deficit scores can be put in more than two

categories and are therefore *polytomous* response variables. Given both PSUM and NSUM are in the range of 0 to 15, one approach is to treat them as continuous variables. However, this was rejected given the high proportion of zeros for both measures. To assist in describing polytomous logistic models, the simpler dichotomous logistic regression will be discussed in the following section on logit models. Dichotomous logistic regression is discussed because the procedures used in providing models for bivariate response variables are applied to regression with more than two response categories by taking one category as a standard and comparing it to the other categories.

6.3 Logit Models

Models that have more than two response categories are modifications of the logit model for binary responses. The treatment of neurological deficit as either present or absent is an example of a binary response. Each patient would be termed as having either a 'success' or 'failure' with outcomes of 1 or 0. A random variable Y has the Bernoulli distribution if

$$P(Y=1)=p \text{ and } P(Y=0)=1-p.$$

When Y_i has a Bernoulli distribution with parameter p_i , the probability function may be written as

$$\begin{aligned} P(Y_i=y_i) &= f(y_i, p_i) = p_i^{y_i} (1-p_i)^{1-y_i} \\ &= (1-p_i) \left[\frac{p_i}{1-p_i} \right]^{y_i} \\ &= (1-p_i) \exp(y_i \log(\frac{p_i}{1-p_i})) \end{aligned}$$

The logit of p is

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

and it represents the log-odds of a success ($Y=1$).

As an example, in the case study of coronary artery bypass graft patients the response variable Y is deficit and there may be R explanatory variable say A , B , C ... A might be sex, B might be age and C might be previous symptoms, and so on. The relationship between Y and the R explanatory variables can be examined by constructing a model for the logit in terms of the R variables.

There are various models that can be proposed such as a no interactions model, two or more interactions models and the fully saturated model with interaction terms for all explanatory variables.

6.4 Logistic Regression

The linear probability model $E(Y) = p(x) = \alpha + \beta x$ for a binary variable Y has a major difficulty in that $p(x)$ is a probability and so must fall between 0 and 1 whilst linear functions can take values over the entire real line. A non-linear relationship between $p(x)$ and x would be expected. To satisfy an assumption of least squares estimation, variability must be constant. However, as $p(x)$ moves towards either 0 or 1, the conditional distribution of Y is more concentrated at a single point, and the variance more towards 0. Therefore, curvilinear models are used instead.

$$p(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \text{ is the logistic regression function.}$$

For $\beta < 0$, $p(x)$ approaches 0 when $x \rightarrow \infty$

For $\beta > 0$, $p(x)$ approaches 1 when $x \rightarrow \infty$

Since

$$\begin{aligned}\frac{\partial p(x)}{\partial x} &= \frac{\beta \exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} - \frac{\beta \exp(\alpha + \beta x)^2}{(1 + \exp(\alpha + \beta x))^2} \\ &= \frac{\beta \exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \left(1 - \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}\right) \\ &= \beta p(x)(1 - p(x)).\end{aligned}$$

The steepest slope occurs when $p(x) = 1/2$.

That is $\exp(\alpha + \beta x) = 1$

$$\alpha + \beta x = 0$$

$$x = \frac{-\alpha}{\beta}.$$

The tangent to the curve at this point has the gradient $\frac{\beta}{4}$.

The odds of making response 1 is

$$\frac{p(x)}{1-p(x)} = \exp(\alpha + \beta x) = e^{\alpha} (e^{\beta})^x.$$

Thus the odds increases multiplicatively by e^{β} or every unit increase in x .

The log odds has a linear relationship to x :

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \alpha + \beta x.$$

Therefore the appropriate link is the logit relating $E(Y)$ linearly to x is the log odds transformation.

Logistic models refer to odds and estimate odds at one value of x divided by the estimated odds of another value of x and this is known as the odds ratio.

It is often helpful to plot sample logits against x . Let n_i be the number of observations at the i th value of x . The number of "1" outcomes are reported

for each value of x . Let Y_i denote the outcome of this binomial random variable. The i th sample logit is the $\log\left(\frac{Y_i}{n_i - Y_i}\right) = \log\left(\frac{p_i}{1 - p_i}\right)$ where $p_i = \frac{Y_i}{n_i}$.

However this is not defined when Y_i equals 0 or n_i and therefore is replaced by an adjusted value

$$\log\left(\frac{Y_i + \frac{1}{2}}{n_i - Y_i + \frac{1}{2}}\right) \text{ called the empirical logit.}$$

Agresti (1990, Table 4.2, p. 88) gives an example of cancer remission as a response, where '1' indicates a success. An explanatory variable is a labelling index (LI). This index measures proliferative activity of cells and represents the percentage of cells that are labelled. There were 27 observations at 14 levels. This table is reproduced below as Table 6.1:

Table 6.1 Cancer Remission & Labelling Index

LI	No. Cases	No.	p (logistic)	p(probit)	p(linear)
Remissions					
8	2	0	0.068	0.053	-.003
10	2	0	0.089	0.075	0.053
12	3	0	0.115	0.103	0.109
14	3	0	0.148	0.138	0.164
16	3	0	0.189	0.181	0.220
18	1	1	0.237	0.231	0.276
20	3	2	0.293	0.288	0.331
22	2	1	0.357	0.350	0.387
24	1	0	0.425	0.417	0.443
26	1	1	0.497	0.487	0.498
28	1	1	0.569	0.557	0.554
32	1	0	0.702	0.689	0.665
34	1	1	0.759	0.748	0.721
36	3	2	0.849	0.846	0.832

(Source: Agresti, 1990, Table 4.2, p.88)

Because of the small sample size the LI were grouped as in Table 6.2.

Table 6.2 Empirical Logits for Table 6.1 and Predicted Number of Remissions

LI	No. Cases	No. Remissions	Empirical Logit
8-12	7	0	-2.71
14-18	7	1	-1.47
20-24	6	3	0.00
26-32	3	2	0.51
34-38	4	3	0.85

The empirical logit for LI at 8 to 12 has $y_i = 0$ and $n_i = 7$ so the empirical logit

$$= \log\left(\frac{\frac{1}{2}}{7 + \frac{1}{2}}\right) = \log\left(\frac{1}{15}\right) = -2.71$$

The empirical logits in Table 6.2 show an increasing trend which would indicate the linear effect of LI be modelled on the logit scale. The maximum likelihood estimates for the logistic regression model are $\hat{\alpha} = -3.777$ and $\hat{\beta} = 0.145$ (ASE = 0.059) where ASE=estimated asymptotic standard error. So for a unit change in LI, the estimated odds of remission are multiplied by $\exp(0.145) = 1.16$.

Now, $p(x) = 0.5$ ($p(x)$ = predicted probability of remission) at $LI = \frac{-\alpha}{\beta} = 26.00$.

The estimated logit $L = \hat{\alpha} + \hat{\beta}$ with an estimated odds $\exp(L)$ and estimated probability $\hat{p} = \frac{\exp(L)}{1 + \exp(L)}$. For $LI = 38$, as an example, the predicted number of

remissions for the 3 cases is $3(0.85) = 2.55$. In making inferences about β from logistic regression, the confidence interval for β is $\hat{\beta} \pm z_{\alpha/2}(ASE)$ where $z_{\alpha/2}$ is the standard normal z score, z_i , and $\Pr(z > z_i) = \frac{\alpha}{2}$ % for a $(1 - \alpha)$ % confidence interval.

Wald's statistic is the square of the ratio of a parameter estimate to its standard error and has approximately, a chi-square distribution with $df=1$ if the parameter is zero. If $Z = \frac{\hat{\beta}}{SE(\hat{\beta})}$, then the observed Z is $\frac{0.145}{0.059} = 2.45$ and the

observed Wald statistic for testing $H_0: \beta=0$ is $Z^2=5.96$. This shows a strong

indication of association and that, in fact, $\beta > 0$. The case discussed above can have explanatory variables that are continuous or categorical.

For a given logit model, we can use model parameter estimates to calculate predicted logits and therefore predict probabilities and expected frequencies $\{m_{ij} = n_{i+} \hat{p}_{j|i}\}$ where m_{ij} is the expected number of counts in the (i,j) th cell and $\hat{p}_{j|i}$ is the probability in the j th column of the i th row. With large enough expected frequencies we can test goodness-of-fit with a Pearson or likelihood ratio χ^2 statistic.

The likelihood ratio χ^2 statistic is denoted as $G^2(M)$ for the model M . $G^2(M) = 2 \sum \sum n_{ij} \log \left(\frac{n_{ij}}{m_{ij}} \right)$. The degrees of freedom equals the number logits less the number of linearly independent parameters in the model.

This section has demonstrated how to model the relationship between a bivariate response variable and one explanatory variable. To analyse relationships between several variables there is a need to look at multi dimensional contingency tables.

Take, for example, multi dimensional contingency tables, at say the simplest level of a three-variable case x , y and z . If we display a table with a cross tabulation of x and y at different levels of z , that is, z is controlled for, it is called the partial table. However, if the partial tables are combined, that is, z is ignored - these are called marginal tables.

Agresti (1990) presents data relating to whether defendants receive death penalties according to race. Table 6.3, which is reproduced from Agresti (1990,

Table 5.1,p. 136), clearly shows that approximately 12% of white defendants and 10% of black defendants received the death penalty. However, when the race of the victim is controlled for, quite different partial tables show that when the victim was white, the death penalty was imposed almost 5% more often if the defendant was white. Similarly, when the victim was black, the death penalty was imposed almost 6% more often on black defendants than on white defendants.

Table 6.3: Death Penalty Verdict by Defendant's Race and Victim's Race

Defendant's Race	Victim's Race	Death Penalty		Percentage
		Yes	No	Yes
White	White	19	132	12.6
	Black	0	9	0.0
Black	White	11	52	17.5
	Black	6	97	5.8

(Source: Agresti, 1990, p.136)

There are three types of dependence - mutual, joint and conditional. For the three variable example x , y and z are described as mutually independent when there are no interaction terms. Where m_{ijk} = expected frequencies, then $\log m_{ijk} = \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z$. Secondly, y is said to be jointly independent of x and z if the model has an interaction term for x and z ($\log m_{ijk} = \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z + \lambda_{ik}^{xz}$). Thirdly, when x and y are independent in the partial table for the k th category of z , then x and y are said to be conditionally independent at level k of z . They are conditionally independent given z

when independent at all levels of z and have interaction levels in x and z and y and z ($\log m_{ijk} = \mu + \lambda_i^x + \lambda_j^y + \lambda_k^z + \lambda_{ik}^{xz} + \lambda_{jk}^{yz}$).

For three dimensional models, to test the goodness-of-fit, likelihood ratio statistic G^2 can be used or the Pearson statistic χ^2 .

$$G^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left(\frac{n_{ijk}}{\hat{m}_{ijk}} \right)$$

$$\chi^2 = \sum_i \sum_j \sum_k \frac{[n_{ijk} - \hat{m}_{ijk}]^2}{n_{ijk}}$$

The degrees of freedom equals the difference in the number of parameters in the general case when the model holds.

The general case has only one constraint

$$\sum_i \sum_j \sum_k p_{ijk} = 1 \text{ so } df = IJK - 1 \text{ where } i=1, \dots, I, j=1, \dots, J, k=1, \dots, K$$

For the mutually independent model (x, y, z)

$$\{p_{ijk} = p_i \cdot p_j \cdot p_k\} \text{ ie. } I-1, J-1 \text{ and } K-1.$$

$$\text{so } df = (IJK - 1) - (I + J + K - 3)$$

$$= IJK - I - J - K + 2$$

Table 6.4 gives the residual degrees of freedom for loglinear models for three-way tables and is reproduced from page 175 of Agresti (1990).

Table 6.4 Residual Degrees of Freedom for Loglinear Models for Three-Way Tables

Model	Degrees of Freedom
(X, Y, Z)	$IJK - I - J - K + 2$
(XY, Z)	$(K-1)(IJ-1)$
(XZ, Y)	$(J-1)(IK-1)$
(XY, YZ)	$J(I-1)(K-1)$
(XZ, YZ)	$K(I-1)(J-1)$
(XY, XZ)	$I(J-1)(K-1)$
(XY, XZ, YZ)	$(I-1)(J-1)(K-1)$
(XYZ)	0

(Source: Agresti, 1990, p. 175)

To test goodness-of-fit, G^2 is calculated for each model and tested for significance for the given df. Those that fit adequately at 0.05 level of significance are those with p values >0.05 .

The Newton-Raphson method in an iterative procedure which can be applied to maximum likelihood estimation of logistic regression. The maximum likelihood estimate is the limit of a sequence of iteratively reweighted least square estimates.

6.5 Comparing Models

$G^2(M)$ is the likelihood ratio statistic for testing the fit of model M . Consider two models M_1 and M_2 where M_2 is simpler than M_1 , that is, model M_2 is nested in M_1 with v_1 and v_2 residual degrees of freedom.

Then, $v_1 < v_2$

and $G^2(M_1) \leq G^2(M_2)$

$G^2(M_2 | M_1) = G^2(M_2) - G^2(M_1)$ is large when M_2 fits poorly compared to M_1 ,

$G^2(M_2 | M_1)$ has an asymptotic χ^2 distribution with $df = v_1 - v_2$.

It is important to check the fit according to unconditional tests for accepting a model with a conditional test. As significance may be attributed to sample effects showing chance variation adjusted significance tests are used.

If there are S nested models, there are $S-1$ tests. If each test has level $1-(1-\alpha)^{1/S-1}$ then the asymptotic Type 1 error probability cannot be greater than α . For example if $\alpha = 0.10$ with a 5 nested model then $1-(0.90)^{0.25} = 0.026$. For 1df, critical $\chi^2 = 4.96$. A simpler model is chosen if the difference is not significant between the M_n and M_{n+1} model. The simpler model is compared with the next simplest model for significance, and so on. Of course, the model fit depends on the choice of the nested set.

Fienberg (1994) notes that it is not possible to get negative values for log likelihood ratio statistics using this method unless the error is due to rounding.

It is also noted by Fienberg (1994) that this technique has not made assumptions with respect to ordering of the categories. Taking order into account provides better fitting models.

6.6 Loglinear Models for Ordinal Variables

The previous discussions have only addressed nominal classification of variables and thus has limitations. If an interchange in cell counts in two rows takes place, the fitted values interchange but there is no change in G^2 and χ^2 . The models do not detect trends in the data.

Agresti (1990, p. 262) suggests that models accounting for ordering have the following three advantages:

1. Model parameters describe types of trends and are simpler to interpret than nominal models.
2. Unsaturated models exist in situations in which nominal models are saturated. For instance, there are unsaturated ordinal models for association in two-way tables and for three-factor interaction in three-way tables. The ordinal models have structured association and interaction terms that contain fewer parameters, hence retaining more residual degrees of freedom than the nominal models.
3. Tests based on ordinal models have improved power for detecting certain types of association and interaction.

A two-way table model is represented by Agresti (1990) with scores (u_i) assigned to rows and (v_j) to columns where $u_1 \leq u_2 \leq \dots \leq u_I$, $v_1 \leq v_2 \leq \dots \leq v_J$ with the model:

$$\log m_{ij} = \mu + \lambda_i^X + \lambda_j^Y + \beta u_i v_j.$$

The independence model is the case where $\beta = 0$. The model above has only one more parameter, β , than the independence model with $df = IJ - [1 + (I-1) + (J-1) + 1] = IJ - I - J$.

This model is a special case of the saturated model $\log m_{ij} = \mu + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$.

For identification, a constraint such as $\sum \lambda_i^X = \sum \lambda_j^Y = 0$ is imposed. A possibility in selection of scores is to attempt take approximate distances between midpoints of categories for an assumed underlying interval range. If scores are equal intervals,

$$\begin{aligned} \text{ie. } u_2 - u_1 &= u_3 - u_2 \dots \dots \dots = u_I - u_{I-1} \text{ and } v_2 - v_1 = v_3 - v_2 \dots \dots \dots = v_J - v_{J-1} \\ \text{so } u_i &= i - \frac{i+1}{2} \text{ and } v_j = j - \frac{j+1}{2}. \end{aligned}$$

For the latter scores, $\sum_i \beta u_i v_j = \sum_j \beta u_i v_j = 0$. It is useful on occasions to linearly transform scores to obtain the standardisation

$$\begin{aligned} \sum u_i p_{i+} &= \sum v_j p_{+j} = 0 \\ \sum u_i^2 p_{i+} &= \sum v_j^2 p_{+j} = 1 \end{aligned}$$

The model can be expressed as $\log m_{ij} = \text{independence} + \beta u_i v_j$. If $\sum u_i = \sum v_j = 0$, then, $\beta u_i v_j$ is a deviation of $\log m_{ij}$ from independence. When $\beta > 0$, $\beta u_i v_j$ is positive for values where X is small and Y is small or values when X is large and Y is large. In these cells, m_{ij} is greater than if X and Y were independent. Vice versa, the expected number of observations is smaller for large X and small Y or small X and large Y . So, deviation from independence

gets larger as the four corner cells are approached. Hence, a monotone trend is forced onto the model with direction dependent on the sign of β . The deviation $\beta u_i v_j$ of $\log m_{ij}$ from independence is linear in Y for fixed X and linear in X for fixed Y .

For given rows $h < i$ and columns $j < k$

$$\log \left(\frac{m_{hj} m_{ik}}{m_{hk} m_{ij}} \right) = \beta (u_i - u_h)(v_j - v_k)$$

The absolute value of log odds ratio is greater for rows and columns that are further apart.

If $u_i - u_h = v_k - v_j = 1$ then the odds ratio = e^β .

Unless scores are unevenly spaced, Agresti (1990) suggests that scores be unit spaced so that $\log \theta_{ij} = \beta(u_{i+1} - u_i)(v_{j+1} - v_j)$

where

$$\theta_{ij} = \frac{m_{ij} m_{i+1,j+1}}{m_{i,j+1} m_{i+1,j}}$$

So, all $\log \theta_{ij} = \beta$ and all $\theta_{ij} = e^\beta$

The log likelihood for this model

$$\begin{aligned} L(m) &= \sum \sum n_{ij} \log(m_{ij}) \sum \sum m_{ij} \\ &= n\mu + \sum n_{i+} \lambda_i^x + \sum n_{+j} \lambda_j^y + \beta \sum \sum u_{ij} v_j n_{ij} - \sum \sum \exp[\mu + \lambda_i^x + \lambda_j^y + \beta u_{ij} v_j] \end{aligned}$$

Differentiating with respect to λ_i, λ_j and β and setting partial derivatives to zero we obtain the likelihood equations:

$$\hat{m}_{i+} = n_{i+}, i=1,2,\dots,I$$

$$\hat{m}_{+j} = n_{+j}, j=1,2,\dots,J$$

and $\sum \sum u_{ij} v_j \hat{m}_{ij} = \sum \sum u_{ij} v_j n_{ij}$.

Maximum likelihood estimates can be obtained using the Newton-Raphson method. If $L * L$ is the linear by linear association model and the independence model is I and if $L \times L$ holds then the significance of the association can be tested by $H_0: \beta=0$.

One test statistic is $G^2(I|L*L) = G^2(I) - G^2(L*L)$ with $df = 1$. When the $L * L$ model holds, $G^2(I|L*L)$ is more powerful than the test based on $G^2(I)$. Agresti (1990, pp. 269-70) further discusses models with nominal x or ordinal y .

6.7 Multinomial Response Models

The previous sections have not differentiated between response and variables. This section will discuss, briefly, generalised logit models for nominal response variables.

One of the simpler ways of examining polytomous logistic models is to treat the response as nominal as described by Hosmer and Lemeshow (1989). They described polytomous logistic regression for a three category response variable which can be further generalised for more categories of response. There are, therefore, two logit models comparing $Y = 1$ with $Y = 0$ and $Y = 2$ with $Y = 0$. They suggest fitting a model for $Y = 1$ versus $Y = 0$ using standard logistic regression methods for a binary response. A similar procedure is adopted for $Y = 2$ versus $Y = 0$. Coefficients for the polytomous model are taken from two separately fitted models for $Y = 2$ versus $Y = 0$. Hosmer and Lemeshow state that coefficients obtained from separately fitted models are close to fits from polytomous models.

They suggest for the three category response variable case to assess the fit of the two logistic regression models and integrate them descriptively. Similarly, Fienberg (1994) suggests that for polytomous response variables, if there is an I category response variable, multinomial likelihood functions can be written as a product of $(I-1)$ binomial likelihoods. This allows use of the maximum likelihood methods to estimate parameters. Estimations are done separately for each logit. Individual χ^2 statistics are summed to get goodness-

of-fit for models. The next section will cover, in greater detail, models for ordinal response variables.

6.8 Ordered Polytomous Models

In ordered polytomous models, there are more than two categories for the response variable, say k , ordered categories.

The k -category ordered logistic model is $\log\left(\frac{p_j}{1-p_j}\right) = \alpha_j + \beta'x$ $j=1, \dots, k-1$ where p_j is the probability of being assigned to one of the categories $j+1, j+2 \dots k$.

x is a vector of independent variables (ie. A, B, C ... R), β is a vector of logistic coefficients (not dependent on j).

6.8.1 Baseline-Category Logits

As an example of baseline-category logits, take a three-way contingency table where A and B are explanatory factors and $p_j | h_i$ is the probability of response j , when A is at h and B is at i . A logit model (AB, AY, BY) is built. If the final category is the baseline, the j th logit for this model is

$$\begin{aligned} \log\left(\frac{p_j | h_i}{p_J | h_i}\right) &= \log\left(\frac{m_{hij}}{m_{hiJ}}\right) \\ &= \alpha_j + \beta_{hi}^A + \beta_{ij}^B \quad \text{for } j=1, \dots, J-1 \\ \alpha_j &= (\lambda_j^Y - \lambda_J^Y) \\ \beta_{hi}^A &= (\lambda_{hj}^{AY} - \lambda_{hJ}^{AY}) \\ \beta_{ij}^B &= (\lambda_{ij}^{BY} - \lambda_{iJ}^{BY}) \end{aligned}$$

Logits for other pairs of response categories depend on the above model's parameters; eg, $\log\left(\frac{p_{a|h_i}}{p_{b|h_i}}\right) = \log\left(\frac{p_{a|h_i}}{p_{J|h_i}}\right) - \log\left(\frac{p_{b|h_i}}{p_{J|h_i}}\right)$

As an example, the effect of category h of A or the log odds of classification in category a of y instead of b is $\beta_{ha}^A - \beta_{hb}^A$.

To identify the model, there are the following constraints

$$\sum_h \beta_{hj}^A = 0 \text{ and } \sum_i \beta_{ij}^B = 0$$

Fitting generalised logit models can be done simultaneously for the J-1 equations using iterative procedures or can be fitted separately for the J-1 pairs. Fitting of ordinal explanatory variables can be done in the same way as for ordinal loglinear models. However for continuous explanatory variables, the Newton-Raphson iterative procedure give estimates for continuous explanatory variables. However, this can be time consuming given sparse data. χ^2 goodness-of-fit tests are adversely affected and are, therefore, not appropriate for model fit testing. However, χ^2 goodness-of-fit can be used to contrast models.

6.8.2 Logits for Ordinal Responses

Three types of logits are discussed:

Adjacent-categories logits

Let $\{p_i(x) \dots p_J(x)\}$ denote response probabilities at value x for a set of response variables. The logits, L_j , for adjacent categories logits are as follows:

$$L_j = \log\left(\frac{p_j(x)}{p_{j+1}(x)}\right) \quad j=1,2,\dots,J-1$$

They are equivalent to baseline logits:

$$L_j^* = \log\left(\frac{p_j(x)}{p_J(x)}\right) \quad j=1,2,\dots,J-1$$

Let the model for the adjacent categories logit model be

$$L_j = \alpha_j + \beta'x \quad j=1,2,\dots,J-1$$

Now
$$L_j^* = L_j + L_{j+1} + \dots + L_{J-1}$$

$$L_j^* = \sum_{k=j}^{J-1} \alpha_k + \beta'(J-j)x \quad j=1,2,\dots,J-1$$

$$= \alpha_j^* + \beta'u_j \quad j=1,2,\dots,J-1$$

where
$$u = (J-j)x$$

Continuous-Ratio Logits

Let the continuous-ratio logits or L_j be

$$L_j = \log\left(\frac{p_{j+1}(x)}{p_{j+1}(x) + \dots + p_J(x)}\right) \quad j=1,2,\dots,J-1$$

$$\text{and } L_j^* = \log\left(\frac{p_{j+1}(x)}{p_1(x) + \dots + p_j(x)}\right) \quad j=1,2,\dots,J-1$$

if $p_j(x)$ denotes the probability of response j , given response j or higher then

$$\rho_j(x) = \frac{p_j(x)}{p_j(x) + \dots + p_J(x)} \quad j=1,2,\dots,J-1$$

$$L_j = \log\left(\frac{\rho_j(x)}{1 - \rho_j(x)}\right) \quad j=1,2,\dots,J-1$$

At a given x , $n_j(x)$ (for $j=1,\dots,J$) are the response counts and

$$n(x) = \sum_j n_j(x)$$

The multinomial mass function for $n_j(x)$, $j=1,\dots,J$ is factorised to

$$b[n(x), n_1(x); p_1(x)] b[n(x) - n_1(x), n_2(x); p_2(x)] \dots \dots \dots$$

$$b[n(x) - n_1(x) - \dots \dots n_{j-2}(x), n_{j-1}(x); p_{j-1}(x)]$$

where $b(n, y; p)$ is the binomial probability of getting y "successes" in n trials when the probability of succession for each trial is p . Each log likelihood is maximised separately and the separate $(J-1)$ G^2 statistics can be aggregated to give an overall goodness-of-fit. Separate fitting using log-linear models can be used for these logits because they are binary responses.

Cumulative Logits

A third way of handling ordered response categories is to form logits of cumulative probabilities;

$$F_j(x) = p_1(x) + p_2(x) + \dots + p_j(x) \quad j=1, \dots, J$$

with cumulative logits

$$L_j = \text{logit}[F_j(x)] = \log\left(\frac{F_j(x)}{1 - F_j(x)}\right)$$

$$= \log\left(\frac{p_1(x) + \dots + p_j(x)}{p_{j+1}(x) + \dots + p_J(x)}\right) \quad j=1, \dots, J-1$$

So each cumulative logit has all J response categories with i to j forming one category and $j+1$ to J forming another category. There are some cumulative logit models that are not equivalent to binary logit, multinomial or loglinear models. Discussion of the proportional odds model follow.

6.9 Proportional Odds Model

Let $L_j(x) = \log[F_j(x)]$, $j=1, \dots, J-1$ where $F_j(x) = P(Y \leq j | x)$ is the cumulative probability for the response category j and x is the vector of explanatory variables. The simplest cumulative logit model is

$$L_j(x) = \alpha_j \quad j=1, \dots, J-1 \text{ which is independent of } x.$$

The α_j 's are non-decreasing in j and called cutpoints. Including the effects of explanatory variables

$$L_j(x) = \alpha_j + \beta'x \quad j=1, \dots, J-1.$$

The model assumes a variable's effect on the odds of response below category j is constant for all j . At $x=x_1$ and $x=x_2$,

$$\begin{aligned} \text{So } L_j(x_1) - L_j(x_2) &= \log \left[\frac{P(Y \leq j | x_1) | P(Y > j | x_1)}{P(Y \leq j | x_2) | P(Y > j | x_2)} \right] \\ &= \beta'(x_1 - x_2) \end{aligned}$$

The odds ratio of the cumulative probabilities is a cumulative odds ratio. The log of the cumulative odds ratio is proportional to the distance between values of the explanatory variables. This proportion applies to each cutpoint; that is, a proportional odds model. Thus the odds of making response $\leq j$ are $\exp[\beta'(x_1 - x_2)]$ times higher at $x=x_1$ than at $x=x_2$. Explanatory variables in the cumulative logit model may be categorical or continuous. For the single explanatory variable case it can be shown that $F_j(x) = F_k[x + \frac{(\alpha_j - \alpha_k)}{\beta}]$. When $\beta_i > 0$ the cumulative probabilities increase as α_i increases so Y tends to be smaller at greater values of x_i . If $\beta_i > 0$ is to imply the usual meaning of Y increasing when x_i is increasing then $L_j(x) = \alpha_j - \beta'x$. The multinomial likelihood is

maximised given the $L_j(x) = \alpha_j - \beta'x$ constraint. The fitting of the model uses Fisher algorithms for iterative calculations of the maximum likelihood estimates of parameters.

The proportional odds model is discussed in papers by Snell (1964), Walker and Duncan (1967) and Williams and Grizzle (1972). The model is described in detail in a classic paper by McCullagh (1980) which promoted reviews of the paper by a number of statisticians. McCullagh (1980) gives a simple example of the use of the model for tonsil size of carriers and non-carriers of *Streptococcus pyogenes*. The use of proportional odds models concludes that tonsil size does tend to be larger in the carrier than in the non-carrier. This conclusion would have also been made using the Wilcoxon test and tests based on partitioning Pearson's or the likelihood ratio χ^2 statistics. The quantitative conclusion that the odds of greatly enlarged tonsils is 1.8 times greater in the carrier than the non-carrier can be arrived at in the parametric model. The paper states that the advantage of this approach will be best demonstrated in the use of the model on more complex examples. The data from the coronary artery bypass graft patients will provide such a demonstration later in this chapter.

In discussion of McCullagh's paper, Dr Daryl Pregibon (McCullagh, 1980, p. 139), from Princeton University, expresses his disappointment that 'no real regression example was considered'. He further states that he has data on post-operative status of 65 open-heart surgery patients with response categorised as no complications, ischaemia or death due to infarct along with a number of explanatory variables. The response variable is an ordinal outcome measure. Dr Pregibon had some reservations about the usefulness of

the model for these data and believed that the deviance contributions were better candidates for residuals than individual cell differences.

Another example of the use of this model on a database similar to this study's database is given by Agresti (1990, p. 325) where mental impairment in four categories (well, mild, moderate or impaired) is modelled on two explanatory variables. The explanatory variables are a binary socio-economic status measure and an index of life events. Agresti (1990) proposed the use of the SAS LOGIST procedure to analyse these data. The LOGIST procedure uses maximum likelihood fitting of cumulative logit models.

The first section of this chapter has provided a broad review of fitting loglinear and logit models, loglinear-logit models for ordinal variables and multinomial response models as background theory for the model fitting proposed for the coronary artery bypass graft patient data. The next section discusses, more specifically, the statistical procedure in SAS LOGIST, and its applications that will be used for these data.

6.10 The SAS LOGIST Procedure and Applications

The abstract for the SAS LOGIST procedure (SAS Institute Inc, 1986) states:

"The LOGIST procedure fits the logistic multiple regression model to a single binary (0-1) dependent variable or to an ordinal dependent variable. The procedure can fit one model or use either backward elimination procedure or a stepwise technique based on a strategy that with very few calculations determines the best variable to be added to the model in any given step. Maximum likelihood estimates (MLEs) are computed by the modified Gauss-Newton method (with step halving). Singularities (redundant variables) are allowed when fitting a model or when testing covariates not in the model. The procedure computes test statistics for assessing lack of fit of the model. LOGIST performs weighted and unweighted analyses."

In other words, LOGIST is able to fit the proportional odds model described in the previous section. If the NO STEPWISE option is used in LOGIST, the procedure calculates MLEs for the parameters associated with each explanatory variable. The MLE χ^2 statistic (Wald statistic) tests the hypothesis that a parameter is zero by computing parameter estimates divided by the

standard error and squaring. This test assumes that the estimates are asymptotically normally distributed. The default level of significance is 0.05.

Two examples of the application of this procedure to proportional odds models are given by Koch (*et al.*, 1985). The first model looks at an ordered response variable for pain relief in five categories - poor, fair, moderate, good and excellent. The second has subjects in different risk groups classified according to increasing severity of some health conditions. In the second example, an ordinal response variable for dumping syndrome following treatment for duodenal ulcer is given in three categories - none, slight or moderate. Operation type is the explanatory variable.

Three methods for fitting are used with their respective parameter estimates compared. The methods are:

- FARM- functional asymptotic regression methodology;
- SAS GENCAT for weighted least squares estimates; and
- SAS LOGIST for maximum likelihood estimates.

FARM is a two-stage procedure where firstly separate maximum likelihood estimates for the elements of the vector β are made with consistent estimates of their covariance matrix. A modified version of SAS CAT MAX was used for this stage. Secondly, reduced models, such as proportional odds models, are fitted to β via weighted least squares methods. The SAS GENSTAT procedure was used for this stage of FARM. A comparison of the three methods is given in Table 6.5.

Table 6.5: Final Model Estimated Parameters and Estimated Standard Errors for the Dumping Syndrome Data

Final Model (Proportional Odds)			
Parameters	Estimates (Standard Errors)		
	FARM	WLS	ML
Intercept for none vs at least slightly (γ_1)	0.659 (0.173)	0.657 (0.172)	0.657 (0.173)
Intercept for at most slight vs moderate (γ_2)	2.413 (0.215)	2.368 (0.214)	2.411 (0.215)
Common slope (γ_3)	-0.226 (0.088)	-0.222 (0.088)	-0.225 (0.088)

(Source: Koch *et al.*, 1985, p. 367)

Further examples in Koch (*et al.*, 1985) demonstrated similar results for the estimates using the three methods.

Examples of the use of ordered polytomous regression have been given by a number of authors, such as Koch (*et al.*, 1985), McCullagh and Nelder (1989), Agresti (1990), Anderson (1984), Anderson and Philips (1981), Peterson and Harrell (1990) and McCullagh (1980). The data sets in each of these studies have been limited to relatively small numbers.

The most relevant use of the LOGIST procedure for the purposes of this research is given by Ashby (*et al.*, 1986), because their study analyses data on alcohol consumption, serum biochemistry and haematology for 7735 middle-

aged British men. The data were collected during screening procedures in 1978 to 1980. Information on frequency of consumption, quantity and alcohol type were elicited through a questionnaire administered by a nurse. There were eight categories of alcohol consumption ordered from non-drinkers to daily drinkers consuming more than the equivalent of 6 units where a unit was half a pint of beer, a glass of wine or a tot of spirits. Table 6.6 gives the distribution of responses and inputted 'average' weekly units of alcohol.

Table 6.6: Ordered Drinking Categories from the Study by Ashby and Others

Category	Inputted "average" weekly units	Number	Percentage
non-drinkers (N)	0	466	6%
occasional drinkers (O)	1	1845	24%
weekend drinkers(1-2units/day) (W1)	3	725	9%
weekend drinkers(3-6units/day) (W2)	9	1234	16%
daily drinkers(1-2units/day) (D1)	11	585	8%
weekend drinkers (>6units/day) (W3)	16	1095	14%
daily drinkers(3-6units/day) (D2)	30	947	12%
daily drinkers(>6units/day) (D3)	56	832	11%
		-----	-----
	Total	7729	100%

(Source: Ashby *et al.*, 1986,p. 290)

Twenty five biochemical and haematological measures were taken from blood samples obtained at the initial screening. Interrelations between the measures were illustrated using a minimum spanning on the correlation matrix. Six subsets of variables were identified. Rather than estimating the

effect of alcohol consumption on each biochemical and haematological variable, Ashby (et al., 1986) estimated the effect of combinations of biochemical variables on alcohol consumption using an ordered polytomous logistic model. The authors recommended a forward stepwise approach to variable selection using the SAS LOGIST procedure.

The results of the procedure would have included 18 variables in the model if the conventional 5% level of significance was applied. Notwithstanding this, the authors found that the main changes in the log likelihood came from the first five variables.

The authors checked the goodness-of-fit by extending the two category model. The probabilities of belonging to each category for a given respondent are:

$$p(j) = \begin{cases} 1-p_j & j=1 \\ p_{j-1}-p_j & j=2\dots k \end{cases}$$

Observed and expected numbers in each category are compared with m equally spaced intervals on a logit scale. The expected number in category j is $E_j = \sum p(j)$ added over all respondents in that interval. Within each of the m logit intervals $\sum E_j = \sum O_j$. For a model with r explanatory variables, $\sum \frac{(O-E)^2}{e}$ summed over k categories and m logit intervals follows a χ^2 distribution with $mk - (k-1+r) - m$ degrees of freedom for a model with r explanatory variables. Allowing for some aggregation of cells because expected frequencies were less than 5 the overall χ^2 had 52 degrees of freedom and was equal to 112.57 with $p\text{-value} < .001$. This method of goodness-of-fit showed the greatest lack of fit was in the lowest logit interval. In this interval, more

regular drinkers had lower levels of serum biochemistry and haematology than expected from the model. Vice versa, there were fewer non-drinkers with lower values of serum biochemistry and haematology than predicted from the model. The authors suggest that this might be explained by ex-drinkers no longer consuming alcohol because of health reasons.

Ashby (et al., 1986) compared three methods for allocating individuals to categories. The first - the highest probability method, proposed by Anderson and Philips (1981) - used $\hat{\alpha}$ and $\hat{\beta}$ in estimates of the posterior probability. The category with the largest estimated posterior probability is taken as the predictor. The second method - the inter category allocation method, also proposed by Anderson and Philips (1986) - allocated individuals to category j if $\hat{\alpha}_{j-1} \leq \hat{\beta}'x \leq \hat{\alpha}_j$. The final method is the proportional allocation method where allocations are made using $\hat{\beta}'x$ with cutoffs chosen to force numbers into each category equal to the observed numbers. Ashby (et al., 1986) compared the three methods for the data on the 7729 British men and concluded that the proportional allocation methods was the most successful, identifying 57% in the top two categories and only 8% not drinking regularly.

Finally, the authors attempted to simplify the model by amalgamation of adjacent categories. The optimal amalgamation was found by minimising the pooled-within-category mean square for $\hat{\beta}'x$.

Ashby and her colleagues have demonstrated that the SAS LOGIST procedure can be applied to large data sets and the results can be easily interpreted. The next section describes the application of this procedure to the data on the 3979 coronary artery bypass graft patients.

6.11 Application of Ordered Polytomous Regression Models to Coronary Artery Bypass Graft Patient Data

Ordered polytomous regression modelling is applied to categorical response variables NSUM and PSUM - the respective measures of neurological and physical deficit for the coronary artery bypass graft patients. The fitting of the model to NSUM is discussed first, followed by fitting of the model to the response variable PSUM. The final section of the chapter briefly addresses the implications of the models to maximising health outcomes.

6.10.1 Model for NSUM

NSUM has been allocated to five categories as illustrated in Table 6.7. The selection of categories was made after consultation and discussion with the perfusionist about appropriate re-classification of the 0 to 15 scale response variable at the study hospital. The categories give the following levels of neurological deficit - none ($\text{NSUM}=0$), mild ($1 \leq \text{NSUM} \leq 2$), moderate ($3 \leq \text{NSUM} \leq 4$), substantial ($5 \leq \text{NSUM} \leq 6$) and severe ($\text{NSUM} \geq 7$).

Table 6.7: Distribution of NSUM Categories

NSUM Deficit	Frequency	Proportion
None	2768	69.5%
Mild	616	15.5%
Moderate	313	7.9%
Substantial	186	4.7%
Severe	96	2.4%
	-----	-----
Total	3979	100%

Seventeen explanatory variables were fitted to the model for NSUM. These explanatory variables are the same as those that have been described in detail and analysed in Chapter 5. These are listed and described, briefly, in Table 6.8.

Table 6.8: Description of Explanatory Variables

Explanatory Variable (code)	Description
Body mass index (BMI)	patient's $\frac{(\text{weight in kgs})}{(\text{height in metres})^2}$
Cross clamp time (XCLAMP)	the time in minutes that the cross clamp was in place during surgery.
Time on bypass (BYPASS)	grouped into the following 30 minute intervals as follows:- less than 60 , 60 to 89 , 90 to 119 , 120 to 149 , 150 to 179, 180 to 239 and 240 or more.
Age(AGE)	grouped into 5 year cohorts as follows in years:- less than 40 , 40 to 44 , 45 to 49 , 50 to 54 , 55 to 59 , 60 to 64 , 65 to 69, 70 to 74, 75 to 79 and 80 or more .
Gender (SEX)	Binary variable for gender (male/female).
marital status (MARRIED)	(yes/no)
Retirement status (RETIRED)	(yes/no)
Angina (ANGINA)	(present/not present).
Previous surgery (OPS)	a nominal variable with categories for cholecystectomy, transurinary resection procedure, vascular surgery, appendectomy, open heart surgery, hysterectomy or no surgery.
Previous symptoms (SYMPT)	a nominal variable with categories for acutemyocardial infarction, diabetes, shortness of breath, stroke, angina and hypertension.
Year of Surgery (YEAR)	a two digit variable for year of surgery
Socio-Economic Status (SEI)	based on the index of Education and Occupation assigned according to the patient's postcode. The index has a mean of 1000 with a standard deviation of 100.
pH (PH)	as described in Chapter 5, a binary variable for whether patients were in or out of valleys with lower mean NSUM scores for given pH as illustrated in Figure 5.5.
Partial Oxygen (PO)	as above partial oxygen levels (Figure 5.6)
Partial Carbon Dioxide (PCO)	as above for partial carbon dioxide levels (Figure 5.7).
haemoglobin (HB)	as above for haemoglobin (Figure 5.9).
base excess (BE)	as above for base excess (Figure 5.10)

A correlation matrix for these 17 variables is given in Table 6.8.

Table 6.9: Pearson Correlation Matrix for Explanatory and Response Variables

	SEI	RETIRED	MARRIED	NSUM	PSUM	AGE	SEX	BMI	PH	
SEI	1.0000	.0010	.0542	-.0137	-.0250	.0886	.0108	-.1038	-.0162	
RETIRED		1.0000	.1087	.0949	-.0488	-.5527	.0566	.0430	.0191	
MARRIED			1.0000	.0427	-.0283	-.0478	.0402	.0161	.0092	
NSUM				1.0000	.1015	-.1140	.0147	.0168	.0252	
PSUM					1.0000	.0419	-.0411	.0040	.0104	
AGE						1.0000	-.1202	-.1403	-.0609	
SEX							1.0000	.1062	.0200	
BMI								1.0000	.0284	
PH									1.0000	
	BE	PCO	PO	HB	ANGINA	YEAR	XCLAMP	BYPASS	SYMPT	OPS
SEI	.0443	-.0221	.0059	-.0271	-.0152	-.0313	-.0371	-.0263	-.0336	.0118
RETIRED	.0840	.0147	.1018	.0943	-.0007	-.0367	-.0295	-.0357	-.0863	-.0350
MARRIED	.0682	.0261	-.0126	-.0395	-.0090	-.0097	.0454	.0411	-.0277	-.0574
NSUM	.0612	.0199	.1298	.0336	.0926	-.2088	-.0016	.0503	-.1074	-.0219
PSUM	.0662	.0180	.0327	-.0094	.5722	-.0367	.0163	.0436	-.0383	.0064
AGE	-.0031	-.0068	-.1126	-.0958	-.0222	.0765	.0281	.0621	.0777	-.0001
SEX	-.0676	.0049	.1034	.1196	-.0442	-.0581	.0838	.0872	-.0492	.0548
BMI	-.0308	-.0086	.0616	.0926	-.0121	.0238	-.0008	.0091	.0468	-.0572
PH	.1258	.0416	.0468	.0263	.0317	-.0239	-.0191	-.0172	-.0321	-.0018
	BE	PCO	PO	HB	ANGINA	YEAR	XCLAMP	BYPASS	SYMPT	OPS
BE	1.0000	-.0161	.1105	.0089	.0778	-.1939	.1509	.2267	-.0837	.0550
PCO		1.000	-.0274	.0248	.0181	-.0105	.0600	.0547	.0044	.0158
PO			1.000	.0520	.0470	-.4693	-.1114	-.0379	-.2634	.0030
HB				1.000	-.0128	-.0248	-.0440	-.0413	.0141	-.0223
ANGINA					1.000	-.0626	-.0109	-.0011	-.0550	-.0039
YEAR						1.000	-.0119	-.1588	.4012	-.0722
XCLAMP							1.000	.7413	-.0032	.0640
BYPASS								1.000	-.0774	.0906
SYMPT									1.000	-.0231
OPS										1.000

The SAS LOGIST procedure was used to model NSUM on these variables with interaction terms for age and sex; in or out of the base excess valley with the variable for in or out of the pH valley, the partial oxygen valley and the haemoglobin valley; and age and year of surgery. The SAS commands for this model are presented in Appendix F. It should be noted that the number of observations in this model has been greatly reduced because information on previous operations, retirement status and marital status was only recorded after 1990. Thus, there were only 529 observations. The results and the model using these explanatory variables are given in Appendix G. Only two explanatory variables are included in this model - age and an interaction term for age and year of surgery.

As previous operations, marital status and retirement status were not included in the model in Appendix G, they were excluded as possible explanatory variables in a subsequent run of the LOGIST procedure. The commands for the fitting of this model are given in Appendix H. In this case the number of observations was increased to 2555. The model for this procedure is presented in Appendix I.

A summary of the fitted model is presented in Table 6.10. In the table $\hat{\beta}$ is the estimate of the parameter for the given explanatory variable and $SE(\hat{\beta})$ is the standard error of the estimate.

Table 6.10: Maximum Likelihood Estimates for Response NSUM with a Five Category Ordered Logistic Model and Five Explanatory Variables

Slope parameters β				Intercepts		
Variable	$\hat{\beta}$	SE ($\hat{\beta}$)	t value	Category j	$\hat{\alpha}$	SE($\hat{\alpha}$)
presence of angina	-.6221	.1470	-4.23	1	-11.50	2.16
				2	-10.54	2.16
				3	-9.67	2.16
in PO ₂ valley	-.4770	.1520	-3.14	4	-8.50	2.16
age*year of surgery	.0018	.0003	6.07			
year of surgery	.1429	.0228	6.27			
time on bypass	-.1282	.0430	-2.98			

6.10.2 Goodness-of-Fit of Model for Response NSUM

The model was checked for goodness-of-fit using the method described by Ashby (et al., 1986). The respondents were grouped in six equally spaced logits. The expected numbers in each category j was calculated as $E_j = \sum p(j)$ summed over all respondents in the interval. The $p(j)$ was calculated as follows:

$$p(j) = \begin{cases} 1-p_j & j=1 \\ p_{j-1}-p_j & j=2\dots k \end{cases}$$

The statistic $\sum \frac{(O-E)^2}{E}$ summed over the k categories and m logit intervals follows a χ^2 with $(m-1)(k-1)-r$ degrees of freedom. Where r is the number of explanatory variables in the model.

For this model, the logit scale was defined as $-10.5425 + \beta'x$ divided into six equal logit intervals of 0.5. Table 6.11 gives observed and expected numbers (in parentheses) of patients in each of the five NSUM categories within the six logit intervals. Each logit interval's contribution to the total χ^2 is included. The total χ^2 has 15 degrees of freedom. Three cells were combined because their expected frequencies were less than five. For the whole table $\chi^2 = 16.35$ has 12 degrees of freedom ($p > .05$). This indicates a good fit for the model.

Table 6.11: Goodness-of-fit for Five Category Model - Observed and Expected Frequencies for Respondents in each NSUM Category for each Logit Interval

Logit Interval	Total	NSUM Categories: Observed (Expected)					χ^2
		none	mild	moderate	substantial	severe	
≤ 1.0	145	72(65.99)	27(33.72)	22(21.73)	14(15.25)	10(8.31)	2.34
≤ 1.5	509	293(295.43)	111(104.44)	56(56.19)	30(35.13)	19(17.81)	1.26
≤ 2.0	795	540(548.88)	132(131.08)	72(61.70)	41(35.88)	10(17.46)	5.79
≤ 2.5	586	452(458.9)	74(71.74)	31(30.50)	18(16.87)	11(7.99)	1.39
≤ 3.0	379	333(323.78)	26(32.35)	8(12.81)	{7(6.86)}	{5(3.2)}	3.69
> 3.0	141	132(127.33)	6(8.19)	{2(3.10)}	1(1.63)	0(0.75)}	1.88
Total	2555	1822	376	191	111	55	16.35

The results of the LOGIST procedure shows that the estimated odds of being in lower NSUM category if angina is present as opposed to angina not being present is $\exp(-0.6221) = 0.537$ - that is, the estimated odds of being in a higher category of NSUM if angina is present is 1.86. The estimated odds of being in a lower NSUM category if the patient was not in the valley for partial oxygen that minimises mean NSUM is $\exp(-0.4777) = 0.62$. This means that if a

respondent is not in the partial oxygen valley, the estimated odds of being in a higher NSUM category is 1.61. The interaction between age and year of surgery shows that for a higher age category and an increase of one in the year of surgery, the estimated odds of being in a lower NSUM category is 1.002. For an increase in one year in year of surgery, the estimated odds of being in a lower NSUM category is 1.15. The estimated odds of being in a lower NSUM category for an increase of 30 minutes on bypass is 0.88 or the estimated odds of being in a higher category of NSUM for 30 minute increase in time on bypass is 1.14.

In summary, the variables that have association with increasing neurological deficit are the presence of angina one year post-surgery, being outside the partial oxygen valley that minimises mean NSUM, an interaction between being younger and having surgery in earlier years, earlier year of surgery and a longer time on bypass.

6.10.3 Model for PSUM

As with the allocation method for categories for NSUM, PSUM was also allocated to five categories. Again, the selection of categories was made after consultation and discussion with the perfusionist about an appropriate re-classification of the 0 to 15 scale response variable. The categories give the following levels of physical deficit - none ($PSUM=0$), mild ($1 \leq PSUM \leq 2$), moderate ($3 \leq PSUM \leq 4$), substantial ($5 \leq PSUM \leq 6$) and severe ($PSUM \geq 7$).

Table 6.12: Distribution of PSUM Categories

PSUM Deficit	Frequency	Proportion
None	2888	72.6%
Mild	531	13.3%
Moderate	311	7.8%
Substantial	150	3.8%
Severe	99	2.5%
Total	3979	100%

The SAS LOGIST procedure was used to fit explanatory variables to PSUM in the above categories. The explanatory variables were as for Appendix F - with the exception of angina as this variable was derived from the PSUM score. The NSUM value in five categories was included as an explanatory variable. The commands for this procedure are presented in Appendix J. The results of fitting the model are in Appendix K. Only the explanatory variables for gender and socio-economic index were included in the model. Again, the number of observations in this model were greatly reduced because marital and retirement status and previous operations data were only recorded after 1990. These variables were excluded in the commands given in Appendix I. The results of this analyses with 2555 observations is given in Appendix M. Table 6.13 presents the fitted model.

Table 6.13: Maximum Likelihood Estimates for Response PSUM with a Five Category Ordered Logistic Model and Six Explanatory Variables

Variable	Slope parameters β			Category j	Intercepts	
	$\hat{\beta}$	SE ($\hat{\beta}$)	t value		$\hat{\alpha}$	SE($\hat{\alpha}$)
age	-.0966	.0276	-3.50	1	.077	.596
sex	.2871	.1071	2.68	2	.925	.596
occ. & ed. status	.0014	.0005	2.80	3	1.786	.599
NSUM	-.2337	.0431	-5.42	4	2.76	.606
previous symptoms	.0537	.0202	2.66			
in base excess valley	-.3773	.1175	-3.21			

6.10.4 Goodness-of-Fit of Model for Response PSUM

The methods used to test goodness-of-fit described for the NSUM response model were applied to the results for PSUM as a response variable. The logit scale used to test goodness-of-fit was defined as $1.7859 + \beta'x$ divided into six equal intervals of 0.25. Table 6.14 gives the observed and expected counts for respondents for each PSUM category within each of the six logit intervals.

Table 6.14: Goodness-of-fit for Five Category Model - Observed and Expected Frequencies for Respondents in each PSUM Category for each Logit Interval

Logit Interval	Total	PSUM Categories: Observed (Expected)					χ^2
		none	mild	moderate	substantial	severe	
≤ 2.15	202	111(112.84)	36(37.79)	34(25.78)	15(15.04)	6(10.55)	4.70
≤ 1.5	279	175(178.60)	47(46.24)	24(28.36)	26(15.44)	7(10.36)	9.07
≤ 2.0	505	352(350.66)	75(74.21)	44(42.82)	19(22.51)	15(14.80)	0.60
≤ 2.5	701	533(522.04)	81(89.19)	50(48.78)	19(24.89)	18(16.10)	2.63
≤ 3.0	635	501(510.86)	67(69.51)	34(36.44)	16(18.19)	17(11.62)	3.20
> 3.0	233	184(192.82)	31(21.08)	7(10.62)	7(5.20)	4(3.28)	7.05
Total	2555	1856	337	193	102	62	27.24

The χ^2 has $(m-1)(k-1)-r$ degrees of freedom where m is the number of logit intervals, k is the number of response categories and r is the number of explanatory variables in the model. So, χ^2 has 14 degrees of freedom with the loss of one degree of freedom because one cell was combined with another as the expected frequency was less than five. For the whole table, $\chi^2=27.24$ with 13 degrees of freedom ($p<.025$). This indicates a deviation from the model. The logit interval with the most obvious lack of fit is the second lowest where more respondents were observed in the substantial physical deficit category than were expected for the model. A lack of fit in another logit interval, not to the same extent, was apparent in the highest logit interval, where more respondents were observed in the mild physical deficit category than were expected from the model.

The results of the LOGIST procedure show that the estimated odds of being in a lower PSUM category for an increase in age of five years is

$\exp(0.0966)=0.908$. In other words, the estimated odds of being in a higher PSUM category given an increase in age of five years is 1.10. The estimated odds of being in a lower PSUM category if the respondent was male rather than female was $\exp(0.2871)=1.33$. The estimated odds of being in a lower PSUM category if a respondent is not in the valley that minimises mean NSUM for given base excess levels is $\exp(-0.3773)=0.686$. Therefore, the estimated odds of being in a higher PSUM category if the respondent is not in the base excess valley is 1.46. An increase of 1 in the socio-economic index, gives an estimated odds of being in a lower PSUM category of $\exp(0.00141)=1.001$. This can be described as the estimated odds of being in a lower PSUM category is approximately 10 for an increase of 10 in the index. The estimated odds of being in a lower PSUM category given an increase in the NSUM category is $\exp(-0.2337)=0.79$ - that is, for an increase in the NSUM category, the estimated odds of an increase in the PSUM category is 1.26. Finally, for a change in the previous symptoms category, the estimated odds of a lower PSUM category is $\exp(0.0537)=1.06$.

In summary, increasing physical deficit is associated with increasing NSUM, increasing age, lower socio-economic status, the nature of previous symptoms, being outside the base excess valley that minimises mean NSUM and being female.

6.12 Discussion of both Models

Coronary artery bypass graft patients would wish to minimise both neurological and physical deficits following surgery. The fitting of separate models to response variables PSUM and NSUM show that there is some

positive association between levels of neurological deficit and physical deficit. However, the model for neurological deficit showed the presence of angina rather than the overall physical deficit category was a better indicator of change in neurological deficit. In terms of variables common to both models, increasing age was associated positively with physical deficit whilst an interaction between increasing age and increasing year of surgery was associated negatively with neurological deficit. An important finding was that socio-economic status appears to have an impact on physical outcome but not on neurological outcome. There was a negative relationship between year of surgery and neurological deficit. This demonstrates a reduction in neurological deficit corresponding to the perfusionist's implementing continuous quality improvement models over those years to improve neurological outcomes. This improvement over time was not evident for physical outcome measure. Quality improvement techniques were not used by the perfusionist to improve physical outcomes in the study period. However, quality improvement techniques to minimise physical deficit have been used since 1993 by the perfusionist and its impact can be tested when the data become available.

6.13 Summary

Chapter 5 indicated, through the use of χ^2 and t-tests, explanatory variables that had significant association with neurological and physical outcome. This chapter has discussed the shortcomings of this univariate analyses and demonstrated that when all explanatory variables are controlled for, using multivariate regression techniques, not all variables with significant associations with the response variables in Chapter 5 are included in the final

models. The technique used to analyse these data was the SAS LOGIST procedure. This procedure uses maximum likelihood estimation to provide a model based on polytomous logistic regression methods. The theory and applications of this technique have been described in detail prior to the application of the procedure to the research data. The discussion of the theory and application of this procedure was extensive because the procedure had been largely untested on large data sets with the exception of the study by Ashby *(et al., 1986)*.

Applying this technique to the response variable NSUM showed that the variables that have association with increasing neurological deficit are the presence of angina one year post-surgery, being outside the partial oxygen valley that minimises mean NSUM, an interaction between being younger and having surgery in earlier years, earlier year of surgery and a longer time on bypass. The application of the model to the response variable PSUM demonstrated increasing physical deficit is associated with increasing NSUM, increasing age, lower socio-economic status, the nature of previous symptoms, being outside the base excess valley that minimises mean NSUM and being female. The results of fitting models to response variables can be incorporated into the use of continuous quality improvement models described in Chapter 7.

Chapter 7 will now describe the continuous quality improvement model used by the perfusionist to improve neurological outcome and it further provides a more general model that can be applied to high cost/high volume surgical procedures.

CHAPTER 7

A MODEL FOR CONTINUOUS QUALITY IMPROVEMENT

CHAPTER 7 A MODEL FOR CONTINUOUS QUALITY IMPROVEMENT

7.1 Introduction

The previous chapter has reported explanatory variables included in two separate models to explain variations in both NSUM and PSUM. The identification of the variables that explain variation in ordered categorical measures of health outcome is only of intellectual curiosity unless it is used in some way to improve those health outcomes. Box (1993, p. 3) in discussing statistical methods and their effect on quality states that the impact 'depends on the system of management'. This chapter describes 'a system of management' where process variables that have significant statistical associations with health outcomes can be used in a model to improve those outcomes that is, a continuous quality improvement model.

The objective of this chapter is to describe a total quality management (TQM) model for surgical procedures. Of course, it is acknowledged that there has been no attempt to cover all aspects of care in an inpatient hospital episode - only those aspects specific to the surgical procedure have been considered. The model can be modified for differences in both hospitals and their staff and in differences within surgical procedures.

This chapter, first, examines literature relating to the introduction of TQM to industry and its associated models for continuous quality improvement (Section 7.2). The fundamental basis for TQM, in any area, is derived from the work of Shewart (Deming, 1986), and Deming and Ishawaka (Walton,

1989) After the broad-ranging literature review presented in Chapter 2, selected examples of continuous quality improvement models used in health care that are relevant to the model to be developed are presented in Section 7.3. The principles for the design and use of models are introduced in Section 7.4 to assist in development of the patient-outcome model proposed by this study. Finally, as the model proposes the use of computer-based patient records, a brief review of recent literature on aspects of computer-based patient records is undertaken. Finally, the development of a specific model for continuous quality improvement for surgical procedures is detailed. The model described involves a cycle of pre-surgery, surgical procedures, post-surgery, follow-up action, data analyses and recommendations to change in practice. This conceptual model is illustrated both with examples from Chapter 4 and with hypothetical results to explain, in particular, reporting systems.

7.2 TQM in Industry

Kaluzny (et al., 1992, p.257) gives a definition of (TQM) as:

'a participative, systematic approach to planning and implementing a continuous organisational improvement process. Its approach is focused on satisfying customers' expectations, identifying problems, building commitment and promoting open decision-making among workers'.

TQM as a tool in industry dates back to the 1920's use by Walter A. Shewart and more recently to the work of W. Edwards Deming in Japan in the 1950's (Walton, 1989). Shewart was a statistician at the Bell Telephone Laboratories

in New York and was known for bringing industrial processes into statistical control by defining limits for random variation in workers' tasks and then studying the cause of outliers. Thus, finding the cause of the outliers (that is, an outcome measure) he improved the process through a continuous cycle, known as the Shewart Cycle illustrated below in Figure 7.1.

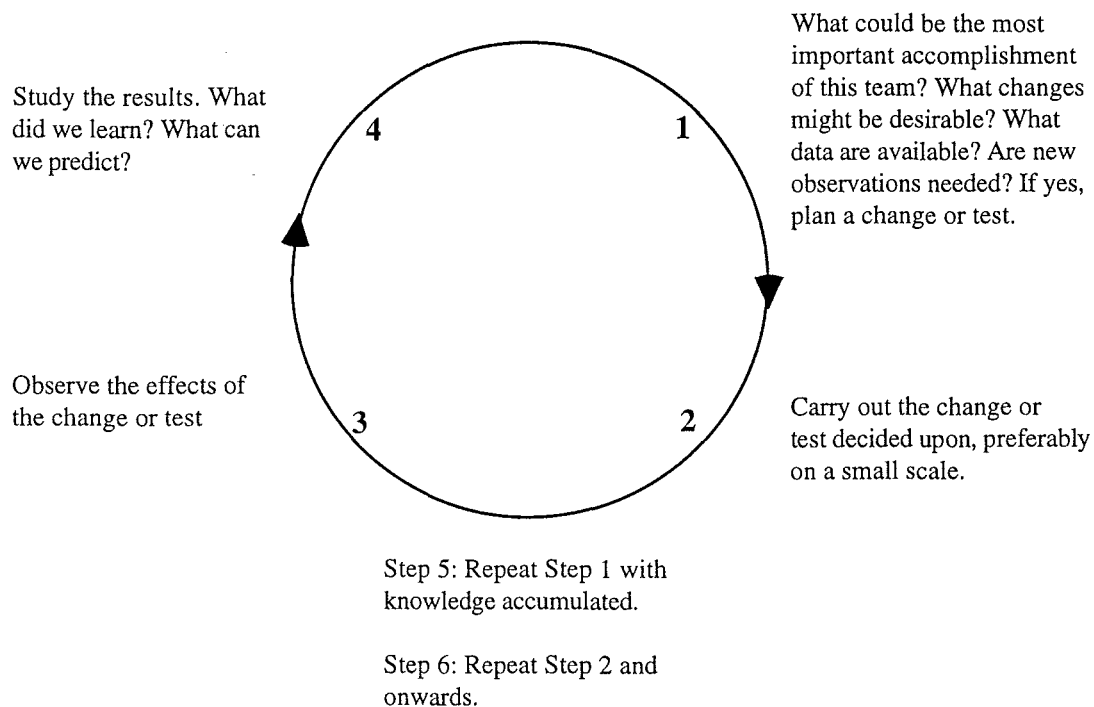


Figure 7.1: The Shewart Cycle

(Source: Based on Deming ,1986, p. 15)

Deming studied the work of Shewart and used those techniques in both the United States and Japan. In the late 1930's, Deming took over sampling for the census bureau in the United States and, later, during World War II used the statistical control techniques for wartime productions. When the United States returned to peace, the country began massive production of consumer goods where there was high demand, no competition and an enormous

unskilled workforce because of post-war migration - factors which did not encourage the use of statistical control techniques. Deming realised that he had convinced the technicians of the benefits of quality control but had failed to convince the managers. He travelled to Japan in the late 1940's, as part of the Allied Occupation, and lectured to Japanese captains of industry about the new principles of management based on developing innovations in quality and productivity. In a lecture to Japanese presidents of major companies in 1950, Deming said that 'the consumer is the most important part of the production line' (Walton, 1989, p.14). The Japanese embraced Deming's philosophy and much of the position of Japan as a world leader in the manufacture of consumer goods has been credited to the Japanese commitment to this philosophy.

The basic tenet of TQM is to assume that the business will be around for a long time and therefore business strategies are developed for the long term rather than for the short term - that is, to create 'constancy of purpose' (Walton, 1989, p. 136). 'Constancy of purpose' can be put in place through: (a) innovation; (b) research and education; (c) continuous improvement of products and services; and (d) maintenance of plant and equipment. Applying these concepts to health care, the 'products and services' are, in most cases, the patients' health outcome. Thus the next section discusses specifically the application and relevance of the continuous cycle of improvement to health care.

7.3 TQM in Health Care

Berwick (1996) describes TQM techniques in health care as 'real time science'. He provides an argument for the use of this technique as opposed to always using randomised clinical trials. This section provides a review of the use of this technique in the health industry. In applying the principle of 'continuous improvement of products and services' to health, process is related to outcome in a continuous cycle for the purpose of refining the process on the basis of outcomes - that is, Continuous Quality Improvement (CQI). More specific references to the use of the quality circle are given by Batalden (et al., 1994) who stated that outcome measurement, process improvement and continual improvement have all been used to improve quality of care but have primarily been used mutually exclusively - not in a cycle.

In health care, one method of continual improvement frequently used in isolation takes the form of developing guidelines for practice. Owens and Nease (1993), in describing the appropriate development of guidelines for health intervention, state:

'The effect of the health intervention should be linked explicitly to the health outcome that the guideline developers aim to improve or prevent' (Owens and Nease, 1993, p. 249).

The Institute of Medicine, USA, described attributes of good practice guidelines as: validity; reliability; applicability; clarity; a multi disciplinary process; review; and documentation (Owens and Nease, 1993). Validity means that the guidelines lead to the appropriate projected health outcomes. Reliability means that given the same evidence and methods for guideline

development, it is expected that the same standards would be produced. Applicability is where the population to which guidelines are to be applied is specifically stated. Clarity refers to the fact that guidelines are easily understood and can not be misinterpreted. All key affected groups or stakeholders must be involved in guideline development and so development in a multi-disciplinary process is essential. A scheduled review ensures that there are statements about when the guidelines are to be reviewed or revised. Finally, careful documentation is essential to ensure the objectives for practice guidelines are that they be accurate, accountable, predictable, defensible and useable. An example of such clinical practice guidelines are those developed by the Agency for Health Care Policy Research in the United States (US Department of Health and Human Services, 1993).

Rosenstein (1994) demonstrated how the use of the complete cycle, or CQI in the form of graphical control charts, can be used in primary health care for prevention, scheduling of appointments and diabetes control. A model in health care proposed by the editor of the Joint Commission Journal on Quality Improvement (Jacquet, 1995) shows such a cycle for improved performance as set out in Figure 7.2.

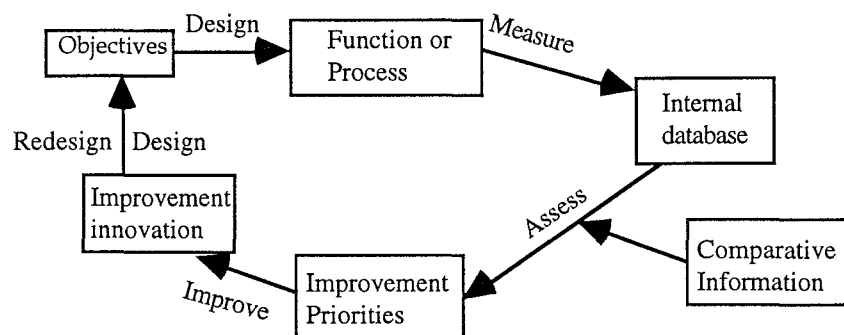


Figure 7.2: Cycle for Improving Performance

(Source: Based on Jacquet, 1995, p. 89)

An example of using such a cycle to alter a process using outcomes in a group of hospitals in Maryland, USA, is given by Kazandjian (*et al.*, 1993). In this study, inpatient outcome assessments are hospital-acquired infections, surgical wound infection, inpatient mortality, neonatal mortality, perioperative mortality, caesarean section rate, unscheduled readmission after ambulatory procedures, unscheduled return to the Intensive Care Unit (ICU) and unscheduled return to the operating room (OR). These data are adjusted for risk factors then analysed for indicator rate trends on a quarterly basis. These data are then related to the process through a communication forum that looks at ways of achieving target rates. The communication forum consists of clinicians who have been involved in the care of these patients.

Another example relating health outcome to process is presented by Krivenko and Chodroff (1994) where appendectomy rates were the outcome measures in the Voluntary Hospitals of America, Pennsylvania, USA. They found the hospitals with the lowest rates and described their 'best practice'. The authors then conducted informal interviews with both the senior staff at hospitals with 'best practice' and with the same staff from other hospitals. Views on policies and procedures, standards and processes of care were exchanged at informal dinners. Discussions of different approaches to diagnosis, monitoring, treatment and peer review were recorded for 'best practice' hospitals. Krivenko and Chodroff claim that an understanding of the relationship between processes/practice and clinical outcomes is gained from the procedure in the non-threatening environment of discussion over dinner and, further, that these comparisons between hospitals improved

outcomes. A similar procedure was used with an outcome measure of caesarean section rates.

Another use of CQI is given by Batalden (*et al.*, 1994) with the 'Serial V' concept that integrates the outcome measurement, process improvement and continual improvement. Batalden and his colleagues maintain the reasons for poor outcomes cannot simply be blamed on the practitioner or the institutions. The causes of adverse outcomes are more complex and relate, as well, to policies, procedures, equipment and techniques. The 'Serial V' concept is reproduced in the model shown in Figure 7.3.

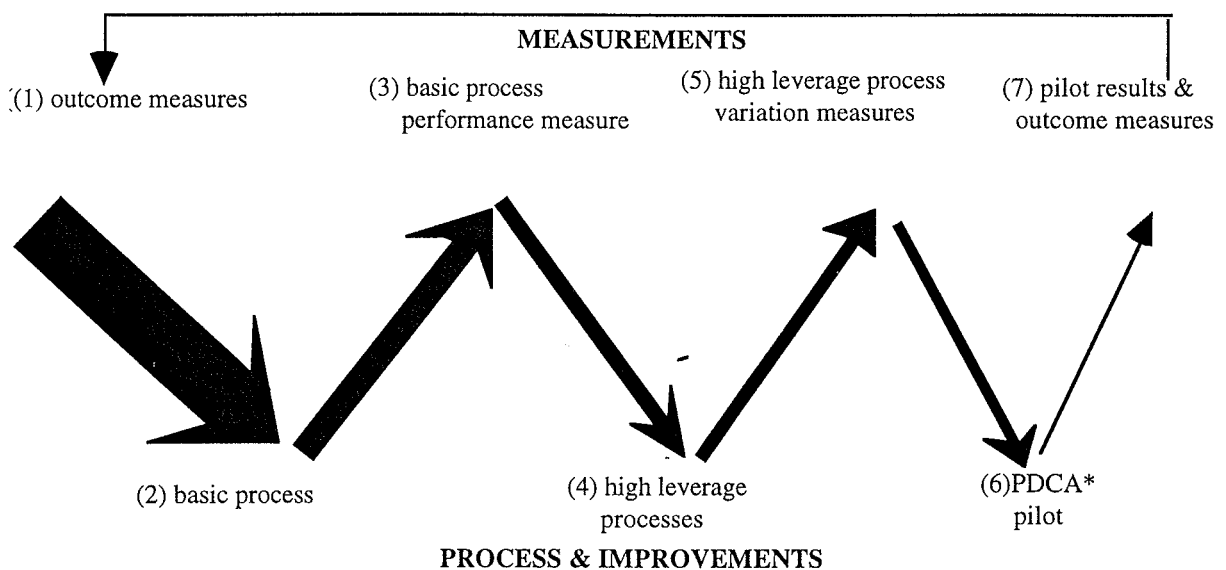


Figure 7.3 Diagrammatic Model of the 'Serial V' Concept

(Source: Based on Batalden *et al.*, 1994, p. 172) *Plan Do Check Act

The outcome measures in the model above can relate to: mortality and morbidity rates; physical, mental and social functioning; satisfaction; quality assurance monitors; and cost/resource usage. The basic process measures might be speed, accuracy, appropriateness and efficiency. The high leverage

processes are more difficult to identify and require both professional and local knowledge and analyses of cause-and-effect. An example of professional knowledge might be that influenza immunisation of high risk groups reduces morbidity and mortality. Local knowledge may show up high leverage processes such as delays in processes causing adverse outcomes. Therefore Batalden and his colleagues suggest that examination of cause-and-effect diagrams reveals high leverage factors. Finally, the PDCA (ie Plan Do Check Act) pilot format occurs where the original outcome measures are re-measured to ascertain the effect of the given pilot test. Figure 7.4 is an example of the 'Serial V' concept for coronary artery bypass graft surgery relevant to this thesis. It is presented below in the "clinical value compass" that is based on Batalden's approach.

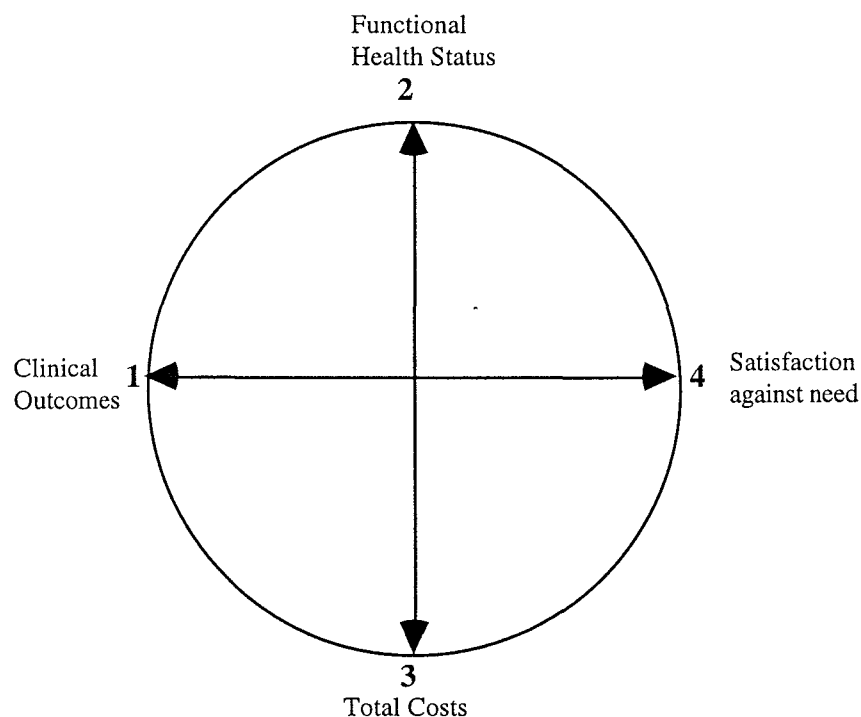


Figure 7.4: Clinical Value Compass

(Source: Based on Batalden *et al.*, 1994, p. 175)

The cardinal points of the compass are represented by the numbers 1 to 4 with the following meaning:

1. clinical outcome = leg wound infections and mortality;
2. functional health status = return to work;
3. patient satisfaction = pain management and communication; and
4. costs = hospital length of stay, ICU length of stay, and readmissions within 30 days.

Carman (et al., 1996) reviewed TQM and CQI in 10 hospitals to identify the factors that indicate success in implementation in CQI in acute hospitals. These measures of success related to the environment and the resources of the hospital. They were gauged by the hospital's revenue, market share, growth, length of stay, cost, productivity, competitiveness, the implementation approach for CQI, physician participation in CQI, the culture, the depth of implementation measured by volume of quality improvement activity and use of outcome and patient satisfaction studies. They found a strong link between culture and change in customer satisfaction scores. This finding was reinforced in a study by (Shortell et al., 1995). The findings of the study by Carman and colleagues suggest, but do not prove, that if the workforce has the perception that it is empowered then there is improved customer satisfaction and efficiency. One of the major implications of the study was that "CQI programs need to encompass pre- and post-acute care and CQI programs need to be integrated across these internal customer and supplier provider groups" (Carman et al., 1996, p. 59).

Another recent study of 15 095 coronary artery bypass graft patients in Maine, New Hampshire and Vermont had as the main outcome measure, observed

and expected hospital mortality (O'Connor *et al.*, 1996). The CQI techniques used were to: firstly, feed back outcome data to the clinicians; secondly, to provide two-day training sessions for executives and four-hour training for general members on the theory and techniques of CQI. The results showed that when coronary artery bypass graft patients' casemix was controlled for, the observed inpatient mortality was statistically significantly lower than the expected mortality levels. The study summarised changes in protocol prompted by the CQI techniques.

A detailed review of TQM literature in health care is given by Gann and Restuccia (1994). Twenty six recent research studies have been assessed with respect to validity and reliability. The review carefully examines each study in terms of principles of TQM but does not place a great deal of emphasis on the measurement of outcome. A less detailed review of TQM in the health sector is given by Motwani (*et al.*, 1996).

The last two sections have provided a summary of applications of TQM to industry and more specifically to health care. Before describing the patient-outcome model, the principles used in the design and use of such models is discussed.

7.4 Principles for Design and Use of Models

Prior to designing a measure of patient-assessed outcome, it is important to note that there are many factors that contribute to the outcome and these, therefore, must be taken into account in the design of the measure of a

patient's health outcome. These contributing factors are summarised in the Figure 7.5.

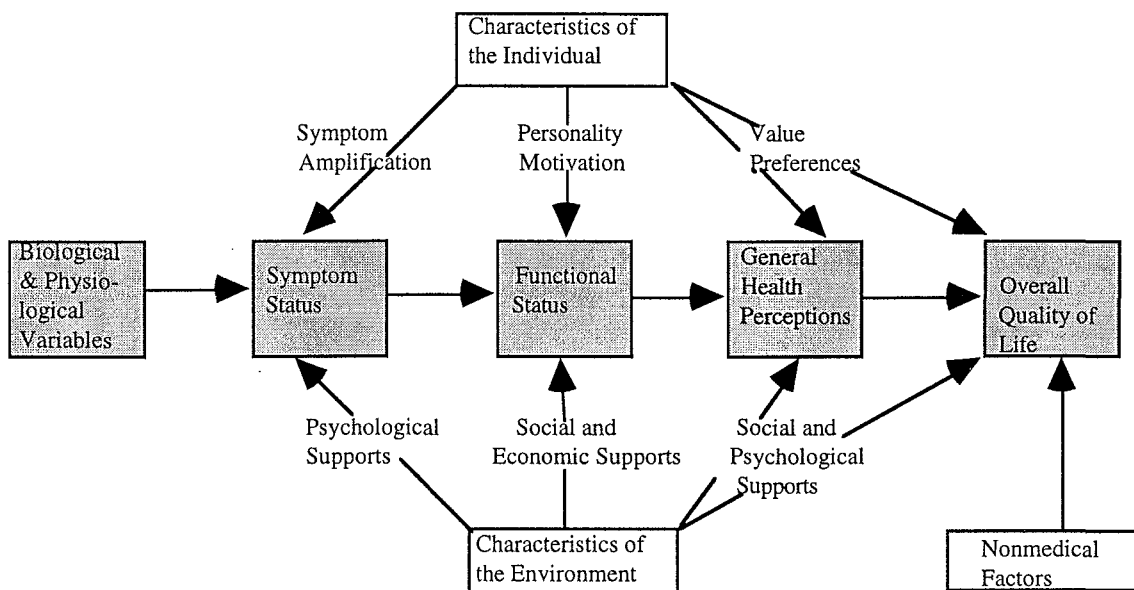


Figure 7.5: Factors Contributing to Patient's Health Outcome

(Source: Based on Batalden *et al.*, 1994, p.1 75)

If patients are to assess their quality of life following some medical intervention, it is important that this assessment is put in the context of the characteristics of the environment and of the individual. Surrogate measures for characteristics of the environment might be demographic factors such as age, gender, income, marital status, family support and employment. Characteristics of the individual, and expectations of the intervention, would be ascertained through an appropriate survey instrument administered pre- and post-intervention.

In Australia, Ewan (*et al.*, 1994) have developed a national framework for incorporating health impact assessment into the process of environmental

impact assessment. The philosophy underlying the proposed framework includes, *inter alia*, the interdependence of human health and the environment (Ewan *et al.*, 1994, p. 7). The study emphasises the need to take an holistic approach to the impact of development on the public health and well-being of the community, including the application of a system approach. Systems analysis is a process for planning and managing complex physical or socio-economic/environmental systems where there are interactions amongst component parts of the system (de Neufville and Stafford, 1974). Applied systems analysis is a 'set of computer-based methods essential for the planning of major projects' (de Neufville, 1990, p. xiii). An important part of systems analysis involves the modelling of the system in question as a basis for systems understanding, prediction and control.

The principles for the design and the use of models will be discussed before outlining the specific patient health outcome model. A general introduction to modelling is provided by Lee (1973) who describes the five major steps required in model design. These are illustrated in Figure 7.6 as problem identification, formulation of initial model, model validation, reformulation of model and application of model.

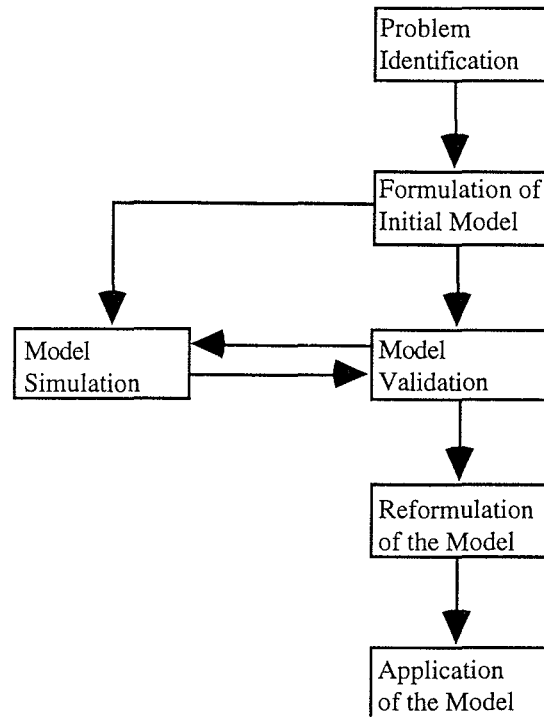


Figure 7.6: Model Development as an Iterative Process

(Source: Based on Lee, 1973, p. 18)

The patient-outcome model that is proposed follows Lee's principles with the further enhancement that the last box in Figure 7.6 - the 'Application of the Model' is included in the 'Reformulation of the Model' box so that model application, modification and validation occur in a continuous cycle. In the model proposed, the general 'problem definition' is defined specifically as 'how health outcomes assessed by surgical patients post-discharge can be related to the surgical process with the purpose of both process and outcome improvement on a continuous basis?'

In terms of 'model formulation', one of the most crucial decisions relate to the selection of appropriate variables to include in the model. In the case of the proposed model, it is important, firstly, to design a way of measuring patient-assessed outcome (as has been covered in Chapter 3) and, secondly, to include all patient-related social and economic environment variables that may influence the patient's assessment of their health outcome. Thirdly, the model needs to include all measurable process variables (as demonstrated in Chapter 5 and 6). In this particular model, these are variables that can be specifically controlled during surgery. There are, of course, many other process variables that can be controlled during an episode of hospitalisation.

In Figure 7.6, 'level of aggregation and method of categorisation' refers to the requirement for appropriate ways to aggregate the results so that mean values for the group can be compared and how data should be adjusted for risk. This is important and will be discussed in more detail in the next section.

'Simulation and validation' of the model, as defined by Lee, refers to trying to reproduce the real world to test the proposed model. However, with the patient outcome model there are no ethical reasons why the initial cycle of data collection and analyses cannot be conducted on real patients, and this factor was taken into account in model design. These principles of model design have lead to the model described in the next section.

7.5 The Continuous Quality Improvement Model

The model presented in Figure 7.7 represents only the outline of the conceptual structure of the final proposed model. Each section of the model is separately discussed with greater elaboration and specific reference to its use in the case study of coronary artery bypass graft patients. The advantage of this model is that data collected at each point in this model are generally available in most pre-existing clinical information systems, and therefore require minimal extra input for data collection.

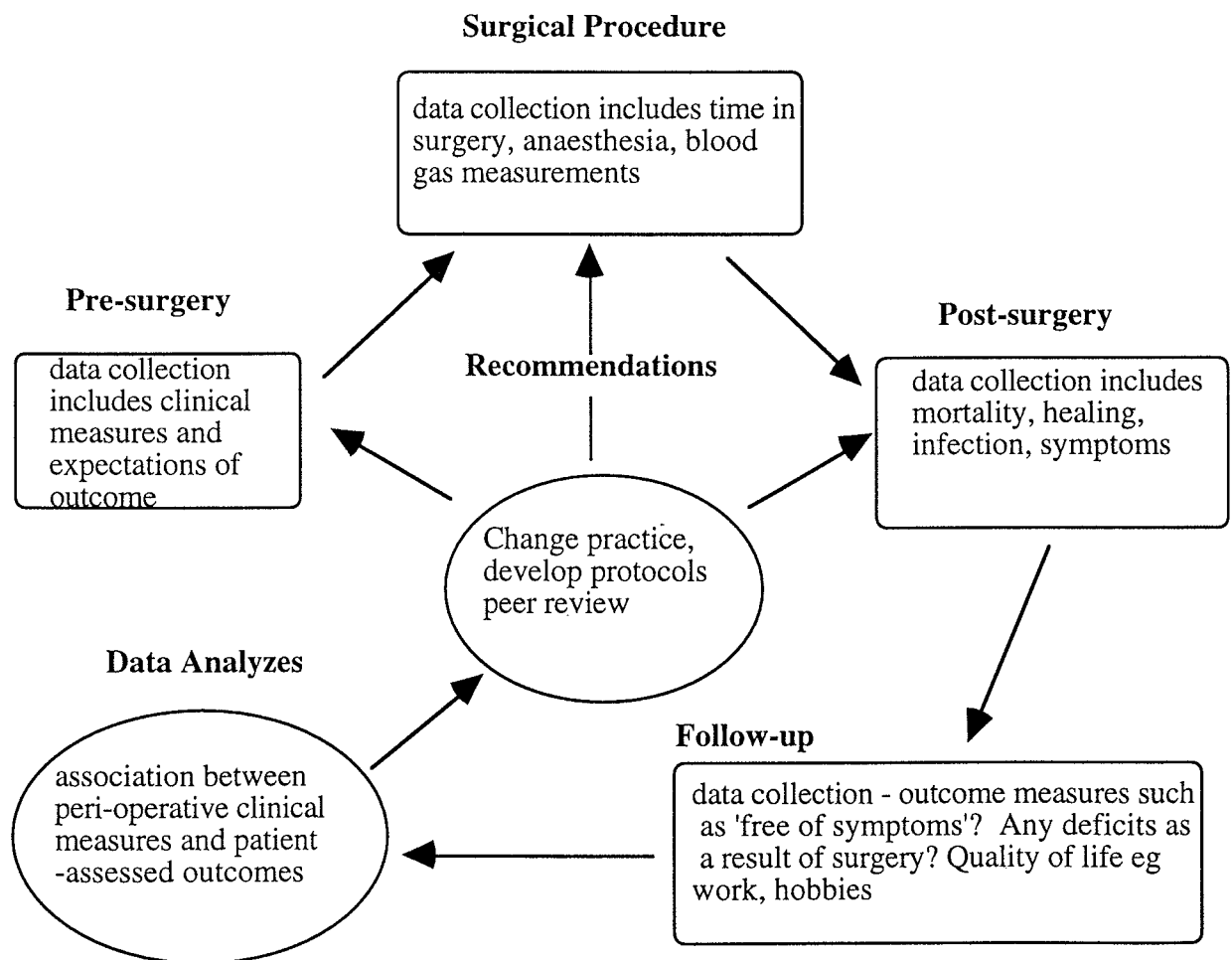


Figure 7.7 Outline of Model to be Developed

As the model presents a cycle of continuous improvement there is no starting point. However, it may be appropriate to choose, from Figure 7.7, the point at which data are collected pre-surgery and follow the cycle in a clockwise fashion.

7.5.1 Pre-Surgery Data Collection Procedure

The variables required in the pre-surgery data collection fall into three categories:

- (a) patient's social and economic environment surrogate measures;
- (b) the symptoms causing the patient to seek surgery and their expectation of that surgery; and
- (c) health status measures.

These data include data collected for all surgical patients and data which are intervention specific. The collection of data in these categories will be discussed with respect to the case study of coronary artery bypass graft surgery; and then with respect to the general model.

7.5.2 Determinants of Social and Economic Environment

The measures that have traditionally been used to explain the patient's social and economic environment are (Australian Bureau of Statistics, 1991):

- age
- gender
- ethnicity

- country of birth
- main language spoken at home
- religion
- home postcode
- marital status
- household composition
- highest level of education
- employment status
- occupation
- comorbidities and/or disabilities
- private health insurance status.

In the case study of coronary artery bypass graft patients, the patient's medical record includes age, gender, home postcode, private health insurance status and comorbidities. These data are not transferred electronically from the hospital's information system but are entered manually onto the appropriate spreadsheet on the perfusionist's PC computer. The PCs, including the terminals forming part of the heart/lung machines' setup, are networked within the pump room but not to the hospital-wide electronic information systems. The unique identifier on the pre-surgery database allowing linkage to other databases is the patient ID number and the date of surgery. The date is required as the patient might have surgery on more than one occasion. The other data, not obtained from the medical records, are collected on cards by the perfusionist during a bedside interview with the patient prior to surgery. The data are then transferred from the cards to the database. The database used by the perfusionist is Access III.

The ideal general model would have the social and economic and environment explanatory variables electronically collected using an integrated record that is accessible by all health services. It is relevant at this point to review and discuss how recent literature on developments in hospital information systems and visions for the future might be incorporated into the general model.

The literature addresses *how* the patient information should be recorded, *what* should be in the record and *why* in terms of how the information is to be used. One of the major problems of how to record computer-based patient data is the vast amount of information required in the record. Bourke (1994), in describing information systems including clinical information, stated that they are better to be based on object-oriented databases (ie use drawing and voice) because the amount of data produced by the human body in one 24 hour period is 'overwhelming'. Certainly, given the vast amounts of clinical data in the system, data entry must be both easy and fast. However, discussion of computer-based patient records by McDonald (1992) highlights the need for structured, as opposed to free form, machine ready information. He states that free text information has merit in the direct care of patients because of the ease of retrieval and display but not in terms of research and decision support.

Levy and Lawrance (1992, p. 127) make two important propositions to be considered and evaluated in data acquisition:

- " • Proposition 1: Insofar as possible, each event should be recorded one and only once as proximate as possible to its occurrence.

- Proposition 2: Data should generally be recorded by the individual first acquiring it."

They also stated that terminals should prompt for data, be clear in their display and have adequate screen size. They also state that the problems for data acquisition are that:

- (a) nomenclature for tests differ from system to system;
- (b) the nature of data needs to be standardised; and
- (c) although laboratories and other highly specialised units need to maintain specialised information systems, it is important that data can be received and understood by other sections of the system.

In terms of the scope and comprehension of data, McDonald points out that there is a requirement to be able to merge data collected from multiple sources, multiple users and at different times. In summary, the systems should be user friendly, accurate and flexible. It is imperative that databases from different sources are able to be integrated.

In terms of what should be included in the record, McDonald (1992, p. 6) lists the data elements suggested by expert users:

- "• the problem list
- medication profile
- lists of acute and minor problems, long-term and major problems
- social and functional measurements such as daily activities, meals, kind of work done, home setting, job satisfaction social community etc.

- manifestations, studies and test orders, results, and clinical course
- psychosocial values, including information about family situation, patient expectations, and so on"

The desirable data elements that Davies (1992) believes would be useful are :

- identifiers
- characteristics such as demographics, referring physician, type of admission
- dates of admissions, test, etc.
- codes such as DRGs, procedure, diagnosis
- clinical indicators
- process of care
- disposition such as referrals, transfers and follow-up
- outcomes.

Davies has placed greater emphasis on the usual data elements included in a medical record whilst McDonald has acknowledged the need for more information on the social environment of the patient and their expectations as is proposed in this study's model.

What has been proposed in recent literature on computer-based patient records in terms of requirements for hardware, software and system design is extensive. It is important therefore to discuss *why* these data need to be collected. McDonald (1992) believes that requests for data come from two sources - clinicians in one group and policymakers, managers and researchers

in another group. Clinicians, he states, need the information for clinical care, such as medications, test results and problem lists. The other group require information to satisfy regulations and quality assurance requirements. He states:

"Policymakers and researchers want information that will help them assess illness severity and judge outcomes. This would include a minimum of information about the patient's mental and functional status, as well as disease and/or procedure-specific information to characterise severity before and after the procedure."

(McDonald, 1992, p. 7)

Ross Davies (1992) defined the broad categories for research using patient records as : quality of care evaluation; patterns of variation service; utilisation studies; and outcomes studies.

In terms of outcome studies, she also believes that it is relevant for patient-assessed health status measurement to be discussed by the Institute of Medicine (IOM) Committee on the Future of the Patient Record in the United States.

Shortliffe (et al., 1992) describes the vision for computer-based patient records as a physician's work station being an integral part of the normal practice environment. He outlines the functions of these records to be fulfilled as :

- "• provide clinicians with readily accessible, intelligent assembly of clinical data that

characterises the condition, prior management, and current treatment plan of a patient

- provide a clearly defined and well-organised database to permit epidemiologic assessment of patient outcomes and patterns of practice, revealing merits of management strategies in specific clinical contexts
- provide interactive decision support tools based on information derived from critically analysed population data
- decrease administrative record keeping and form submission responsibilities that burden health care providers" (Shortliffe *et al.*, 1992, p. 273)

Both the articles by Shortliffe and colleagues and by Davies emphasise the importance, in the future, of outcomes in decision-making based on computer-based patient records.

In agreement with the literature, this general model, therefore, proposes that data sources using this integrated system would include the patient's outpatient detail, relevant information from general practitioner's patient files and further details recorded at the time of admission. This clinical information would be linked to all areas of the hospital including radiology, pathology, medical and surgical wards, ICU and theatres. There would also be a link to the management information system. If, however, it was not possible for all data to be collected on admission, there would be a bedside terminal where data can be entered directly to the system during, or

immediately following, the patient interview. Entering data on admission, or at the pre-surgery interview, would require, at most two keystrokes for each variable with the only exception of the four digit postcode, As an example, highest level of education would have pre-coded categories as follows: (1) primary school education only; (2) less than 4 years high school; (3) School Certificate or equivalent; (4) Higher School Certificate or equivalent; (5) trade certificate; (6) post-secondary diploma; (7) undergraduate degree; and (8) postgraduate degree. Data capture would be via menu prompts for the appropriate code for the given variable.

7.5.3 Symptoms Prompting Surgery and Expectations of Surgery

A reasonable presumption is that the intervention has been sought to alleviate symptoms. Thus, in the pre-surgery interview, the symptoms and their impact on the patient's activities of daily living need to be ascertained. Therefore, the most important question to be addressed is '*why did the patient seek surgical intervention?*' The patient's choice, on the advice of their physician, to seek surgical intervention with the accompanying general anaesthetic should not be taken lightly given the physical and emotional impact on the patient and the high cost of the procedure.

At the same time, the patient's expectation of their needs should be determined so that Calman's Gap, which measures the difference between expectation of an health intervention and the patient's perception of the fulfilment of the expectation post-intervention, can be assessed. Calman's Gap has already been discussed in Chapter 2, Section 2.2.

In the pre-surgery interview for coronary artery bypass graft patients, conducted by the perfusionist, the patient's symptoms are recorded on cards. The symptoms are most frequently one of, or a combination of, shortness of breath (recorded in the spreadsheet as SOB), angina (recorded as such) and acute myocardial infarction (recorded as AMI). These are later transferred from the cards to the spreadsheet where they are respectively recorded as 'SOB', 'angina' and 'AMI'.

Detailed on the cards, at the same time, are the patient's hobbies and leisure activities and whether the symptoms prevent him or her from participating in these activities. The patients are then questioned on their expected participation in employment, hobbies, sport and social activities following their recovery from the surgery. The expectations of patients undergoing surgery can differ widely. As an example, a 75 year old woman may wish to climb one flight of stairs to her home, be independent enough to shop, cook and clean for herself and be capable of catching public transport to visit members of her family. By comparison, a 40 year old woman working as a lawyer with in a busy practice may wish to continue her practice on recovery and take up, again, her hobby of marathon running. These data about leisure activities are recorded in an abbreviated manner on the cards and later entered into a spreadsheet under such headings as 'hobbies1' and 'hobbies2'.

The optimum situation for the general model would be one where the medical records for the consulting physician, who referred the patient for surgery, would be electronically linked to the hospital's clinical information system. The patient's symptoms would be recorded by the physician on his or her database with a unique patient identifier that coincides with the

hospital's identifier. All symptoms would be pre-coded. In the case of cardiothoracic consultants, they would have pre-coded symptoms as a menu item on their computer screen such as a '1' for 'shortness of breath', '2' for 'angina', '3' for 'acute myocardial infarction' and '4' for 'other'.

With respect to recording expectations of health outcome, data would be captured by a bedside interview and entered onto the relevant database. The interviewer would be prompted on pre-coded questions, such as current employment status, expected employment status post-recovery, or other activities relating to retirement, where appropriate. Again, a minimum number of keystrokes would be required for pre-coded categories of sports and hobbies.

7.5.4 Physiological Measures

In the case study of CABG patients weight in kilograms, height in centimetres, body area, blood pressure and ejection fraction are taken from the patient's notes and transferred to the perfusionist's cards. These data are later transcribed from the cards to the appropriate database.

The appropriate physiological measures for the procedure should be transferred electronically from the hospital's information system to the database in the general model.

7.5.5 Patient Follow-up

The survey instrument and method of administration for the CABGs patient in the case study have been presented in Chapter 4 Section 4.2. This interview is conducted by telephone one year after the graft surgery. Patients are specifically asked if they have any deficits as a result of the surgery.

In the general model the timing of the follow-up interview will depend on the surgical procedure. The most appropriate method would be telephone interviews with three follow-up attempts at different times if there is no response. The questions should initially be open-ended, as in the case study of coronary artery bypass graft patients. A suggested format has been designed and this is given over the page in Figure 7.8. Data capture will occur through a menu-driven program that will prompt for responses.

I am ringing for Dr X at W hospital who was the surgeon for you (type of surgery) months ago. This is a routine follow-up call to see how you are. How are you?

Record actual comments eg 'better than ever'
'not myself'

Are you happy with your recovery since the operation and with your current health?
Yes/No

If No, probe whether problems relate to
(1) surgical procedure;
(2) pre-existing condition;
or
(3) or new condition since surgery.

In an interview prior to the operation, you stated your symptom/s before surgery were .

Are you free of the symptom/s?
Yes/No

In the same interview, you stated that you hoped to when you made a full recovery. Have you been able to participate in these activities?
Yes/No

Have you noted any
changes in your memory? Y/N
changes in your moods? Y/N
ability to concentrate? Y/N
level of depression? Y/N
irritability? Y/N
social life? Y/N
sex life? Y/N

Were you satisfied with treatment at Hospital X? Yes/No
If No why?

Have you experienced any financial difficulties because of the operation? Y/N

Thank you for your time, this information will be used to improve the services at Hospital X.

Figure 7.8: Suggested Format for Follow-Up Interview Instrument

7.5.6 Data Analyses

The continuous quality improvement program implemented by the perfusionist has been to plot on a 3-dimensional graph at intervals when sufficient data have been collected. The variables analysed in each graph were:

- (a) various blood gas measurements including pH, PO₂, PCO₂, haemoglobin (on the x-axis);
- (b) time on bypass measured at time 0, 10 minutes on bypass and then every 20 minutes until 150 minutes (on the y-axis); and
- (c) average NSUM, a score for neurological damage (on the z-axis).

Peaks occur in the three dimensional graphs where the average NSUM is high. It would therefore be expected that the patients would wish to minimise neurological damage and would prefer to be in the valleys where NSUM is lowest. At six monthly intervals, the perfusionist produced such graphs which were based on the total number of patients at that time for each of the blood gas measurements. Subsequently, he attempted to maintain the blood gas variables in the valleys for all patients. When another six months had passed, new graphs were produced and the optimum levels for the blood gas variables were further modified.

An example of such a graph is given in Figure 7.9 and is for all patients who had CABG surgery up to, and including, the end of 1992. Figure 7.9 takes pH as an example and demonstrates that patients would want their pH not to be greater than 7.4mm Hg when they are initially put on bypass, after 10 minutes

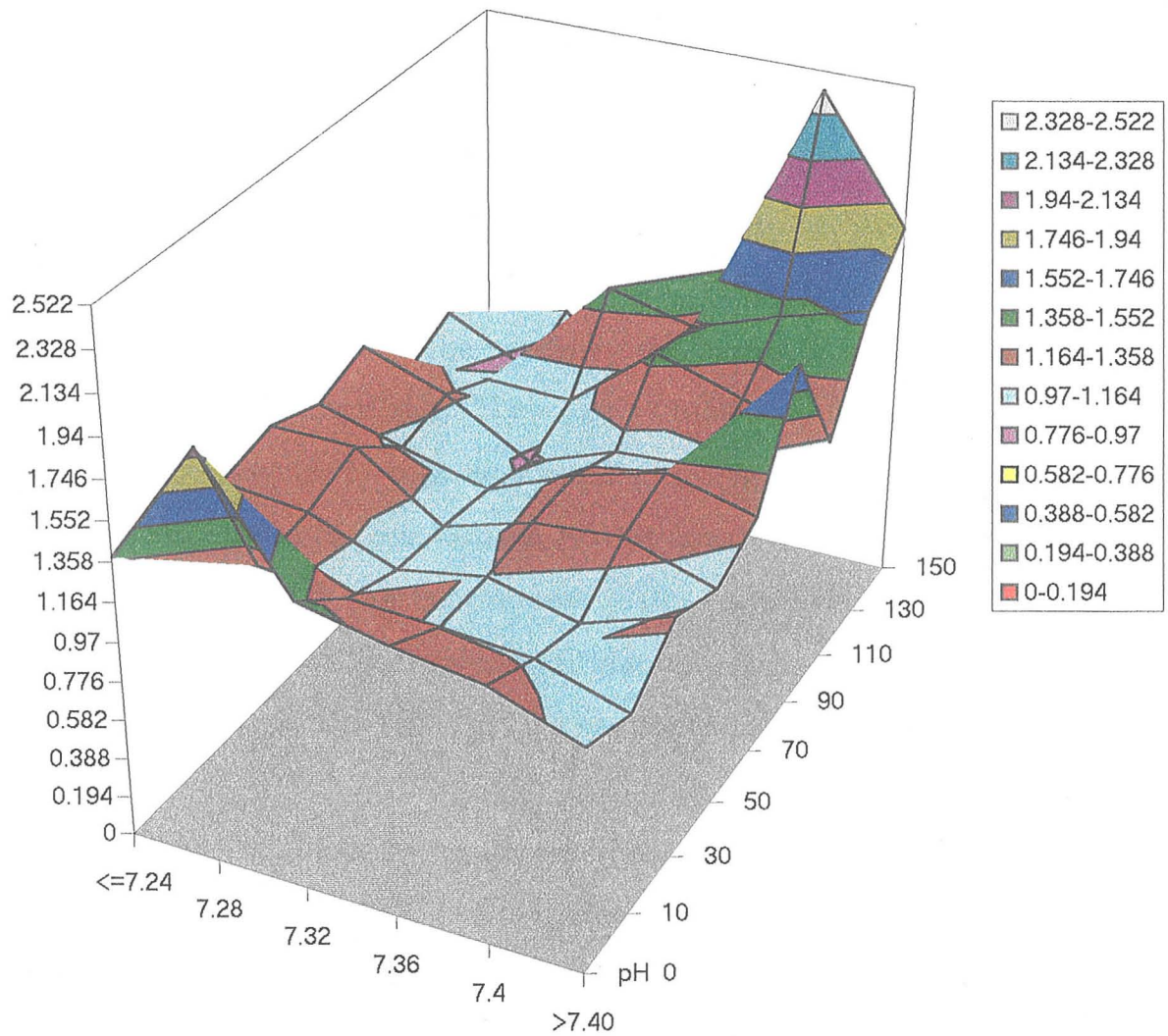


Figure 7.9: pH by Time by Mean NSUM

they would wish it to be around 7.36 mmHg, and so on. It should be noted that this path follows neither the recommended paths of alpha-stat or pH-stat which are recommended in the literature on cardiopulmonary bypass (Davies, 1993) which suggests a strong case for using patient outcomes as an indicator of the correct level for pH rather than the conventional options for pH.

The monitoring and control of pH during cardiopulmonary bypass provides one example of how the perfusion process can be used to improve neurological outcome. This continuous quality improvement process has been used (and continues to be used), by the perfusionist for a number of other variables including PaO₂. The success of this strategy has already been mentioned and has clearly been demonstrated in Figure 5.11 which shows an increasing proportion of patients with NSUM equal to zero corresponding to the implementation of the continuous quality improvement cycle of the perfusionist. Other data analyses include graphing rates of angina and rates of physical damage (PSUM). These graphs have been presented in Chapter 5.

A general model that can be readily modified for other surgical procedures has been devised, and this is now described in detail with reference, where necessary, to CABG surgery.

Data analyses in the general model should be simple and easily understood by the person conducting the analyses. The outline of the questionnaire described in the patient follow-up general model allows three important measures of outcome to be derived:

- (a) a dichotomous response Yes/No as to whether they have achieved their expected outcome in terms of activities of daily living;
- (b) a score can be derived for neurological damage; and
- (c) a score can be derived for physical damage.

Depending on the procedure, it may be desirable to just have dichotomous measures of neurological and physical damage rather than scores - that is, no damage and some damage.

In this simple model variables fall into categories - outcome measures, social, economic and environment measures, symptoms and expectations of surgical procedure and physiological measures.

The aim of the data analyses is to determine which controllable variables (that is, measures that can be altered during the procedure) have an association with outcome. Before concluding that a variable is influencing outcome, it is important to ensure that the controllable variable is not highly correlated with social or economic or environmental measures or symptoms. The diagram in Figure 7.10 gives a simple example of data analyses.

Examining Figure 7.10 from left to right, patients outcomes are categorised as either 'good' or 'bad'. At a scheduled time, ensuring a sample size of at least 25 to 30, the data on outcome for patients should be aggregated. For those patients with a 'good' outcome, the mean for the controllable variable should be calculated. A similar calculation for the patients with a 'poor' outcome

should be completed. A t-test should be conducted to see if there is a difference between the means. If there is no statistically difference between the means at a level of significance of say 5%, the mean for another controllable variable should be calculated. However, if there is a significant difference, it is necessary to ensure that the controllable variable is not associated with a non-controllable variable and therefore acting as a surrogate for this variable.

An example of this might be that age of the patient (a non-controllable variable) is highly correlated with a controllable variable, haemoglobin, and therefore the 'poor' outcomes are not caused by haemoglobin but by age. Ideally, outcome measures should be standardised for non-controllable variables that affect outcome. If the results show that there is a significant difference in the means of the controllable variable for the two outcome levels and there is no association between the controllable variable and non-controllable variables, then the procedure indicates that the controllable variable should be run at the mean level that achieves the 'good' outcome.

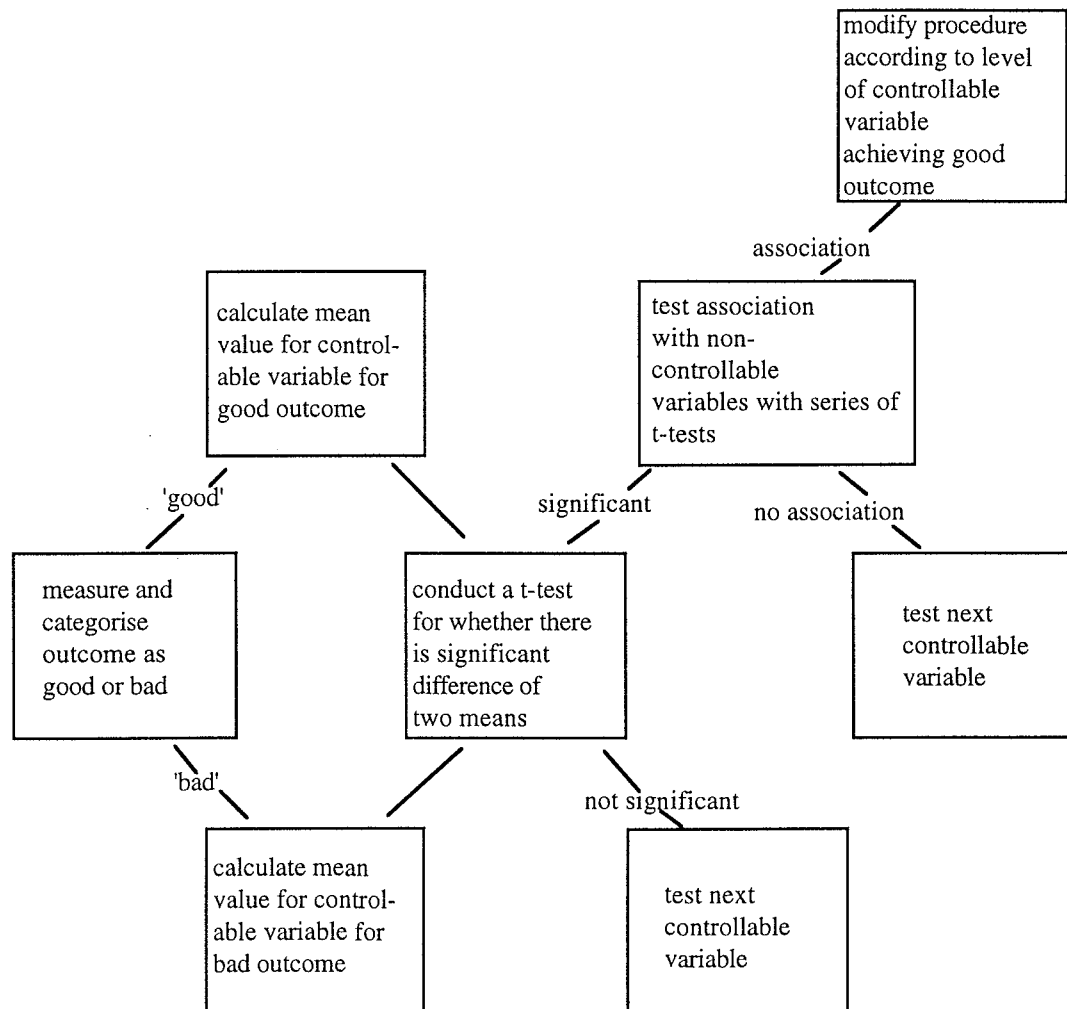


Figure 7.10 : Model for Data Analyses

The model in Figure 7.10 assumes that the controllable variable is continuous. If the controllable variable is discrete a chi-square test rather than a t-test should be used. Examples of variables to be modified can be ascertained using outcome as a response in the regression model described in the previous chapter. In the case of NSUM, as an outcome response variable, the procedure would be modified according to optimal partial oxygen levels over time that minimise neurological deficit.

An example of use of this model is presented for a hypothetical sample of 100 patients undergoing a surgical procedure. For simplicity, assume the only possible variables to be measured are those in Figure 7.11 - that is, gender, age and haemoglobin level.

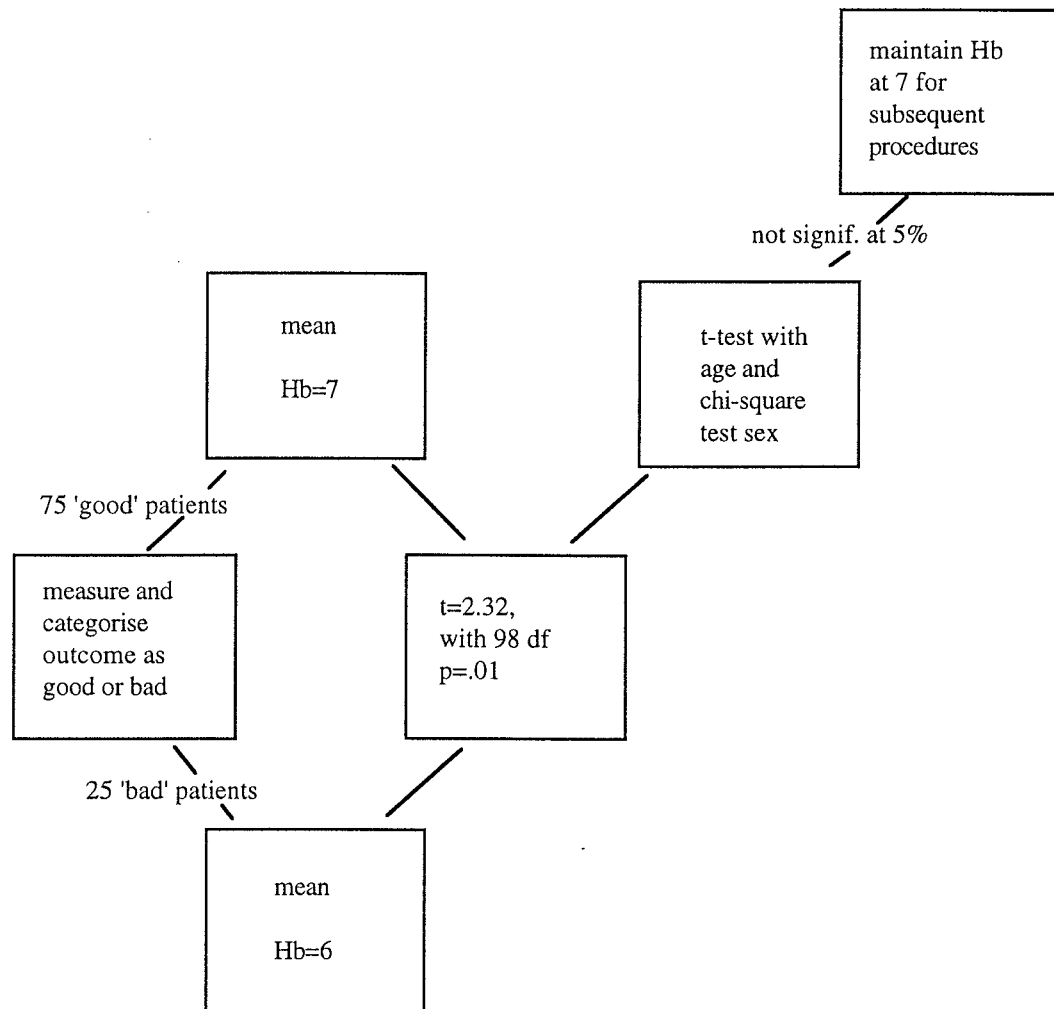


Figure 7.11 Example of Use of Model

Application of the Model for Data Analysis in Figure 7.10 leads to the results shown in Figure 7.11. The example is very elementary in terms of variables measured and level of sophistication of data analyses and requires no more than a spreadsheet software package. It is essential that the person

undertaking data analyses and interpretation is familiar with the way the controllable variables are measured and how they can be modified. The data analyses step in this model needs to be adapted to a user with a reasonable balance between knowledge of how the variables are measured, or modified, and knowledge of basic univariate statistical analyses techniques. The technique will fail if sophisticated statistical analyses are conducted on the data by a statistician with little understanding of how the data are derived or how they can be sensibly altered. Vice versa, the technique will also fail if the health professional, who may have an excellent understanding of the manipulation of measures taken during the procedure, but finds the results of the complex data analyses a mystery.

Other analyses should include examining patterns in outcome measures over time. Hopefully, a pattern of proportionally fewer patients with unsatisfactory outcomes should emerge over time, given modification in procedure. However, if this pattern does not emerge then there is a need to verify that the characteristics of the patients, which may affect outcome, have not changed over time.

7.5.4 Action Based on Data Analyses

The analyses of outcomes in relation to process variables leads to recommendations to change practice and/or develop guidelines. However, change in practice cannot be initiated through presentation of data analyses alone. The culture of an organisation needs to be changed. Batalden (*et al.*, 1994, p. 169) stated that just telling a surgeon about his or her mortality rate for CABGs 'is a necessary but insufficient step towards improvement'.

Analyses of the patient outcome data in the case study of CABG patients, has lead the perfusionist to follow the path that minimises adverse outcomes. Taking pH level as an example, Figure 7.9 demonstrates that patients would want their pH to be greater than 7.4 mmHg when the are put on bypass, after ten minutes greater 7.36 mmHg, and so on. As more patients are added to the database, new three-dimensional graphs are produced and the process of minimising adverse outcomes by further refinement of the optimum path continues. There is, therefore, a continuous process of improvement which has been demonstrated by Figure 5.11 in Chapter 5.

Also, the perfusionist attempts to keep the other blood gas measurement for the patients within their respective valleys. The control of these variables is constrained by the fact that blood needs to be extracted at the scheduled intervals for blood gas analyses. There are presently no mechanisms for constant blood gas analyses. However, given the demonstrable improvement in patient outcome over time, the perfusionist at the study hospital is presently in negotiation with a medical equipment manufacturer to provide constant blood gas measurement throughout the perfusion process.

In this case study, apart from the on-going involvement of the hospital perfusionists, there is no formal discussion of the results of the data analyses with other health professionals. Notwithstanding the lack of formal discussion, the surgeons at the hospital are aware that their patient outcomes are being monitored and that analyses are being undertaken in relation to physiological variables that are under their control whilst the patient is on the heart-lung machine.

In the general model, a number of ways of presenting the relationships between outcomes and process are proposed. The methods fit into three broad categories: peer review; exception reports; and finding the lowest rates for outcome and describing 'best practice'. Each method will be discussed separately and simple examples of applications will be given. In the actual implementation of the model, these techniques for proposed action as a result of outcome data analyses are not discrete but would, most likely overlap to optimise the process.

One method - *peer review* is to present surgeons with comparative rates for poor outcomes for their patients on a quarterly basis. Ideally, the method used to measure outcome would have been designed in consultation with the surgeons. As emphasised earlier in this chapter, it is important that these measurements are standardised for patients' social and economic circumstances. The surgeons must be confident that the rates for outcomes that they are presented with provide a true measure of outcome. CQI will not be successful unless there is a commitment amongst the health professionals to quality improvement.

A measure of the difficulty of changing the culture amongst clinicians is clearly demonstrated in a study by Berwick (et al., 1992). Two hundred medical directors of health management organisations (HMOs) and 451 physicians were surveyed about their attitudes to TQM. More than 8 in 10 (81%) of the HMO medical directors placed 'a great deal of emphasis' on TQM compared to only 34% of physicians who agreed with the statement. Responses to other questions gave further insight into the gap in attitudes

towards TQM between managers and clinicians. The respondents were asked 'which of the following statements most closely reflects your views?' The first statement was 'quality management... makes more sense in an industrial setting than it does in health delivery'. Eleven percent of managers agreed with this statement compared to 48% agreement amongst physicians. Similar attitudes were revealed in their respective responses to another statement - 'health delivery is a production process well-suited to the application of quality management' with 87% of medical directors agreeing with the statement and 51% of physicians agreeing.

The rates should be presented for discussion, in confidence, amongst surgeons only and are not presented in a punitive and public manner. Following the discussion, the measurement of outcome may be refined according to any agreement amongst surgeons that there is some bias in the measurement tool that needs to be adjusted. CQI not only involves continuously improving process but should also be used to refine the measurement of outcome itself. Rates should be presented in aggregate form on a regular basis. Suggested intervals for reporting are quarterly or six monthly ensuring the rate is based on at least 20 to 30 patients' outcomes.

Formats for presentation have been designed and are given in Figures 7.12 and 7.13. The aim in Figure 7.13 is to present an example of a computer-based reporting form where the percent of adverse outcomes are accumulated over time for each surgeon undertaking the operation.

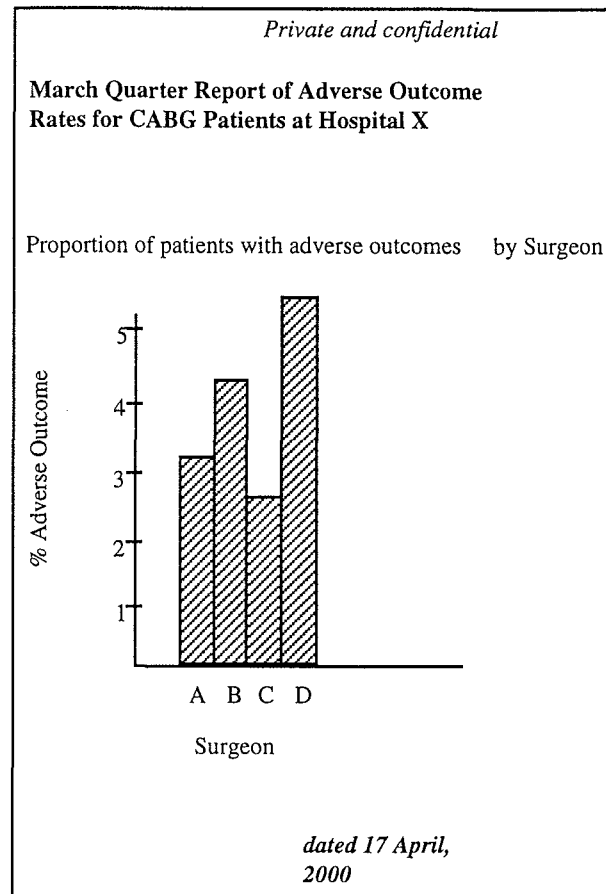


Figure 7.12: Report Format A for Peer Review

A productive result of the presentation of Figure 7.12 might be that Surgeon C and Surgeon D compare the characteristics of their patients and their process for the March Quarter (Figure 7.13). If there are perceived differences in the characteristics of patients, it would be recommended that all outcome data be analysed with respect to the given patient characteristic. A recommendation might be that this type of surgery is contraindicated for these types of patients. If there are differences in a process variable, the recommendation might be for Surgeon D to trial Surgeon C's procedure in the next quarter. In summary, peer review provides a mechanism for improving outcome through informal agreement on appropriate alterations to practice.

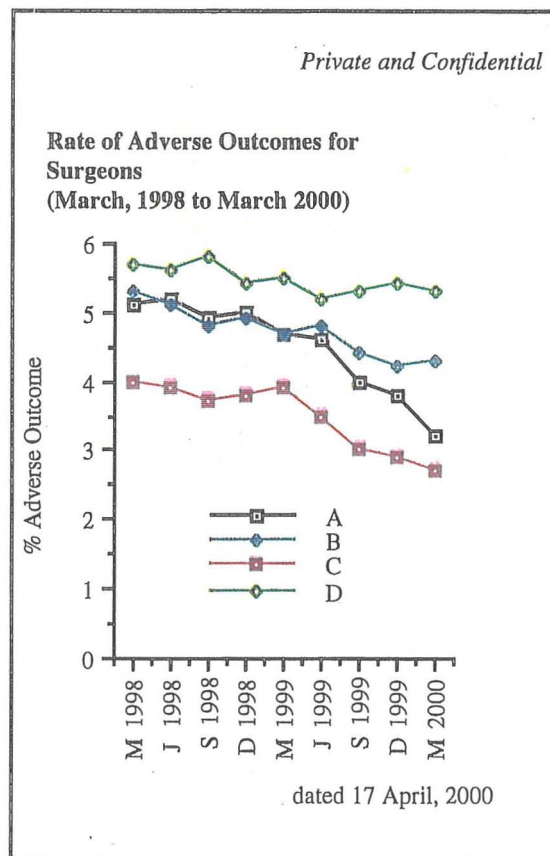


Figure 7.13: Report Format B for Peer Review

Exception reporting is a more formal technique than peer review. In general, the method used is to present data in graphical form. The results are plotting showing areas of acceptable and non-acceptable outcomes.

As an example, a measure of outcome for a given surgical procedure is the proportion of patients in a quarter reporting one or more negative outcomes relating to the procedure in their follow-up. This rate will be standardised for age, gender and other variables known to contribute to negative outcomes. An exception report will be produced if this standardised rate exceed 10% in

the quarter. Action must be taken when an exception report is made to ascertain process variables that are associated with the outlier.

As an example, the standardised rate for poor outcomes for the March Quarter might be 12% as shown in Figure 7.14. An example of a process variable that might be examined for association with outcome is surgical wound infection rate. The infection rates are examined in the same way as outcomes.

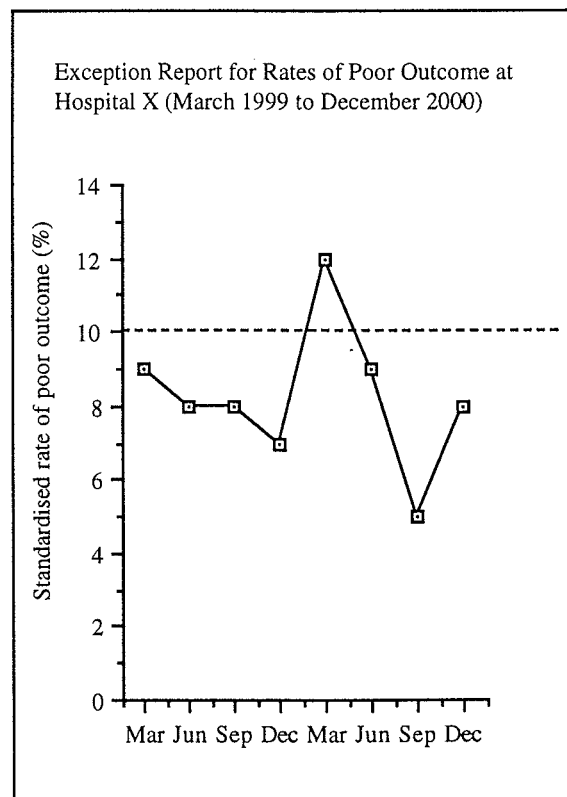


Figure 7.14: Exception Report Sample Format A

Figure 7.15 shows an association between high infection rates and standardised poor outcome rates. Action, therefore, to reduce poor outcomes would be to lower infection rate.

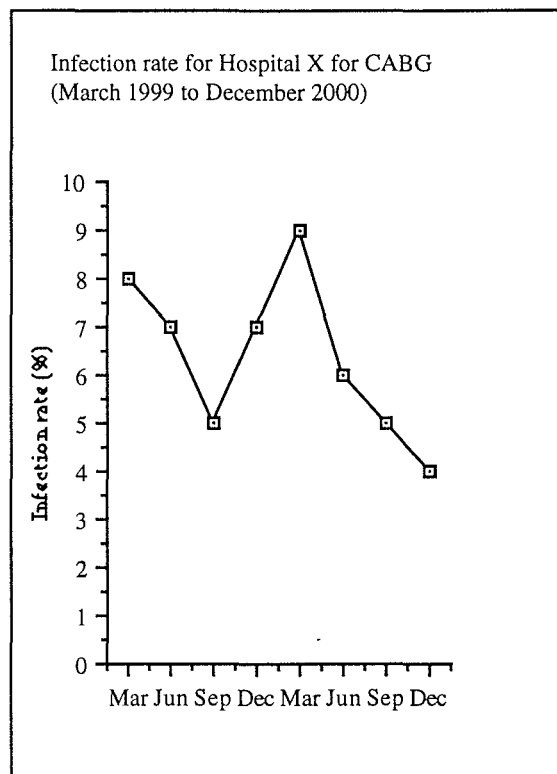


Figure 7.15: Exception Report Sample Format B

Although exception reports provide a more rigorous form of CQI than peer review, exception reports could also be used informally as part of the peer review process.

The US Department of Health and Human Services, through their Agency for Health Care Policy and Research (AHCPR), has developed a number of *practice guidelines* for high volume/high cost procedures. These guidelines have been developed in consultation with expert panels. An example of the consultation process in the development of guidelines is described in the manual for cataracts in adults (US Department of Health and Human Services, 1993). In developing the guidelines a multi disciplinary team of ophthalmologists, nurses, optometrists, registrars, a psychiatrist, an anaesthesiologist, a general practitioner, a clinical social worker and a patient representative was convened. The panel reviewed, firstly, the process of care and, secondly, reviewed literature in the area. Next draft guidelines were drawn up and peer reviewed. Finally, the guidelines were refined in the light of feedback from the peer review.

In the general model, it is also proposed that all health professionals seen to influence patient outcome would have an important contribution to the development of guidelines since implementation of guidelines will not be successful without their commitment. However, the procedure described above had feedback from only one patient advocate. The experiences of all patients needs to be addressed in the formulation of guidelines. The paradigm adopted by AHCPR would be used in the proposed model but the stage where literature is reviewed will place paramount emphasis on the perceived outcome of the procedure by their patients and the analysis of how the outcomes relate to the process.

7.6 Summary

All stages in the continuous quality improvement cycle for surgical processes, illustrated in Figure 7.7, have been covered in the previous sections. A detailed description of a system where the model has been successfully implemented for CABG patients has been described. A more general model to optimise patient-assessed outcomes for surgical procedures has been described for each stage of the cycle with simple examples of the application of the general model is included in these sections.

In summary, this chapter has shown that continuous quality improvement methods will occur if there is a collaborative and collegial approach from all contributors to the surgical process. The clinician must, firstly, be confident that the measure of patient-assessed outcome is appropriately adjusted for risk so that the measurement in no way disadvantaged his or her patients. Secondly, the clinician must have an understanding of how the controllable process variables are found to be associated with outcome - that is they must have some basic knowledge of test of statistical significance. Finally, success is dependent on the recognition that total quality management procedures are not punitive but of benefit to all stakeholders in the process.

The development of this model is significant because the use of patient-assessed outcomes is unique for this type of model. The potential applications of the model to other surgical procedures are elaborated fully in the Conclusions.

CHAPTER 8

CONCLUSIONS

CHAPTER 8 CONCLUSIONS

An issue facing health service providers is whether improved information systems make their services more effective. This thesis argues that the use of improved information systems can be more effective in terms of measurable patient outcomes. It has demonstrated, conclusively, that such a link between clinical information systems (procedures for the perfusion process with coronary artery bypass graft surgery patients and a continuous quality improvement model) and better health outcomes post-surgery (as measured by indices of neurological and physical deficit) can be made based on a statistical analysis of a large data set of patients' records. The database includes not only characteristics of the patient and their environment but also extends to data collected from pre-surgery, through the peri-operative procedure, to the post-surgery inpatient episode concluding with a one year after surgery follow-up of patients.

The conventional model for an information system for patients undergoing surgery (Figure 1.1) only contains a post-operative assessment - mortality, wound healing and infection - and not any patient-assessed outcome related to quality of life. Those studies that do follow-up surgical patients when they return to normal activities (for example, Pinna Pintor, *et al.*, 1992) collect data in one-off situations. The information system model proposed in this thesis recognises the importance of patient-assessed outcomes. The measures of patient-assessed outcomes are fed into a model which contains process variables related to the surgical procedures and analyses to develop protocols for surgical practice (Figure 1.2).

In order to develop such a conceptual model, a review of the literature was undertaken (Chapter 2). This covered four substantive areas. First, definitions and measurements of patient (or health) outcomes, and the core concepts and domains of health-related quality of life. The second substantive area was to review specific studies of conventional medical approaches to measuring outcomes for coronary artery bypass graft surgery. Thirdly, the health information literature which has some loose links with patient outcomes was reviewed. Finally, those health information systems which have, in only relatively recent times, attempted to make direct links with health outcomes (for example, Shortliffe, 1991) are identified.

Outcome measurement has been introduced to solve spiralling health costs (Linder, 1991, p. 23) but there is some resistance from the medical profession (Gulliford, 1992; and Sage, 1993) not least because of the difficulty in designing appropriate outcome measures (Ellwood, 1988,p. 1554) that are multi-level (Von Korff, 1992; and Williams, et al., 1992) Subjective measures of patient outcomes are even more contentious (for example, Aharony and Strasser, 1993) although research by Ware (1993) involves the search for reliable measures of the way patients perceive their health. The review of health information systems confirms that their success has focussed on cost reduction and not on patient management for better outcomes (Classen, et al., 1991).

The literature reveals very few published studies linking health outcomes to health information systems and there are no Australian studies that have been identified. The major proponent of this approach is Shortliffe (1991) who has devised an outcome management process. Shortliffe and colleagues

in Medical Informatics at Stanford University, USA have provided most of the literature in this area. This review led to research gaps being identified. The literature acknowledges that patient-assessed outcomes are important but there are few tools for their measurement. Evidence of medical effectiveness in the literature depends largely on the results of randomised clinical trials, not on observations of data collected in real-time. Those studies that follow up patients' outcomes have been conducted on an ad hoc basis not on routinely collected data.

Originally the candidate proposed to collect data but it was found that, within Australia, there was a perfusionist, at the case study hospital, who had been collecting patient-assessed outcome data since 1983. With agreement of the Research and Ethic Committee of the study hospital, these data were made available to the candidate. The purpose of the data collection was to improve health outcomes for patients, in a continuous cycle, the process and performance of coronary artery bypass graft surgery. However, no rigorous statistical analyses of these data had been completed. Because of the critical importance of understanding the source of data in any statistical analyses, Chapter 3 has described perfusion - the passage of blood through a membrane oxygenator when patients are on cardiopulmonary bypass. The configuration of the heart-lung machine at the study hospital, the cardiopulmonary bypass circuit and photographs of the set-up of the equipment were described in sufficient detail as befits any experimental-type research project. The chapter also explained the critical problem facing the perfusionist: determining the best acid-base management in real time during surgery when there is a significant time delay in obtaining results of energy plasma electrolyte analysis.

Chapter 4 described the methods used for quality improvement in the case study hospital. As the issue of the validity of the measurement of patient-assessed outcomes is contentious amongst health professionals (see, for example, Testa and Nackely, 1994) the development of any instrument requires careful documentation. Therefore the instrument, for coronary artery bypass graft surgery was critically assessed based on the purpose of the measure, level of measure, construction of measures and scaling, variability of scores, reliability and validity. The perfusionist's PC database, using ACCESS III, contains a pre-operative interview, the patient's medical record and data recorded during the perfusion process - such as arterial line blood pressure whilst on bypass, length of time on bypass and intra-operative blood gas analyses at specified times. The purpose of perfusion analysis is to see whether there is a relationship between the patient-assessed outcomes recorded one year after surgery and the peri-operative variables. The perfusionist has defined two important indices - neurological damage (NSUM) and physical damage (PSUM). Both the PSUM and NSUM scores range from a score of 0 indicating no deficit to a maximum possible score of 15 indicating substantial deficit.

For this research data were transferred from the perfusionist's PC using MACLink Plus/PC software as MAC Excel files to an external hard drive. Statistical analysis involved the initial univariate analyses (Chapter Five) and the more complex multivariate analyses (Chapter Six). Univariate analyses were carried out using SPSSX for MAC and Cricket Graph for graphics. The analyses reported in Chapter 5 included descriptive results using frequency tables and graphs, and inferential results (chi-square and t-tests) for

relationships between the outcome variables NSUM and PSUM and variables from the peri-operative process and other explanatory variables (patient's demographic, pre-operative and other relevant data from the medical record).

Chapter 5 presents a description of the characteristics of the patients followed up in this study and shows that they have a similar profile by gender and age to other reported studies. Then the results of analyses of associations of various measures with the presence or absence of neurological deficit are given. Age, length of time on bypass, the presence of angina 12 months post-surgery, whether the patient was retired and the patient's symptoms pre-surgery had significant statistical associations with the presence of neurological deficit. A similar analysis for associations with the presence of physical deficits was also presented. In this case age, gender, length of time on bypass, whether the patient was retired, the nature of previous operations, previous symptoms and socio-economic status were statistically significantly associated with the presence or absence of physical deficit.

The shortcomings of the univariate analysis, using NSUM and PSUM as binary responses, were exposed in Chapter 6, especially that such analyses fail to account for the size of the deficit score either for NSUM or PSUM and for the possible interactions between explanatory variables. To overcome such problems, ordered polytomous logistic regression was used. As this technique has been used only once to a large database of relevant medical studies (Ashby, *et al.*, 1986) Chapter 6 built up the theory using first principles for this procedure, starting with the simpler, and more widely applied, dichotomous logistic regression.

Applying the technique appropriate to ordered polytomous logistic regression to the response variable, NSUM, showed that the variables that have association with increasing neurological deficit are the presence of angina one year post-surgery, being outside the partial oxygen valley that minimises mean NSUM, an interaction between being younger and having surgery in earlier years, earlier year of surgery and a longer time on bypass. The application of the model to the response variable PSUM demonstrated increasing physical deficit is associated with increasing NSUM, increasing age, lower socio-economic status, the nature of previous symptoms, being outside the base excess valley that minimises mean NSUM and being female. It is understood that identification of explanatory variables associated with poor outcomes does not imply causality. The identification of explanatory variables associated with outcomes provides a basis for future research. Of particular interest is the influence of lower socio-economic status on physical health outcomes whilst there was no association with neurological outcome. Because patient-assessed quality of life measures are usually ordered categorical responses, the application of the multivariate technique described in Chapter 6 will have widespread relevance in health care.

Chapter 7 presents a synthesis of the philosophy of continuous quality improvement described in Chapter 4 together with rigorous statistical analyses as described in Chapters 5 and 6. A general model for total quality management (TQM) has been developed for surgical procedures. It can be modified for differences in hospitals, the hospital's staff and the surgical procedures. To do this, literature relating to the introduction of TQM, in general, and in the health industry, in particular, and its associated models for

continuous quality improvement, were reviewed. Principles for the design and use of such models were established. A model for surgical procedures that incorporates continuous quality improvement was proposed (Figure 1.2) and this has been refined with specific examples of applications of the model. It contains a cycle of pre-surgery, surgical procedures, post-surgery, follow-up action, data analyses and recommendations to change practice and protocols. The model was illustrated both with findings from Chapter 4 on a continuous quality improvement process for coronary artery bypass graft surgery and with some hypothetical results so as to explain proposed reporting systems.

Areas of particular significance arising from this research are summarised as follows. The use of multivariate statistical techniques has demonstrated that the link between patient-assessed outcomes, other clinical-based outcome measures and the surgical procedure for coronary artery bypass graft patients was not just due to chance since variables such as aging were controlled for in the model. Outcomes, both qualitative and quantitative, can be attributed to the process. The validation of self-assessed measures (which are ordinal and categorical), therefore, has provided an important contribution to the measurement of outcome. The statistical techniques developed can be generalised to other procedures. An outcome management process, where specific post-discharge measures of outcome have been used to continually develop protocols for the peri-operative process represents significant contribution in areas of both health outcomes and of health information systems.

There are a number of areas where further research could build on the findings of this thesis. There will be costs involved in routinely collecting,

analysing and incorporating the post-discharge, patient-assessed outcomes into continuous quality improvement models. Mahler and Kulik (1990) have found that active patient participation results in better health outcomes. Similarly, a report by the United States Agency for Health Care Policy and Research (1995) demonstrates, with reference to ten areas, that quality can reduce costs. Therefore, there will also be benefits to the public in terms of the resulting better health outcomes. An economic study of the cost/benefit of implementing such practices is an important area for further research.

A recent report by Shroyer (et al., 1995) describes a proposed study that will be enrolling up to 6 000 patients by the end of 1996 - the Processes, Structures and Outcomes of Care in Cardiac Surgery (PSOCS). The study was designed to ascertain linkages between processes and structures of care with risk-adjusted outcomes for this surgery. There is an opportunity for collaboration with this group with particular emphasis, firstly, on the design of instrument to measure outcome (McCarthy et al., 1995). The second area for collaboration is in the statistical analyses given the statistical procedures proposed for these data by Marshall (et al., 1995).

Patient-assessed outcomes in terms of quality of life invariably are ordered categorical responses. As demonstrated in this thesis multivariate statistical techniques have enormous applications in providing explanations of variations in patient outcomes. As argued by Griffith and his colleagues, based on lessons learned through coronary artery bypass surgery, there will be a growing likelihood of a need for outcome measures for all of the surgical sub-specialities and that surgeons must "learn the language and methods of some of these statistical models" (Griffith et al., 1995, p. 599). The

multivariate statistical techniques used in this study will inevitably play an important role in future analyses of health outcomes for surgical procedures and further case studies are planned.

For instance, there are other procedures for cardiopulmonary bypass at the study hospital where data have been collected. Other areas of priority to examine include the high cost and high volume procedures such as hip replacement. The development of surgery-specific quality improvement models in the health industry will meet the societal challenges imposed on hospitals to provide objective evidence of quality and effectiveness.

These suggestions for further research and the contents of this thesis are contributions to what may be termed "real-time science in medicine". There are two schools of thought about quality improvement in health care. One is proposed by the Cochrane Collaboration which is described as evidence-based medicine where only results of clinical trials are used for quality improvement. The other side of the debate is represented by a discussion in the editorial in JAMA (Berwick, 1996, p. 877) where "journal reviewers demand evidence of proper control and statistical analyses". Berwick contends that this ignores the role of what he describes as "real-time science" - the routine collection of data and the analysis of those results in terms of patient outcomes. In such a "science", this thesis adds weight to the argument that rigorous statistical techniques are a fundamental key to successfully unlocking the paradigm that routinely collected data can be used effectively in a cycle of continuous quality to improve health outcomes and thus the quality of life experienced by those members of the community who have undergone surgery.

BIBLIOGRAPHY

Bibliography

Aharony L. and Strasser S. (1993) "Patient satisfaction: what we know about it and what we still need to explore", Medical Care Review ,Vol. 50, No. 1, pp 49-79.

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

Akers P.A.S. (1991) "An algorithmic approach to clinical decision making", Oncology Nursing Forum, Vol. 18, No. 7, pp 1159-1163.

Alavi M. and Joachimsthaler E. (1992) "Revisiting DSS implementation research: a meta analysis of literature and suggestions for researchers", MIS Quarterly, March, pp 95-117.

Alemi F., Stephens R., Parran T., Llorens S., Bhatt P., Ghadiri A. and Eisenstein E. (1994), "Automated monitoring of outcomes: application to treatment of drug abuse", Medical Decision Making, Vol. 14, No. 2, pp 180-7.

Anderson G. M. and Lomas J. (1989) "Regionalisation of coronary artery bypass surgery: effects on access", Medical Care, Vol.27, No. 3, pp 288-296.

Anderson J.A. (1984) "Regression and Ordered Categorical Variables", Journal of the Royal Statistical Society B Vol. 46, No. 1, pp 1-30.

Anderson J.A. and Philips P.R. (1981) "Regression, Discrimination and Measurement Models for Ordered Categorical Variables", Applied Statistics, Vol. 30, No. 1, pp 22-31.

Andreassen S., Jensen I.V. and Olesen K.G. (1991) "Medical expert systems based on causal probabilistic networks", International Journal of Biomedical Computing, Vol. 28, pp 1-30.

Ashby D., Pocock S. and Shaper G. (1986) "Ordered polytomous regression: an example relating serum biochemistry and haematology to alcohol consumption", Applied Statistics, Vol. 35, No.3, pp 289-301.

Australia, Department of Human Services and Health (1994) Better Health Outcomes for Australians: National Goals, Targets and Strategies for Better Health Outcomes in the Next Century, Canberra, National Health Goals.

Australian Bureau of Statistics (1991) 1989-90 National Health Survey - User's Guide, Canberra, Australian Bureau of Statistics, Catalogue 4363.0.

Australian Bureau of Statistics (1993) Information Paper: 1991 Census Socio-Economic Indexes for Areas, Canberra, Australian Bureau of Statistics, Catalogue 2912.0.

Bartlett R.H. and Harken D.E. (1976) "Instrumentation for cardiopulmonary bypass - past, present and future", Medical Instruments, Vol. 10, pp 119-124.

Batalden R., Nelson E. and Roberts J. (1994) "Linking outcomes measurement to continual improvement: the serial 'V' way of thinking about improving clinical care", Joint Commission Journal of Quality Improvement, Vol. 20, No. 4, pp 167-80.

Black D. and Cameron I. (1993) "Focusing on the customer - the role of market research techniques", unpublished invited paper St Vincents' Hospital Division of Pathology Colloquia on Thursday, 1 July, 1993.

Black D. (1994) "Community Attitudes to NSW Health: A report for Unisearch", Kensington, Unisearch Ltd.

Black D., Grant C., Lapsley H. and Rawson G. (1994) "The social and service needs of people with multiple sclerosis in NSW, Australia", The Journal of Rehabilitation, Vol. 60, No. 4, pp 60-65.

Bergner M., Bobbit R., Carter, and Gilson B. (1981) "The Sickness Impact Profile: Development and final revision of a health status measure", Medical Care, Vol. 19, No. 8, pp 787-805.

Bernelot Moens H.J. and van der Korst J.K. (1991) "Comparison of rheumatological diagnoses by a Bayesian program and by physicians", Methods of Information in Medicine, Vol. 30, pp 187-193.

Bero L. and Rennie D. (1995) "The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care", JAMA, Vol. 274, No. 24, pp 1935-38.

Berwick D., Baker M. and Kramer E. (1992) "The state of quality management in HMOs", HMO Practice, Vol. 6, No. 1, pp 26-32.

Berwick D. (1996) "Harvesting knowledge for improvement", JAMA, Vol. 275, No. 11, pp 877-8.

Bishop Y., Fienberg S. and Holland P. (1976) Discrete Multivariate Analyses: Theory and Practice, Cambridge, Massachusetts, MIT Press.

Blum B.I. (1986) Clinical Information Systems, New York, Springer-Verlag.

Blumenthal J., Madden D., Burker E., Croughwell N., Schniebolk S. Smith R., White W., Hlatky M. and Reves J. (1991) "A preliminary study of the effects of cardiac procedures on cognitive performance", International Journal of Psychomatics, Vol 38, pp 13-16.

Booth D., Deupree R., Hultgren H., De Maria A., Scott S. and Luch R. (1991) "Quality of life after bypass surgery for unstable angina; 5-year follow-up results of Veterans Affairs Cooperative Study", Circulation, Volume 83, pp 87-95.

Bourke M. K. (1994) Strategy and Architecture of Health Care Information Systems, New York, Springer-Verlag.

Box G. E. P. (1996) "Role of statistics in quality and productivity performance", Journal of Applied Statistics, Vol. 23, No. 1, pp 3-20.

Brook R. and Appel F. (1973) "Quality of care assessment: choosing a method for peer review", New England Journal of Medicine, Vol. 288, No. 25, pp 1323-1329.

Caine N., Harrison S., Sharples L. and Wallwork J. (1991), "Prospective study of quality of life before and after coronary artery bypass grafting", British Medical Journal, Vol. 302, pp 511-517.

Carman J., Shortlee S., Foster R., Hughes E., Boerstler H., O'Brien J. and O'Connor E. (1996) "Keys for successful implementation of TQM in hospitals", Health Care Management Review, Vol. 21, No. 1, pp 48-60.

Carmel S. (1985) "Satisfaction with hospitalisation: a comparative analysis of three types of service", Social Science and Medicine, Vol. 21, No. 11, pp 1243-1249.

Carr-Hill R.A. and Morris J. (1991) "Current practice in obtaining the 'Q' in QALYs: a cautionary tale", British Medical Journal, Vol. 303, pp 699-701.

Carrier M., Pineault R., Tremblay N. and Pelletier C. (1993) "Outcome of rationing access to open-heart surgery: affect of the wait for elective surgery on patient outcome", Canadian Medical Association Journal, Vol. 149, No. 8, pp 1117-1129.

Chapman S., Smith B. and Black D. (1988) "The Great Australian Salt Search", Australian Medical Journal, Vol 149, pp 714.

Chen M. and Black D. (1993) "Cervical cancer screening rates among Chinese women in Sydney", paper presented to Public Health Association 25th Annual Conference, Sydney, October 1, 1993.

Chute C.G. (1992) "Clinical data retrieval and analysis: I've seen a case like that before", Annals New York Academy of Sciences

Classen D.C., Burke J.P., Pestonik S., Evans R.S. and Stevens L.E. (1991) "Surveillance for quality assessment: IV surveillance using a hospital information system", Infection Control and Hospital Epidemiology, Vol. 12, No. 4, pp 239-244.

Cleary P.D., Keroy L., Karapanos G. and McMullen W. (1989) "Patient assessments of hospital care", Quality Review Bulletin, Vol. 15, pp 172-179.

Coombs C., Dawes R. and Tversky A. (1970) Mathematical Psychology: An Elementary Introduction, Englewood Cliffs, New Jersey, Prentice-Hall Inc.

Corder M.P. and Ellwein L.B. (1984) "Comparison of treatments for symptomatic Hodgkin's lymphoma employing decision analysis", American Journal of Clinical Oncology, Vol. 6, pp 33-38.

Cowie C., Black D. and Ferson M. (1995) "Lead levels in pre-school children in Sydney Eastern Area", paper presented to Public Health Association 27th Annual Conference, Cairns, September 27, 1995.

Cox D. R., Fitzpatrick R., Fletcher A. E., Gore S. M., Spiegelhalter D.J. and Jones D. R. (1992) "Quality-of-life assessment: can we keep it simple?", Journal of the Royal Statistical Society Series A, Vol 155, No. 3, pp 353-393.

Davies A., Doyle M., Lansky D., Rutt W., Stevie M. and Doyle J. (1994) "Outcome assessment in clinical settings: a consensus on principles and best practices", Joint Commission Journal of Quality Improvement, Vol. 20, No. 1, pp 6-16.

Davies L. K. (1993) "Hypothermia: physiology and clinical use", in Gravlee G., Davis R. and Utley J. (eds.) Cardiopulmonary Bypass: Principles and Practice, Baltimore , Williams and Wilkins, pp. 140-154.

Davies R. (1992) "Health care researchers' needs for computer-based patient records" in Ball M. J. and Collen M. F. (eds.) Aspects of the Computer-based Patient Record, New York, Springer-Verlag, pp 46-56.

De Lone W.H. and McLean E.R. (1992) "Information systems success: the quest for the dependent variable", Information Systems Research, Vol. 3, No. 1, pp 60-95.

De Neef P.D. (1987) "Using decision analysis to communicate the importance of quality assurance studies", Quality Review Bulletin.

de Neufville R. and Stafford J. (1974) Systems Analysis for Engineers and Managers, New York, McGraw-Hill Book Company.

de Neufville R. (1990) Applied Systems Analysis: Engineering Planning and Technology Management, New York, McGraw-Hill Book Company.

Degeling P., Black D., Palmer G. and Walters J. (1995) "A national survey of the knowledge and attitudes of hospital staff about casemix", Report No. 1. A Report to the Australian Department of Human Services and Health.

Deming W.E. (1986) Out of Crisis, Cambridge, Massachusetts, Cambridge University Press.

DesHarnais S., McMahon L. and Wroblewski R. (1991) "Measuring outcomes of hospital care using multiple risk-adjusted indexes", Health Services Research, Vol. 26, No. 4, pp 425-45.

Dillard D. and Miller D. (1993) Atlas of Cardiac Surgery, New York, Macmillan Publishing Company.

Donabedian A. (1981) "Criteria, norms and standards of quality: what do they mean?", American Journal of Public Health, Vol. 71, No. 4, pp 409-412.

Edwards G., Compton P. and Malor R. (1993) "PEIRS: a pathologist-maintained expert system for the interpretation of chemical pathology reports", Pathology, Vol. 25, No. 1, pp 27-34.

Elfström J., Stubberöd A. and Troeng T. (1996) "Patients not included in medical audit have a worse outcome than those included", International Journal for Quality in Health Care, Vol. 8, No. 2, pp 153-7.

Ellwood P. (1988) "Special report - Shattuck Lecture- outcomes management", New England Journal of Management, Vol. 318, No. 23, pp 1549-1556.

Epstein A.M. (1990) "The outcome movement - will it get us where we want to go?", The New England Journal of Medicine, Vol. 323, No. 4, pp 266-270.

Evans S., Larsen R.A., Burke J.P., Gardner R.M., Meier F.A., Jacobson A., Conti M.T., Jacobson J.T. and Hulse R.K. (1986) "Computer surveillance of hospital-acquired infections and antibiotic use", Hospital-Acquired Infections, Vol. 256, No. 5, pp 1007-1011.

Ewan C., Young A., Bryant E. and Calvert D. (1994) National Framework for Environmental and Health Impact Assessment, Canberra, Australian Government Publishing Service.

Favaloro R. (1968) "Saphenous vein autograft replacement of severe segmental coronary artery occlusion", Annals of Thoracic Surgery, Vol. 5, pp 334-339.

Fienberg S.E. (1994) The Analysis of Cross-Classified Categorical Data, Cambridge, Massachusetts, The MIT Press.

Fisher K., Black D. and Ferson M. (1995) "Working parents caring for sick children: The impact of child care arrangements on childhood illness and workplace absenteeism", A Report to the Australian Department of Human Services and Health.

Fletcher A., Gore S., Jones D., Fitzpatrick R., Spiegelhalter D and Cox D. (1992) "Quality of life measures in health care II: Design, analysis and interpretation", British Medical Journal, Vol. 305, pp 1145-1148.

Fries J.F. and Spitz P.W. (1990) "The hierarchy of patient outcomes", in B. Spilker (ed.) Quality of Life Assessment in Clinical Trials, New York, Raven Press, pp. 25-36.

Furber S.E. and Black D. A.(1995) "Uptake of the Haemophilis Influenza Type b (Hib) Vaccine in Children Attending Childcare Centres", Health Promotion Journal of Australia, Vol 5, No. 2, pp. 45-48.

Gage J. S. (1993) "Probability and medical decision making - a politically incorrect view", M D Computing, Vol. 10 No. 2 pp 84-86.

Gann M. and Restuccia J. (1994) "Total quality management in health care; a view of current and potential research", Medical Care Review, Vol. 51, No. 4, pp 467-500.

Garnick D., Hendricks A. and Comstock C. (1994) "Measuring quality of care: fundamental information from administrative datasets", International Journal for Quality in Health Care, Vol. 6, No. 2, pp 163-77.

Garrow J. (1981) Treat Obesity Seriously: A Clinical Manual, London, Churchill Livinstone.

Gilbert J. and Schoolfield J. (1991) "The outcome index and system outcome score: a method of QA through outcome analysis in the special care area", International Journal of Clinical Monitoring and Computing, Vol. 8, pp 85-94.

Gilman S. (1990) "Neurological complications of open-heart surgery", Annals of Neurology, Vol. 28, No. 4, pp 475-476.

Gordon T., Burleyson G., Tielsch J. and Cameron J. (1995) "The effects of regionalization on cost and outcome for one general high-risk surgical procedure", Annals of Surgery, Vol. 221, No. 1, pp 43-9.

Govier A.V. and Reves J.G. (1989) "Cerebral blood flow: Autoregulation during cardiopulmonary bypass", in M. Hilberman (ed.) Brain Injury Protection during Heart Surgery, Boston, Martin-Nijhoff, pp 27-45.

Griffith B., Hattler B., Hardest R., Kormos R., Pham S. and Bahnson H. (1995) "The need for accurate risk-adjusted measures of outcome in surgery: lessons learned through coronary artery bypass", Annals of Surgery, Vol. 222, No. 4, pp 593-9.

Guadagnoli E. and McNeil B. (1994) "Outcomes research: hope for the future or the latest rage?", Inquiry, Vol. 31, pp 14-24.

Gulliford M. (1992) "Evaluating prognostic factors: implications for measurement of health care outcome", Journal of Epidemiology and Community Health, Vol. 46, pp 323-6.

Hagen M. (1992) "Decision analysis: a review", Family Medicine, Vol. 24, pp 349-354.

Hammermeister K., Shroyer A., Sethi G. and Grover F. (1995) " Why is it important to demonstrate linkages between outcomes of care and processes and structures of care.", Medical Care, Vol. 33, No. 10, pp OS5-OS16.

Harmon P. and Sawyer B. (1990) Creating Expert Systems for Business and Industry, New York, John Wiley and Sons Inc.

Harrison M., Schneidau A., Ho R., Smith P., Newman S. and Treasure T. (1989) "Cerebrovascular disease and functional outcome after coronary artery bypass surgery", Stroke, Vol. 20, No. 2, pp 235-7.

Harvey R. (1991) "Making it better: strategies for improving the effectiveness and quality of health services in Australia", National Health Strategy Background Paper, No. 8, October (Canberra).

Hatcher M.E., Green L.W., Levine D.M. and Flagle C.E. (1986) "Validation of a decision model for triaging hypertensive patients to alternative health education interventions", Science and Medicine, Vol. 8, pp 813-819.

Hendrickson G., Kelly J.B., and Citrin L. (1991) "Computers in oncology nursing: present use and future potential", Oncology Nursing Forum, Vol. 18, No. 14, pp 715-723.

High K.M., Snider M.T. and Bashein G. (1993) "Principles of oxygenator function: gas exchange, heat transfer, and blood-artificial surface interaction", in Gravlee G., Davis R. and Utley J. (eds.) Cardiopulmonary Bypass: Principles and Practice, Baltimore, Williams and Wilkins, pp 28-54.

Hines J. and Black D. (1993) "Dementia - a challenge for rural health", Proceedings of the 2nd National Conference on Rural Health, pp 61-65 .

Hosmer D.W. and Lemeshow S. (1989) Applied Logistic Regression, New York, John Wiley and Sons.

Hovenga E. (1996) "Health care services and information systems", in Hovenga E., Kidd M. and Cesnik B (eds), Health Informatics: An Overview, Melbourne, Churchill Livingstone.

Jaquet G.J. (ed.) (1995) "Using patient input in a cycle for performance improvement", Joint Commission Journal on Quality Improvement, Vol. 21, No. 2, pp 87-96.

Johnston M., Langton K., Haynes B. and Mathieu A. (1994) "Effects of computer-based clinical decision support systems on clinician performance and patient outcome: A critical appraisal of research", Annals of Internal Medicine, Vol. 120, No. 2, pp 134-142.

Kaluzny A., McLaughlin C. and Simpson K. (1992) "Applying total quality management concepts to public health organisations", Public Health Reports, Vol. 107, No. 3, pp 257-64.

Kaplan R. (1994) "An outcomes-based model of directing decisions in women's health care", Clinical Obstetrics and Gynecology, Vol. 37, No. 1, pp 192-206.

Kazandjian V., Lawthers J., Cernak C. and Pipesh F. (1993) "Relating outcomes to processes of care: the Maryland hospital association's quality indicator project", The Joint Commission Journal of Quality Improvement, Vol 19, No. 11, pp 530-8.

Keller C. (1991) "Seeking normalcy: the experience of coronary artery bypass surgery", Research in Nursing and Health, Vol. 14, pp 173-178.

Keller R.B., Chapin A.M. and Soule D.N. (1990) "Informed inquiry into practice variations: the Maine Medical Assessment Foundation", Quality Assurance in Health Care, Vol. 2, No. 1, pp 69-75.

Kerr J.K. and Jelinek R. (1990) "Impact of technology in health care and health administration: Hospitals and alternative care delivery systems", The Journal of Health Administration Education, Vol. 8, No. 1, pp 5-10.

Koch G.G., Amara I.A. and Singer J. M. (1985) "A two-stage procedure for the analyses of ordered categorical data" in Sen P. K. (ed.) Biostatistics: Statistics in Biomedical Public Health and Environmental Science, North Holland, Elsevier Science Publishers, B.V. pp 357-87.

Knaus W. (1988) "Predicting and evaluating patient outcomes", Annals of Internal Medicine, Vol. 109, No. 7, pp 521-522.

Kresel J.J., Hutchings H.C., MacKay D.N., Weinstein M.C., Read J.L., Taylor-Halvorsen K. and Ashley H. (1987) "Application of decision analysis to drug selection for formulary addition", Hospital Formulary, Vol. 122, pp 658-676.

Krivenko C. and Chodroff C. (1994) "The Analysis of clinical outcomes: getting started in benchmarking", The Joint Commission Journal of Quality Improvement, Vol. 20, No. 5, pp 260-6.

Kuperman G., James B., Jacobsen J. and Gardner R. M. (1991) "Continuous quality improvement applied to medical care: experiences at LDS Hospital", Medical Decision Making, Vol. 11, No. 4, pp S60-S65.

Lau F., Vincent D., Fenna D., Goebel R. and Modry D. (1991) "Designing an outcome-oriented computer decision support system for cardiovascular ICU - a preliminary report", Journal of Medical Systems, Vol. 15, No. 5, pp 359-377.

Lawson J.S. and Black D.A. (1993a) "Socio-economic status - the prime determinant of premature death in Australia", Journal of Biosocial Science, Vol. 25, pp 539-552.

Lawson J.S. and Black D. (1993b) "The Socio-economic and health link", New Doctor, No. 60, p 7.

Lawson J.S. and Black D.A. (1995) "Education for health services management - a student survey of an Australian program", Journal of Health Administration Education, Vol. 13, No. 2, pp 335- 343.

Lawson J., Tutt D. and Black D (1992) "Colossal alcohol consumption among some Australian rugby footballers - implications for health promotion", Health Promotion Journal of Australia, Vol. 2, No.3

Lee C. (1973) Models in Planning: An Introduction to the Use of Quantitative Models in Planning, Oxford, Pergamon Press.

Lehmann H.P. and Shortliffe E.H. (1991) "THOMAS: Building Bayesian statistical expert systems to aid in clinical decision making", Computer Methods Programs in Biomedicine, Vol. 35, pp 251-260.

Levy A.H. and Lawrance D.P. (1992) "Data acquisition for the computer-based patient record", in Ball M. J. and Collen M. F. (eds.) Aspects of the Computer-based Patient Record, , New York, Springer-Verlag, pp 125-139.

Lillehei C. W. (1993) "Historical developments of cardiopulmonary bypass" in Gravlee G., Davis R. and Utley J. (eds.) Cardiopulmonary Bypass: Principles and Practice, Baltimore, Williams and Wilkins, pp 1-27.

Lindal E. (1990) "Post-operative depression and coronary bypass surgery", International Disability Studies, Vol. 12, pp 70-74.

Linder J.C. (1991) "Outcomes measurement: compliance tool or strategic initiative?", Health Care Management Review, Vol. 16 No. 4 , pp 21-33.

Lipsitz S. R., Fitzmaurice G. M. and Molenberghs G. (1996) " Goodness-of-fit tests for ordinal regression models", Applied Statistics - Journal of the Royal Statistical Society Series C, Vol. 45, No. 2, pp 175-190.

Louagie Y., Jamart J., Buche M., Euchler P., Schoevaerds D., Collard E., Gonzales M., Marchandise B. and Schoevaerds J-C. (1995) "Operation for unstable angina pectoris: factors influencing adverse in-hospital outcome", Annals of Thoracic Surgery, Vol. 59, pp 1141-49.

Luce R., Bush R. and Galanter E. (1963) Handbook of Mathematical Psychology, New York, Wiley and Sons.

Luft H. (1980) "The relation between surgical volume and mortality: an exploration of causal factors and alternative models", Medical Care, Vol. 19, No. 9, pp 940-959.

McCarthy M., Shroyer A., Sethi G., Moritz T., Henderson W., Grover F., London M., Gibbs J., Lansky D., Miller D., Clark J. and Hammermeister K. (1995) "Self-reported measures for assessing treatment outcomes in cardiac surgery patients", Medical Care, Vol. 33, No. 10, pp OS76-85.

McCullagh P. (1980) "Regression models for ordinal data", Journal of the Royal Statistical Society, Vol. 42, No. 2, pp 109-142.

McCullagh P. M. and Nelder J. A. (1989) Generalised Linear Models, Chapman and Hall

McDonald C.J. (1992) "Physicians' needs for computer-based patient records", in Ball M. J. and Collen M. F. (eds.) Aspects of the Computer-based Patient Record, New York, Springer-Verlag, pp 3-11.

McGee H., Graham T., Crowe B. and Horgan J. (1993) "Return to work following coronary artery bypass surgery or percutaneous transluminal coronary angioplasty", European Heart Journal, Vol. 14, pp 623-28.

Macklin J. (1992) "Enough to make you sick: How income and environment affect health", National Health Strategy Research Papers No 1, Canberra.

Mahler H. and Kulik J. (1990) "Preferences for health care involvement, perceived control and surgical recovery: A prospective study", Social Science and Medicine, Vol. 31, No. 7, pp 743-751.

Marsden J. and Pingry D (1993) "Theory of decision support systems portfolio evaluation", Decision Support Systems, Vol. 9, pp 183-199.

Marshall G., Henderson W., Moritz T., Shroyer A., Grover F. and Hammermeister K. (1995) "Statistical methods and strategies for working with large data bases", Medical Care, Vol. 33, No. 10, pp OS35-OS42.

Martin J.B. (1990) "The environment and future of health information systems", The Journal of Health Administration Education, Vol. 8, No. 1, pp 11-24.

Mirin S.M. and Namerow M.J. (1991) "Why study treatment outcomes?", Hospital and Community Psychiatry, Vol. 42, No. 10, pp 1007-1013.

Moss M. and O'Connor S. (1993) "Outcomes management in perioperative services", Nursing Economics, Vol. 11, No. 6, pp 364-9.

Motwani J., Sower V. and Brashier L. (1996) "TQM in the health sector", Health Care Management Review, Vol. 21, No. 1, pp 73-82.

Murkin J.M. (1989) "Cerebral hyperfusion during cardiopulmonary bypass: the influence of PaCO₂", in M. Hilberman (ed.) Brain Injury Protection during Heart Surgery, Boston, Martin-Nijhoff, pp 47-66.

New South Wales Department of Health (1993) NSW Health Department Public Health Bulletin, Vol. 4, No. 5.

Nissen T. (1993) "The information solution: public, private efforts seek to improve care, reduce costs through data collection", Minnesota Medicine, Vol. 76, pp 14-20.

Nord E. (1992) "An alternative to QALYs: the saved young life equivalent (SAVE)", British Medical Journal, Vol. 303, pp 875-877.

Newman S. (1989) "The incidence and nature of neuropsychological morbidity following cardiac surgery", Perfusion, Vol. 4 No. 2, pp 93-100.

Newman S., Klinger L., Venn G., Smith P., Harrison M. and Treasure T. (1989) "Subjective reports of cognition in relation to assessed cognitive performance following coronary artery bypass surgery", Journal of Psychosomatic Research, Vol. 33, No. 2, pp 227-33.

Newman S., Klinger L., Venn G., Smith P., Harrison M. and Treasure T. (1990) "The persistence of neuropsychological deficits twelve months after coronary artery bypass surgery", in Willner A. and Roderwald G. (eds.) Impact of Cardiac Surgery on the Quality of Life: Neurological and Psychological Aspects, New York, Plenum Press, pp 173-9.

Nykänen P., Chowdhury S. and Wigertz O. (1991) "Evaluation of decision support systems in Medicine", Computer Methods and Programs in Biomedicine, Vol. 34, pp 229-238.

O'Connor G., Plume S., Olmstead E., Morton J., Moloney C., Nugent W., Hernandez F., Clough R., Leavitt B., Coffin L., Marrin C., Wennberg D., Birkmeyer J., Charlesworth D., Malenka D., Quentin H. and Kasper J. (1996) "A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery", JAMA, Vol. 275, No. 11, pp 841-6.

Opit L. (1991) "The measurement of health service outcomes", in Holland, Detels and Knox (eds.) Oxford Textbook of Public Health, Oxford, Oxford University Press, pp 159-72.

Owens D.K. and Sox H.C. (1990) Medical decision making: probabilistic medical reasoning", in Shortliffe E., Perreault E.L., Wiederhold G. and Fagan L.M. (eds.) Medical Informatics: Computer Applications in Health, Massachusetts, Addison Wesley Publishing Company, pp 70-116.

Owens D.K. and Nease R.F. (1993) "Outcome-based practice guidelines: a method for structuring problems and synthesising evidence", Joint Commission Journal of Quality Improvement, Vol. 19, No. 7, pp 248-63.

Oxford Reference Concise Medical Dictionary (1992) Oxford Reference Concise Medical Dictionary, Oxford, Oxford University Press.

Pascoe G. C. (1983) "Patient satisfaction in primary health care", Evaluation and Program Planning, Vol. 6, pp 185-210.

Patrick D. and Erickson P. (1993), Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation, New York, Oxford University Press.

Pauker S.G. and Pauker S.P. (1987) "Prescriptive models to support decision making in genetics", Birth Defects, Vol. 23, No. 2, pp 279-296.

Peterson B. and Harrell F.E. (1990) "Partial proportional odds models for ordinal response variables", Applied Statistics- Journal of the Royal statistical Society Series C, Vol. 39, No. 2, pp 205-217.

Perry C.A. (1990) "Knowledge bases in medicine: A review", Bulletin of the Medical Librarian Association, Vol. 78, No. 3, pp 271-282.

Pinna Pintor P., Torta R., Bartolozzi S., Borio R., Caruzzo E., Cicolin A., Giammaria M., Mariana F., Ravarino G., Triumbari F., Alfieri O and Ravizza L. (1992) "Clinical outcome and emotional behaviour status after isolated coronary surgery", Quality of Life Research, Vol. 1, pp 177-185.

Pryor D. and De Long E. (1994) "The programmed outcomes research teams (PORTs) and implications for clinical practice", The American Journal of Cardiology, Vol. 73, pp 34b-38b.

Pryor T. (1994) "The use of medical logic models at LDS hospital", Computing Biology and Medicine, Vol. 24, No. 5, pp 391-5.

Redeker (1992) "The relationship between uncertainty and coping after coronary bypass surgery", Western Journal of Nursing Research, Vol. 14, No. 1, pp 46-68.

Reinertsen J. (1994) "Outcomes management and continuous quality improvement: The compass and the rudder", Quality Review Bulletin, January, pp 5-7.

Renwick M. and Sax S. (1991) "Variations in practice patterns", The Medical Journal of Australia, Vol. 155, pp 359-360.

Rigby E., O'Connor K. and Black D. (1994) "Accounting education for health service managers: finding out what they need to know", Journal of Health Administration Education, Vol. 12, No. 1, pp 59-69.

Rind D., Safran C., Phillips R., Wang Q., Calkins D., Delbanco T., Bleich H. and Slack W. (1994) "Effect of computer-based alerts on the treatment and outcome of hospitalised patients", Archives of Internal Medicine, Vol. 154, pp 1511-17.

Rosenstein A. (1994) "Cost-effective health care: tools for improvement", Health Care Management Review, Vol. 19, No. 2, pp 53-61.

Rosser R. and Kind P. (1978) "A scale of valuations of illness: Is there a social consensus?", International Journal of Epidemiology, Vol. 7, No. 4, pp 347-358.

Rotem A., O'Connor K., Black D., Bauman A., Dewdney J. and Hodgkinson A. (1995) "The Public Health Workforce Education and Training", a report to the the Australian Department of Human Services and Health.

Safran C., Rind D., Davis R., Ives D., Sands D., Currier J., Slack W., Cotton D. and Makadon H. (1996) "Effects of a knowledge-based electronic patient record on adherence to practice guidelines", M.D. Computing, Vol. 13, No. 1, pp 55-63.

Sager M., Franke T., Inouye S., Landefeld S., Morgan T., Rudberg M., Siebens H. and Winograd C. (1996) "Functional outcomes of acute medical illness and

hospitalization in older persons", Archives of Internal Medicine, Vol.156, pp 645-52.

Sager J.T., Haug P.J., Turner C.W. and Hebertson C.W. (1991) "Developing models to evaluate pregnancy outcomes", Medical Decision Making, Vol. 11, No. 4, pp S33-S36.

Sarnquist F.H. (1989) "Neurological Outcomes after 'low flow low pressure' cardiopulmonary bypass", in M. Hilberman (ed.) Brain Injury Protection during Heart Surgery, Boston, Martin-Nijhoff, pp 13-25.

SAS Institute Inc. (1986) SUGI Supplemental Library User's Guide Version 5 ed., North Carolina, Cary, pp 269-293.

Schipper H.S., Clinch J. and Powell V. (1990) "Definitions and Conceptual Issues", in B. Spilker (ed.) Quality of Life Assessment in Clinical Trials, New York, Raven Press, pp 11-24.

Schwartz C., Cole B. and Gelber R. (1995) "Measuring Patient-centered Outcomes in Neurological Disease: extending the Q-TWIST method", Archives of Neurology, Vol. 52, pp 754-62.

Scott A. and Smith R. (1994) "Keeping the customer satisfied: issues in the interpretation and use of patient satisfaction surveys", International Journal for Quality for Health Care, Vol. 6, No. 3, pp 353-9.

Sergeant P., Lesaffre E., Flameng W., Suy R. and Blackstone E. (1991) "The return of clinically evident ischemia after coronary artery grafting", European Journal of Cardiothoracic Surgery, Vol. 5, pp 447-457.

Shanon I. (1992) "Managing outcomes with on-line expertise", Healthcare Informatics, Vol. 1, pp 60-62.

Shaw P., Bates D., Cartlidge N., Heaviside D., Julian D. and Shaw D. (1985) "Early neurological complications of coronary artery bypass graft surgery", British Medical Journal, Vol. 291, pp 1384-1386.

Shaw P.J. (1989) "The incidence and nature of neurological morbidity following cardiac surgery: a review", Perfusion, Vol. 4, No. 2, pp 83-92.

Shortell S., O'Brien J., Carman J., Foster R., Hughes E., Boerstler H., and O'Connor E. (1995) "Assessing the impact of continuous quality improvement/total quality management" concept versus implementation", Health Services Research, Vol. 30, No. 2, pp 377-401.

Shortliffe E.H. (1991) "Medical informatics and clinical decision making: the science and the pragmatics", Medical Decision Making, Vol. 11, No. 4, pp S2-S14.

Shortliffe E. H., Tang P. C., Amatayakul M. K., Cottington E., Jencks S.F., Martin A., MacDonald R., Morris T. Q. and Nobel J.J. (1992), "Future vision and dissemination of computer-based patient records" in Ball M. J. and Collen

M. F. (eds.) Aspects of the Computer-based Patient Record, New York, Springer-Verlag, pp 46-56.

Shroyer A., London M., VillaNeuva C., Sethi G., Marshall G., Moritz T., Henderson W., McCarthy M., Grover F. and Hammermeister K. (1995) "The processes, structures and outcomes of care in cardiac surgery study protocol", Medical Care, Vol. 33, No. 10, pp OS17-OS25.

Sigurgeirsson B. (1991) "CANEST: A microcomputer program for estimating cancer in a cohort", Computer Methods and Programs in Biomedicine, Vol. 35, pp 193-201.

Silverman W. (1995) "Effectiveness, efficiency ... and subjective choice", Perspectives in Biology and Medicine, Vol. 38, No. 3, pp 480-95.

Sjöland H., Wiklund I., Caidahl K., Haglid M., Westberg S. and Herlitz J. (1996) "Improvement in quality of life and exercise capacity after coronary bypass surgery", Archives of Internal Medicine, Vol. 156, pp 265-71.

Sonneberg F.A. and Eckman M.H. "Health outcomes research and its interfaces with medical decision making", Medical Decision Making, Vol. 11, No. 4, pp S15-S31.

Snell E.J. (1964) "A scaling procedure for ordered categorical data", Biometrics, Vol. 20, pp 592-607.

Spiehler V.R. (1989) "Computer-assisted interpretation in forensic-toxicology: Morphine involved deaths", Journal of Forensic Science, Vol. 34, No. 5, pp 1104-1115.

Spotnitz W., Sanders R., Hanks J., Nolan S., Tribble C., Bergin J., Zacour R., Abbott R. and Kron I. (1995) "General surgical complications can be predicted after cardiopulmonary bypass", Annals of Surgery, Vol. 221, No. 5, 489-97.

Sprague R.H. and Watson H.J. (1989) Decision Support Systems: Putting Theory into Practice, Englewood Cliffs, New Jersey, Prentice Hall.

Stearns S.C. (1991)"Hospital discharge decisions, health outcomes, and the use of unobserved information on casemix severity", Health Services Research, Vol 26, No. 1, pp 27-51.

Stein D., Troudart T., Hymowitz Z., Gotsman M. and Kaplan de-Nour A. (1990) "Psychosocial adjustment before and after coronary bypass surgery", International Journal of Psychiatry in Medicine, Vol. 20, No. 2, pp 181-192.

Stitt F.W., Ying L., Dickinson G.M. and Klimas N.G. (1991) "Automated severity classification of AIDS hospitalisations", Medical Decision Making, Vol. 11, No. 4, pp S41-S45.

Tanenbaum S. (1994) "Knowing and acting in medical practice: The epistemological politics of outcomes research", Journal of Health Politics, Policy and Law, Vol. 19, No. 1, pp 27-44.

Tarlov A. R., Ware J.E., Greenfield S., Nelson E.C., Perrin R. and Zubkoff M. (1989) "The Medical Outcome Study: An application of methods for monitoring the results of medical care", Journal of American Medical Association, Vol. 262, No. 7, pp 925-930.

Tarlov A. (1996) "Status of the Patient Outcomes Research Teams (PORTs): An interview with Richard J. Greene, MD, PhD", Medical Outcomes Trust Bulletin, Vol. 4, No. 2, p 3.

Testa M. and Nackley J. (1994) "Methods for quality of life studies", Annual Review of Public Health, Vol. 15, pp 539-59.

Turban E. (1988) Decision Support and Expert Systems: Managerial Perspectives, New York, MacMillan Publishing Company.

Underwood M., Firmin R. and Jehu D. (1993) "Aspects of psychological and social morbidity in patients awaiting coronary artery bypass grafting", British Heart Journal, Vol. 69, pp 382-384.

United States Agency for Health Care Policy and Research (1995) Better Quality Can Cost Less: The Evolving Role of AHCPR - Interim Report to The National Advisory Council, AHCPR Publication No. 95-R011.

US Department of Health and Human Services (1993) Manual for Cataracts in Adults Rockville, Maryland, AHCPR Publications.

Veldhuyzen Van Zanten S. (1991), "Quality of life as an outcome measure in randomised clinical trials: An overview of three general medical journals", Controlled Clinical Trials, Vol. 12, pp 234S-242S.

Venn G.E. (1989) "Cerebral vascular autoregulation during cardiopulmonary bypass", Perfusion, Vol. 4, No. 2, pp 101-104.

Von Korff M., Koepsell T., Curry S. and Diehr P. (1992) "Multi-level analysis in epidemiological research on health behaviours and outcomes", American Journal of Epidemiology, Vol. 135, No. 10, pp 1077-1082.

Walton M. (1986) The Deming Management Method, New York, Perigee Books.

Walker S.H. and Duncan D.B. (1967) "Estimation of the probability of an event as a function of several independent variables", Biometrika, Vol. 54, pp 167-178.

Wannathemee G. and Shaper A. (1991) "Self-assessment of health status and mortality in middle-aged British men", International Journal of Epidemiology, Vol. 20, No. 1, pp 239-245.

Ware J. and Sherbourne C. (1992) "The MOS 36-Item Short Form health survey 1: Conceptual framework and item selection", Medical Care, Vol. 30, No. 6, pp 473-483.

Ware J.E. (1993) "Measuring patients' views: The optimum outcome measure", The British Medical Journal, Vol. 306, pp 1429-1430.

Weisman E. and Koch N. (1989) "Progress notes: Special patient satisfaction issue", Quality Review Bulletin, Vol. 15, pp 166-167.

Wennberg J.E., Bunker J.P. and Barnes B. (1980) "The need for assessing the outcome of common medical practices", Annual Review of Public Health, Vol. 1, pp 277-295.

Wennberg J. and Gittelsohn A. (1982) "Variations in medical care among small areas", Scientific American, Vol. 246, No. 4, pp 100-111.

Wilkin D., Hallam L. and Doggett M-A. (1993) Measures of Need and Outcome for Primary Health Care, New York, Oxford University Press.

Williams D. and Grizzle J.E. (1972) "Analysis of contingency tables having ordered response categories", Journal of the American Statistical Association, Vol. 67, pp 55-63.

Williams J., Naylor C., Cohen M., Goel V., Besinski S., Ferris L., and Llewellyn-Thomas H. (1992) "Outcomes and the management of health care", Canadian Medical Association Journal, Vol. 147, No. 12, pp 1775-1786.

Willner A.E. (1990) "Can one predict the patient's postoperative psychiatric or neurological condition?" in Willner A. and Roderwald G. (eds.) Impact of

Cardiac Surgery on their Quality of Life: Neurological and Psychological Aspects, New York, Plenum Press, pp 399-410.

Wilson I. and Cleary P. (1995) 'Linking clinical variables with health related quality of life', JAMA, Vol. 273, No. 1, pp 59-65.

Wong E., Pryor T., Huff S., Haug P. and Warner H. (1994) "Interfacing a stand-alone diagnostic expert system with a health information system", Computers and Biomedical Research, Vol. 27, pp 116-129.

Wrigley N. (1990) "Information technologies and health care services: cities and new technologies", Proceedings of a Conference Nov 26-27, Paris, Organisation for Economic Cooperation and Development's Group on Urban Affairs, pp 159-182.

APPENDICES

Appendix A

SPSSX Command File for pH

```
get file='matchall.sys'.
recode ph0 ph1 ph2 ph4 (missing=0).
do if (ph0 ge 7.4 and (ph1 ge 7.34 or ph1 eq 0) and (ph2 ge 7.28 or
ph2 eq 0) and ((ph3 ge 7.28 and ph3 le
7.36) or ph3 ge 7.4 or ph3 eq 0)
and (ph4 gt 7.26 and ph4 le 7.36 or ph4 eq 0 )).
compute phflag=1.
else.
compute phflag=2.
end if.
t-test groups=phflag(1,2)/nsum.
```

Appendix B

SPSSX Command File for Partial Oxygen

```
recode po0 po1 po2 po4 (missing=0).  
do if (po0 ge 280 and (po1 ge 500  
or po1 eq 0) and (po2 ge 480 or po2 eq 0)  
and (po3 ge 450 or po3 eq 0)  
and (po4 ge 440 or po4 eq 0 )).  
compute poflag=1.  
else.  
compute poflag=2.  
end if.  
t-test groups=poflag(1,2)/nsum.
```

Appendix C

SPSSX Command File for Partial Carbon Dioxide

```
get file='matchall.sys'.
recode pco0 pco1 pco2 pco4 (missing=0).
do if ((pco0 le 39 or (pco0 ge 41 or pco0 le 47)) and (pco1 ge 59 or pco1
le 36 or (pco1 ge 41 and pco1 le 52) or pco1 eq 0) and ((pco2 ge 28 and
pco2 le 34) or (pco2 ge 37 and pco2 le 52) or pco2 eq 0)
and ((pco3 ge 28 and pco3 le
33) or (pco3 ge 39 and pco3 le 56) or pco3 eq 0)
and ((pco4 ge 38 and pco4 le 61) or pco4 eq 0 )).
compute pcoflag=1.
else.
compute pcoflag=2.
end if.
t-test groups=pcoflag(1,2)/nsum.
```

Appendix D

SPSSX Command File for Haemoglobin

```
get file='matchall.sys'.
recode hb0 hb1 hb2 hb4 (missing=0).
do if ((hb0 ge 7.2) and ((hb1 ge 3.4 and hb1 le 7.5)
  or hb1 eq 0) and ((hb2 ge 3.9 and hb2 le 7) or (hb2 ge 8.5 and
hb2 le 9.4) or hb2 eq 0)
  and ((hb3 ge 5.5 and hb3 le
7.5) or (hb3 ge 8.4 and hb3 le 9.4) or hb3 eq 0)
  and ((hb4 ge 4.6 and hb4 le 7.5) or (hb4 ge 8.4 and
hb4 le 9.3) or hb4 eq 0 )).
compute hbflag=1.
else.
compute hbflag=2.
end if.
t-test groups=hbflag(1,2)/nsum.
```

Appendix E

SPSSX Command File for Base Excess

```
get file='matchall.sys'.
recode be0 be1 be2 be4 (missing=0).
do if ((be1 ge -9 or be1 eq 0) and (be2 ge -9 or
be2 eq 0) and (be3 ge -8.2 or be3 eq 0)
and (be4 gt -5.5 or be4 eq 0 )).
compute beflag=1.
else.
compute beflag=2.
end if.
t-test groups=beflag(1,2)/nsum.
```

Appendix F

SAS LOGIST Commands for NSUM with all Variables

```
FILENAME A1 'GRP$77:[S8871572]finalgo1.doc';
DATA J1;
INFILE A1;
INPUT newpcode icode
newretir newmarrd nsum psum age newsex bmi phflag
beflag pcoflag poflag hbflag newmflag angina nyear
xc_mins bt_mins newpreil newop;
agexyear=age*nyear;
agexsex=age*newsex;
if phflag=9 then phflag=A;
if beflag=9 then beflag=A;
if poflag=9 then poflag=A;
if pcoflag=9 then pcoflag=A;
if hbflag=9 then hbflag=A;
phxbe=phflag*beflag;
hbxbe=hbflag*beflag;
poxbe=poflag*beflag;
if bmi=9 then bmi=A;
if bt_mins=9 then bt_mins=A;
if nyear < 90 then newretir=A;
if nyear< 90 then newmarrd=A;
if newpreil=0 then newpreil=A;
if icode=9 then icode=A;
if newop=0 then newop=A;
missing A;
proc logist k=4;
model nsum=newmarrd newretir angina newop
age newsex poflag phflag beflag pcoflag beflag
phxbe hbxbe poxbe
agexyear agexsex icode bmi psum nyear
bt_mins newpreil/stepwise;
run;
ENDSAS;
```

Appendix G

List File for SAS LOGIST Procedure for NSUM with All Variables

1
Tuesday, July 30, 1996 1

The SAS System

12:44

The LOGISTIC Procedure

Data Set: WORK.J1
Response Variable: NSUM
Response Levels: 5
Number of Observations: 529
Link Function: Logit

Response Profile

Ordered			
Value	NSUM	Count	
1	0	450	
2	1	39	
3	2	21	
4	3	11	
5	4	8	

WARNING: 3450 observation(s) were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Step 0. Intercepts entered:

Residual Chi-Square = 33.0551 with 20 DF (p=0.0333)

Step 1. Variable AGEXYEAR entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 1.0154 with 3 DF (p=0.7975)

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept and		Chi-Square for
		Intercept Only	Covariates	
	AIC	644.730	636.240	.
	SC	661.814	657.595	.
	-2 LOG L	636.730	626.240	10.490 with 1 DF (p=0.0012)
	Score	.	.	10.778 with 1 DF (p=0.0010)

Residual Chi-Square = 20.0706 with 19 DF (p=0.3903)

1 The SAS System 12:44
Tuesday, July 30, 1996 2

The LOGISTIC Procedure

Step 2. Variable NYEAR entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 13.7103 with 6 DF (p=0.0330)

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept and		Chi-Square for
		Intercept Only	Covariates	
	AIC	644.730	629.158	.
	SC	661.814	654.784	.
	-2 LOG L	636.730	617.158	19.572 with 2 DF
(p=0.0001)	Score	.	.	19.193 with 2 DF (p=0.0001)

Residual Chi-Square = 11.6134 with 18 DF (p=0.8665)

NOTE: No (additional) variables met the 0.05 significance level for entry into the model.

Summary of Stepwise Procedure

Chi-Square	Step	Variable		Number Removed	Score In	Wald Chi-Square	Pr > Chi Square
		Entered					
	1	AGEXYEAR		1	10.7785	.	0.0010
	2	NYEAR		2	8.5458	.	0.0035

1
Tuesday, July 30, 1996 3 The SAS System 12:44

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter		Wald Chi-Square	Pr > Chi-Square	Standardized Chi-Square	Odds Estimate
		Estimate	Error				
INTERCP1	1	-52.6474	18.1021	8.4586	0.0036	.	0.000
INTERCP2	1	-51.8608	18.0984	8.2110	0.0042	.	0.000
INTERCP3	1	-51.0640	18.0969	7.9620	0.0048	.	0.000
INTERCP4	1	-50.1695	18.0976	7.6848	0.0056	.	0.000
AGEXYEAR	1	0.00285	0.000836	11.5860	0.0007	0.228003	1.003
NYEAR	1	0.5799	0.1994	8.4630	0.0036	0.215859	1.786

Responses Association of Predicted Probabilities and Observed

Concordant = 60.4% Somers' D = 0.276
Discordant = 32.8% Gamma = 0.296
Tied = 6.9% Tau-a = 0.074
(37597 pairs) c = 0.638

Appendix H

SAS LOGIST Commands for NSUM with Selected Variables

```
FILENAME A1 'GRP$77:[S8871572]finalgo1.doc';
DATA J1;
INFILE A1;
INPUT newpcode icode
newretir newmarrd nsum psum age newsex bmi phflag
beflag pcoflag poflag hbflag newmflag angina nyear
xc_mins bt_mins newpreil newop;
agexyear=age*nyear;
agexsex=age*newsex;
if phflag=9 then phflag=A;
if beflag=9 then beflag=A;
if poflag=9 then poflag=A;
if pcoflag=9 then pcoflag=A;
if hbflag=9 then hbflag=A;
phxbe=phflag*beflag;
hbxbe=hbflag*beflag;
poxbe=poflag*beflag;
if bmi=9 then bmi=A;
if bt_mins=9 then bt_mins=A;
if nyear < 90 then newretir=A;
if nyear< 90 then newmarrd=A;
if newpreil=0 then newpreil=A;
if icode=9 then icode=A;
if newop=0 then newop=A;
missing A;
proc logist k=4;
model nsum=angina
age newsex poflag phflag beflag pcoflag beflag
phxbe hbxbe poxbe
agexyear agexsex icode bmi psum nyear
bt_mins newpreil/stepwise;
run;
ENDSAS;
```

Appendix I

List File for SAS LOGISTProcedure on NSUM with Selected Variables

1

The SAS System

12:51

Tuesday, July 30, 1996 1

The LOGISTIC Procedure

Data Set: WORK.J1

Response Variable: NSUM

Response Levels: 5

Number of Observations: 2555

Link Function: Logit

Response Profile

Ordered Value	NSUM	Count
1	0	1822
2	1	376
3	2	191
4	3	111
5	4	55

WARNING: 1424 observation(s) were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Step 0. Intercepts entered:

Residual Chi-Square = 188.0863 with 18 DF
(p=0.0001)

Step 1. Variable NYEAR entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 0.2311 with 3 DF (p=0.9724)

[illegible][illegible]

The LOGISTIC Procedure

Step 2. Variable AGEXYEAR entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 0.4730 with 6 DF (p=0.9982)

Criteria for Assessing Model Fit

Covariates	Intercept and Covariates		Chi-Square for
	Criterion	Only	
AIC	4790.312	4639.865	.
SC	4813.695	4674.940	.
-2 LOG L	4782.312	4627.865	154.447 with 2 DF
Score	.	.	151.527 with 2 DF (p=0.0001)

Residual Chi-Square = 44.7182 with 16 DF (p=0.0002)

Step 3. Variable ANGINA entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 4.5883 with 9 DF (p=0.8686)

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept and Only Covariates		Chi-Square for
		Intercept Only	Covariates	
	AIC	4790.312	4624.905	.
	SC	4813.695	4665.826	.
	-2 LOG L	4782.312	4610.905	171.406 with 3 DF
(p=0.0001)	Score	.	.	169.512 with 3 DF (p=0.0001)

Residual Chi-Square = 26.6993 with 15 DF (p=0.0313)

Step 4. Variable POFLAG entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 6.8045 with 12 DF (p=0.8703)

1 The SAS System 12:51
Tuesday, July 30, 1996 3

The LOGISTIC Procedure

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept and Only Covariates		Chi-Square for
		Intercept Only	Covariates	
	AIC	4790.312	4617.447	.
	SC	4813.695	4664.214	.
	-2 LOG L	4782.312	4601.447	180.865 with 4 DF
(p=0.0001)	Score	.	.	174.458 with 4 DF (p=0.0001)

Residual Chi-Square = 18.2603 with 14 DF (p=0.1952)

Step 5. Variable BT_MINS entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 8.2545 with 15 DF (p=0.9132)

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept and Intercept Only Covariates		Chi-Square for
		Only	Covariates	
	AIC	4790.312	4610.685	.
	SC	4813.695	4663.297	.
	-2 LOG L	4782.312	4592.685	189.627 with 5 DF
(p=0.0001)	Score	.	.	181.834 with 5 DF (p=0.0001)

Residual Chi-Square = 9.2986 with 13 DF (p=0.7500)

NOTE: No (additional) variables met the 0.05 significance level for entry into the model.

1 The SAS System 12:51
Tuesday, July 30, 1996 4

The LOGISTIC Procedure

Summary of Stepwise Procedure

Chi-Square	Step	Variable Entered	Number Removed	Score In	Wald Chi-Square	Pr > Chi Square
	1	NYEAR	1	113.5	.	0.0001
	2	AGEXYEAR	2	40.0819	.	0.0001
	3	ANGINA	3	18.1056	.	0.0001
	4	POFLAG	4	9.0872	.	0.0026

5 BT_MINS 5 8.8705 0.0029

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-11.5029	2.1593	28.3771	0.0001	.	0.000
INTERCP2	1	-10.5425	2.1573	23.8828	0.0001	.	0.000
INTERCP3	1	-9.6673	2.1566	20.0940	0.0001	.	0.000
INTERCP4	1	-8.5022	2.1586	15.5143	0.0001	.	0.000
ANGINA	1	-0.6221	0.1470	17.9189	0.0001	-0.091931	0.537
POFLAG	1	-0.4777	0.1520	9.8742	0.0017	-0.104421	0.620
AGEXYEAR	1	0.00184	0.000303	36.8089	0.0001	0.151635	1.002
NYEAR	1	0.1429	0.0228	39.2451	0.0001	0.175450	1.154
BT_MINS	1	-0.1282	0.0430	8.9088	0.0028	-0.071373	0.880

Association of Predicted Probabilities and Observed Responses

Responses

Concordant = 64.8% Somers' D = 0.306
Discordant = 34.2% Gamma = 0.309
Tied = 1.0% Tau-a = 0.141
(1507569 pairs) c = 0.653

Appendix J

SAS LOGIST Commands on PSUM with All Variables

```
FILENAME A1 'GRP$77:[S8871572]finalgo1.doc';
DATA J1;
INFILE A1;
INPUT newpcode icode
newretir newmarrd nsum psum age newsex bmi phflag
beflag pcoflag poflag hbflag newmflag angina nyear
xc_mins bt_mins newpreil newop;
agexyear=age*nyear;
agexsex=age*newsex;
if phflag=9 then phflag=A;
if beflag=9 then beflag=A;
if poflag=9 then poflag=A;
if pcoflag=9 then pcoflag=A;
if hbflag=9 then hbflag=A;
phxbe=phflag*beflag;
hbxbe=hbflag*beflag;
poxbe=poflag*beflag;
if bmi=9 then bmi=A;
if bt_mins=9 then bt_mins=A;
if nyear < 90 then newretir=A;
if nyear< 90 then newmarrd=A;
if newpreil=0 then newpreil=A;
if icode=9 then icode=A;
if newop=0 then newop=A;
missing A;
proc logist k=4;
model psum=newmarrd newretir newop
age newsex
agexyear agexsex icode bmi nsum nyear
bt_mins newpreil/stepwise;
run;
ENDSAS;
```

Appendix K

List File for SAS LOGIST Procedure on PSUM for Selected Variables

1
Tuesday, July 30, 1996 1

The SAS System

12:57

The LOGISTIC Procedure

Data Set: WORK.J1
Response Variable: PSUM
Response Levels: 5
Number of Observations: 648
Link Function: Logit

Response Profile

Ordered Value	PSUM	Count
1	0	480
2	1	73
3	2	48
4	3	23
5	4	24

WARNING: 3331 observation(s) were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Step 0. Intercepts entered:

Residual Chi-Square = 25.5339 with 13 DF (p=0.0196)

Step 1. Variable ICODE entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 5.1001 with 3 DF (p=0.1646)

Criteria for Assessing Model Fit

	Criterion	Intercept Only	Intercept and Covariates	Chi-Square for
Covariates				
	AIC	1176.506	1172.429	.
	SC	1194.402	1194.799	.
	-2 LOG L	1168.506	1162.429	6.077 with 1 DF
(p=0.0137)	Score	.	.	5.989 with 1 DF (p=0.0144)

Residual Chi-Square = 19.4987 with 12 DF (p=0.0772)

1 The SAS System 12:57
Tuesday, July 30, 1996 2

The LOGISTIC Procedure

Step 2. Variable NEWSEX entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 7.5897 with 6 DF (p=0.2697)

Criteria for Assessing Model Fit

	Criterion	Intercept Only	Intercept and Covariates	Chi-Square for
Covariates				
	AIC	1176.506	1170.453	.
	SC	1194.402	1197.296	.
	-2 LOG L	1168.506	1158.453	10.053 with 2 DF
(p=0.0066)	Score	.	.	10.299 with 2 DF (p=0.0058)

Residual Chi-Square = 15.1968 with 11 DF (p=0.1737)

NOTE: No (additional) variables met the 0.05 significance level for entry into the model.

Summary of Stepwise Procedure

	Variable	Number	Score	Wald	Pr >
Step	Entered	Removed	In	Chi-Square	Chi Square

Chi-Square

1	ICODE	1	5.9895	.	0.0144
2	NEWSEX	2	4.1505	.	0.0416

1 The SAS System 12:57
 Tuesday, July 30, 1996 3

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Variable	Parameter	Standard	Wald	Pr >	Standardized	Odds
	DF	Estimate	Error	Chi-Square	Chi-Square	Estimate Ratio
INTERCP1	1	-2.0955	1.0942	3.6674	0.0555	. 0.123
INTERCP2	1	-1.3738	1.0933	1.5790	0.2089	. 0.253
INTERCP3	1	-0.5851	1.0961	0.2849	0.5935	. 0.557
INTERCP4	1	0.1244	1.1044	0.0127	0.9103	. 1.132
NEWSEX	1	0.3788	0.1898	3.9819	0.0460	0.094839 1.460
ICODE	1	0.00249	0.00103	5.8565	0.0155	0.123554 1.002

Association of Predicted Probabilities and Observed

Responses

Concordant = 56.7%	Somers' D = 0.148
Discordant = 41.9%	Gamma = 0.150
Tied = 1.4%	Tau-a = 0.064
(90383 pairs)	c = 0.574

Appendix L

SAS LOGIST Commands for PSUM with Selected Variables

```
FILENAME A1 'GRP$77:[S8871572]finalgo1.doc';
DATA J1;
INFILE A1;
INPUT newpcode icode
newretir newmarrd nsum psum age newsex bmi phflag
beflag pcoflag poflag hbflag newmflag angina nyear
xc_mins bt_mins newpreil newop;
agexyear=age*nyear;
agexsex=age*newsex;
if phflag=9 then phflag=A;
if beflag=9 then beflag=A;
if poflag=9 then poflag=A;
if pcoflag=9 then pcoflag=A;
if hbflag=9 then hbflag=A;
phxbe=phflag*beflag;
hbxbe=hbflag*beflag;
poxbe=poflag*beflag;
if bmi=9 then bmi=A;
if bt_mins=9 then bt_mins=A;
if nyear < 90 then newretir=A;
if nyear< 90 then newmarrd=A;
if newpreil=0 then newpreil=A;
if icode=9 then icode=A;
if newop=0 then newop=A;
missing A;
proc logist k=4;
model psum=pcoflag poflag phflag beflag hbflag phxbe hbxbe poxbe age
newsex
agexyear agexsex icode bmi nsum nyear
bt_mins newpreil/stepwise;
run;
ENDSAS;
```

Appendix M

List File for SAS LOGIST Procedure for PSUM with Selected Variables

1
Monday, August 5, 1996 1

The SAS System

11:03

The LOGISTIC Procedure

Data Set: WORK.J1
Response Variable: PSUM
Response Levels: 5
Number of Observations: 2555
Link Function: Logit

Response Profile

Ordered Value	PSUM	Count
1	0	1856
2	1	337
3	2	193
4	3	102
5	4	67

WARNING: 1424 observation(s) were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Step 0. Intercepts entered:

Residual Chi-Square = 83.9707 with 18 DF (p=0.0001)

Step 1. Variable NSUM entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 12.9980 with 3 DF (p=0.0046)

Criteria for Assessing Model Fit

M1

	Criterion	Intercept Only	Intercept and Covariates	Chi-Square for
Covariates				
	AIC	4701.843	4674.911	.
	SC	4725.226	4704.140	.
	-2 LOG L	4693.843	4664.911	28.932 with 1 DF
(p=0.0001)	Score	.	.	31.194 with 1 DF (p=0.0001)
Residual Chi-Square = 52.9953 with 17 DF (p=0.0001)				

1 The SAS System 11:03
Monday, August 5, 1996 2

The LOGISTIC Procedure

Step 2. Variable AGE entered:

Score Test for the Proportional Odds Assumption
Chi-Square = 25.7401 with 6 DF (p=0.0002)

Criteria for Assessing Model Fit

	Criterion	Intercept Only	Intercept and Covariates	Chi-Square for
Covariates				
	AIC	4701.843	4664.474	.
	SC	4725.226	4699.549	.
	-2 LOG L	4693.843	4652.474	41.368 with 2 DF
(p=0.0001)	Score	.	.	43.796 with 2 DF (p=0.0001)
Residual Chi-Square = 40.2799 with 16 DF (p=0.0007)				

Step 3. Variable BEFLAG entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 31.6973 with 9 DF (p=0.0002)

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept Intercept and		Chi-Square for
		Only	Covariates	
(p=0.0001)	AIC	4701.843	4654.681	.
	SC	4725.226	4695.601	.
	-2 LOG L	4693.843	4640.681	53.162 with 3 DF
	Score	.	.	55.955 with 3 DF (p=0.0001)

Residual Chi-Square = 27.7290 with 15 DF (p=0.0233)

Step 4. Variable ICODE entered:

Score Test for the Proportional Odds Assumption

Chi-Square = 38.8664 with 12 DF (p=0.0001)

1 The SAS System 11:03
Monday, August 5, 1996 3

The LOGISTIC Procedure

Criteria for Assessing Model Fit

Covariates	Criterion	Intercept Intercept and		Chi-Square for
		Only	Covariates	
(p=0.0001)	AIC	4701.843	4648.881	.
	SC	4725.226	4695.648	.
	-2 LOG L	4693.843	4632.881	60.961 with 4 DF
	Score	.	.	63.861 with 4 DF (p=0.0001)

AIC 4701.843 4639.767 .
 SC 4725.226 4698.225 .
 -2 LOG L 4693.843 4619.767 74.075 with 6 DF
 (p=0.0001)
 Score . 76.923 with 6 DF (p=0.0001)

Residual Chi-Square = 6.5877 with 12 DF (p=0.8836)

NOTE: No (additional) variables met the 0.05 significance level for entry into the model.

Summary of Stepwise Procedure

Step	Variable Entered	Number Removed	Score In	Wald Chi-Square	Pr > ChiSquare
1	NSUM	1	31.1939	.	0.0001
2	AGE	2	12.3713	.	0.0004
3	BEFLAG	3	12.1532	.	0.0005
4	ICODE	4	7.7107	.	0.0055
5	NEWSEX	5	6.2249	.	0.0126
6	NEWPREIL	6	7.0773	.	0.0078

1 The SAS System 11:03
 Monday, August 5, 1996 5

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	0.0766	0.5956	0.0165	0.8977	.	1.080
INTERCP2	1	0.9249	0.5963	2.4059	0.1209	.	2.522
INTERCP3	1	1.7859	0.5987	8.8974	0.0029	.	5.965
INTERCP4	1	2.7578	0.6061	20.7031	0.0001	.	15.765
BEFLAG	1	-0.3773	0.1175	10.3103	0.0013	-0.073789	0.686
AGE	1	-0.0966	0.0276	12.2797	0.0005	-0.088844	0.908
NEWSEX	1	0.2871	0.1071	7.1853	0.0074	0.063505	1.333
ICODE	1	0.00141	0.000508	7.7241	0.0054	0.069164	1.001
NSUM	1	-0.2337	0.0431	29.4453	0.0001	-0.123490	0.792
NEWPREIL	1	0.0537	0.0202	7.0349	0.0080	0.065184	1.055

Responses

Association of Predicted Probabilities and Observed

Concordant = 58.9%	Somers' D = 0.186
Discordant = 40.3%	Gamma = 0.188
Tied = 0.8%	Tau-a = 0.083
(1458789 pairs)	c = 0.593

Appendix N

HOSPITAL, SYDNEY

Place Patient Identification Label Here	SURNAME OF PATIENT		HOSPITAL NUMBER	
	TITLE	CHRISTIAN OR GIVEN NAMES	AGE (YRS)	SEX (✓) M F
	ADDRESS OF PATIENT			
	POSTCODE	DATE OF BIRTH	CLASSIFICATION	MEDICAL OFFICER IN CHARGE

SEX: AGE: • WEIGHT kg HEIGHT cm BODY SURFACE AREA m²

OPERATION:

PERFUSIONIST:

ANAESTHETIST:

SURGEON:

ASSISTANT:

Haemoglobin:

K:

Creatinine:

ALLERGIES:

PREVIOUS ILLNESS & OPERATIONS:

OXYGENATOR:

PUMP PRIME (litres):

Haemaccel

Hartmann's Solution

Heparin units

Fruzemide

Other

CHECK LIST:

Oxygenator

Cardiotomy Reservoir

Tubing Circuit

Connectors

Pump Prime

Level Alarm

Pump Stopper

Check	Double Check
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

Heater Cooler

Gas Supply

Pump Occlusion

Pressure in Circuit

Check	Double Check
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

TIME: ON BYPASS

AORTIC CROSS CLAMP TIME()

URINE OUTPUT()

OFF BYPASS()

BYPASS TIME()

)

TEMPERATURE

°C

x

CO₂ FLOW RATE
mL/min

O₂ FLOW RATE
L/min

▽

BLOOD FLOW
L/min

○

Halothane %

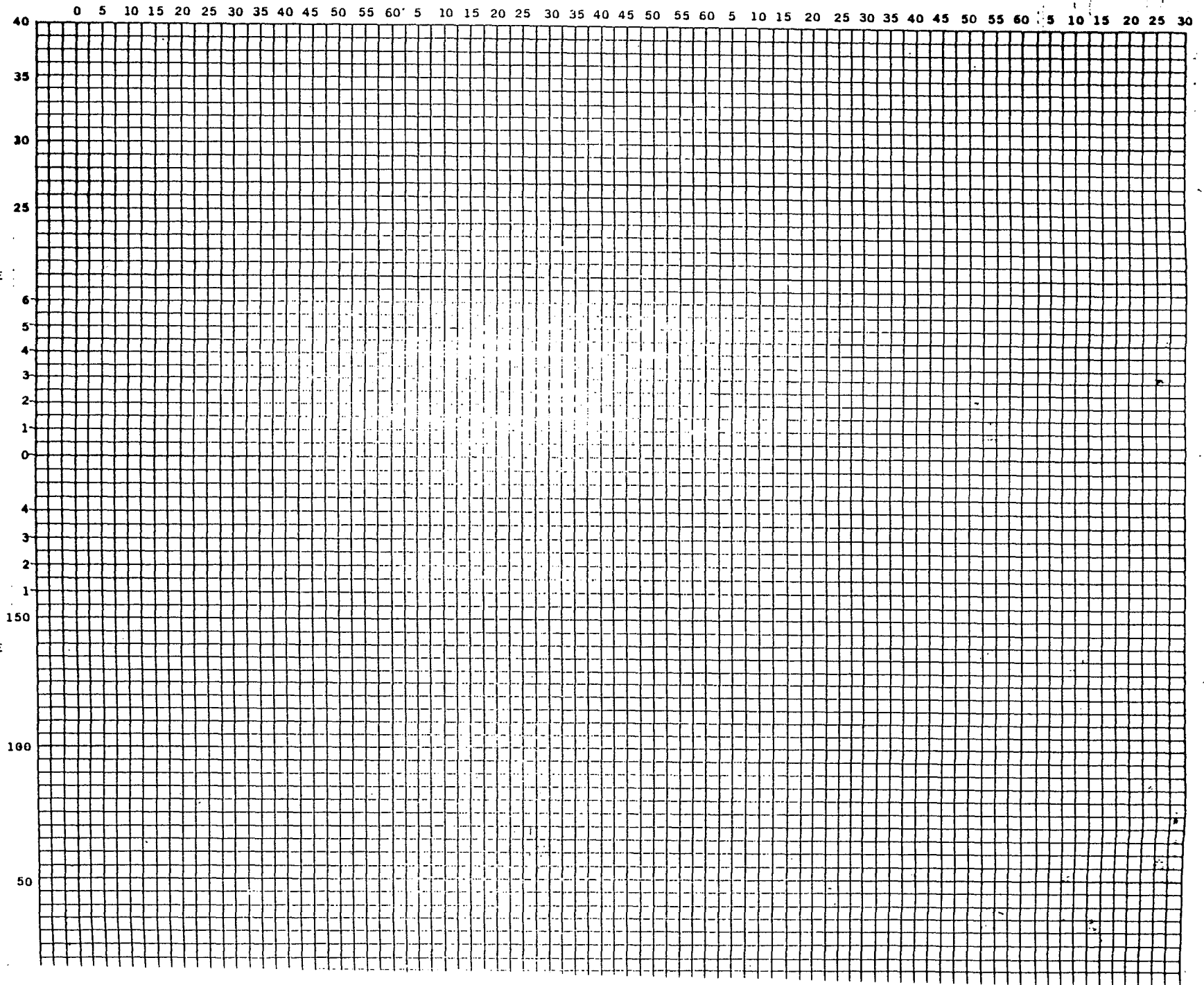
ARTERIAL LINE
PRESSURE
mmHg

•

ARTERIAL
PRESSURE
mmHg

V

VENOUS



BLOOD GASES:

TIME TEMP HB pH PCO₂ PO₂ BE SBC SAT O₂ CONT O₂ TRANSFER CO₂ TRANSFER

PRE

N3

ACTIVATED CLOTTING TIME:

TIME COLLECTED	SECONDS	HEPARIN ADDED
PRE		

FLUID ADDED DURING BYPASS (litres)

Cardioplegic solution
Haemaccel
Hartmann's Solution
Blood
Other

DRUGS ADMINISTERED DURING BYPASS

Diazepam
Fentanyl
Metaraminol
Sodium bicarbonate mmol
Sodium nitroprusside
Calcium chloride
Other

POST-OPERATIVE RECORD

Day	0	1	2	3	4	5	6	7	8	9	10
Hemoglobin											
Blood Loss											
Blood Transfusion											
Respiratory Function											
BLOOD GASES											
Coagulation Studies											
Extubated-Memory of Tube											
Fluid Balance											
Pain											
Conscious State, Memory											
Movements Sensation											
Vision Hearing											
Confusion Hallucination Dreams											

REMARKS