A system for automatic alarm limit setting in anesthesia

A System for Automatic Alarm Limit Setting in Anesthesia

J.H.M. van Oostrom

A system for automatic alarm limit setting in Anesthesia

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van
de Rector Magnificus, prof. dr. J.H. van Lint, voor
een commissie aangewezen door het College
van Dekanen in het openbaar te verdedigen op
woensdag 1 december 1993 te 16:00 uur

door

JOHANNES HUGO MARIA VAN OOSTROM
geboren te Utrecht

Dit proefschrift is goedgekeurd
door de promotoren

prof. dr. ir. J.E.W. Beneken

en

prof. J.S. Gravenstein M.D. Dr. h.c.
(University of Florida)

en door de co-promotor

dr. ir. J.J. van der Aa
(University of Florida)

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Oostrom, Johannes Hugo Maria van

A system for automatic alarm limit setting in anesthesia /
Johannes Hugo Maria van Oostrom. - [S.l. : s.n.]. - Fig., tab.
Thesis Eindhoven. - With ref. - With summary in Dutch
ISBN 90-9006582-2
NUGI 743
Subject headings: anesthesiology / anesthesiology instrumentation /
anesthesiology monitoring.

Aan mijn ouders

To Beatriz

Acknowledgments

Many people have contributed to the research that is embodied in this dissertation, and words of thanks are not enough.
Prof. J.S. Gravenstein, your clinical knowledge, ideas, enthusiasm and financial support of this research have been invaluable. Thank you.
Prof. J.E.W. Beneken, you sent me to Gainesville in 1988, and I had no idea where it was located. Now it is my home, and I thank you for the opportunity you offered me. I also thank you for your guidance, suggestions and constructive criticism that have made this dissertation what it is today.
Dr. Jan van der Aa, thank you for giving me the opportunity to do the research presented in this dissertation, for keeping me on track in finishing it, and for the opportunity to keep up my Dutch.
I also want to thank the expert anesthesiologists whose knowledge of anesthesia is represented in this dissertation: Prof. J.S. Gravenstein, Dr. G.L. Gibby, and Dr. M.L. Good. Thank you for your help and support.
My thanks also go to Christoph Westerteicher of the Hewlett Packard in Böblingen, Germany, for his interest and the financial funding of H.P. of part of this research.
The whole department of Anesthesiology has been very helpful in achieving my goals, especially the engineering group at 'the ranch', with Kelly Spaulding at the center of the office.
And finally I want to thank my wife Beatriz, who has patiently helped me get through the final stages of this dissertation, and for the love and happiness she gives me every day.

Contents

Acknowledgments ..... **i**
Contents ..... **iii**
List of Figures ..... **v**
List of Tables ..... **vii**
1. Introduction **1**
1.1. Anesthesia 1
1.1.1. The preoperative period 2
1.1.2. The anesthetic period 4
1.1.3. The postoperative period 7
1.2. Current status of Alarms in Monitoring of Patients 7
1.2.1. Literature 7
1.2.2. Overview of commercially available systems 11
1.3. Project Objective 12
1.3.1. Problem definition 12
1.3.2. Problem solution 13
1.4. Description of the chapters 13
2. Intraoperative data **15**
2.1. The anesthesia record 15
2.1.1. Manual recording 16
2.1.2. Automatic recording 18
2.2. Measuring real time data 19
2.2.1. Definition of data types 20
2.2.2. A real time data collection tool 21
2.3. Current practice of alarm settings 31
2.4 Summary 40
3. Approaches to Automatic Alarm Limit Setting **41**
3.1. Introduction 41
3.2. Expert system approach 44
3.3. Neural net approach 47
3.4. Patient clustering approach 49
3.5 Summary 51
4. Selection of patient groups **53**
4.1. Introduction 53

4.2. Grouping Patients 54
4.2.1. An introduction to clustering techniques 54
4.2.2. Selection of a clustering method 65
4.2.3. Applying the K-means clustering technique 66
4.2.4. Applying clustering techniques to patient data 69
4.3 Summary 79
5. A System for Alarm Limit Setting in Anesthesia 81
5.1 Introduction 81
5.2. Gathering Alarm Limits 81
5.2.1. Methods for limit selection 83
5.3 Group limit assignment 84
5.4 Database system design 87
5.4.1 Introduction to databases 87
5.4.2 Definition of database 90
5.4.3 Definition of data flow 92
5.5 Summary 94
6 Analysis of performance 95
6.1 Introduction 95
6.2 Validation hypotheses 95
6.2.1 Patient sample is a representative sample of the population 97
6.2.2 No differences between full and limited preoperative evaluation. 99
6.2.3 Different limits are set for different patient groups 100
6.2.4 Performance 105
6.3 Summary 108
7 Discussion and Conclusions 109
7.1 Discussion 109
7.2 Conclusions 114
7.3 Recommendations 115
Appendix A: Food Example Data 117
Appendix B: List of Abbreviations 119
References 121
Summary 127
Samenvatting 129
Curriculum Vitae 131

List of Figures

Figure 1.1: The five major components of the anesthesia machine with breathing circuit 5
Figure 2.1: Handwritten anesthetic record. 18
Figure 2.2: OSI layers. 22
Figure 2.3: Monitor communication data in monitor data definition file. 25
Figure 2.4: Setup file of the application program. 26
Figure 2.5: Data representations. 27
Figure 2.6: Signal data in monitor definition file. 27
Figure 2.7: MS Windows implementation of an intraoperative data collection tool. 30
Figure 3.1: Schematic of the alarm limit setting process. 41
Figure 3.2: Importance of preoperative data items. 44
Figure 3.3: Expert system rule base example. 45
Figure 3.4: Three-layer feedforward ANN. 48
Figure 3.5: Graph of fat contents versus energy. 50
Figure 4.1: Protein contents vs phosphor contents for twelve foods. 54
Figure 4.2: Example of PCA. 58
Figure 4.3: Projection on new axis. 59
Figure 4.4: Two-dimensional projection of food data 61
Figure 4.5: Surface plot of food data after application of a 10x10 grid. 61
Figure 4.6: Contour plot of the food data 61
Figure 4.7:Two different types of clustering. 63
Figure 4.8: Data separation in 2 groups and in 3 groups. 66
Figure 4.9: Analysis of the error vs. the number of groups for the example data. 67
Figure 4.10: Result screen of the KFRONT program. 68
Figure 4.11: Assignment of ASA importance index (ASAi). 72
Figure 4.12: Two-dimensional projection (with PCA) of the patient data 74
Figure 4.13: Three-dimensional plot of the patient data after a PCA projection on two axis. 74
Figure 4.14: Contour plot of the patient data. 74
Figure 4.15: Analysis of the error vs. the number of groups. 76
Figure 5.1: Definition of different limits. 81
Figure 5.2: Questionnaire used by the faculty and residents 86
Figure 5.3: Three level architecture of a database system 88

Figure 5.4: Table linking of the relational data model 89
Figure 5.5: Record linking of the network data model 89
Figure 5.6: Data flow. 93
Figure 6.1: Data sets and how they are used to test the hypotheses. 96
Figure 6.2: Hypothesis testing. 98
Figure 6.3: Closeness of suggested limits to actual limits 107

List of Tables

Table 1.1: Preoperative evaluation. 3
Table 1.2: Definition of ASA classification 4
Table 1.3: ASA monitoring standard 8
Table 1.4: Vendors and integrated monitoring systems. 12
Table 2.1: Anesthetic record. 16
Table 2.2: Data formats of an intraoperative record 16
Table 2.3: Currently available automatic record-keepers. 19
Table 2.4: C language basic data types 20
Table 2.5: Computerized basic data formats of intraoperative data. 21
Table 2.6: Data collection similarities and dissimilarities of intraoperative monitors. 23
Table 2.7: MS Windows communication functions. 24
Table 2.8: Parameters of RS-232 protocol. 24
Table 2.9: Additional communication parameters. 25
Table 2.10: Signal identification parameters. 27
Table 2.11: Data needed per signal. 28
Table 2.12: Defined signal identifiers. 28
Table 2.13: Data format types. 28
Table 2.14: Data structure to store signal data. 29
Table 2.15: Monitors compatible with the intraoperative data recorder with available data. 31
Table 3.1: Permanent and temporary knowledge in the example rule base. 46
Table 4.1: Definitions of different data scales. 55
Table 4.2: Eigenvalues of the food example 61
Table 4.3: Correlation of food data 62
Table 4.4: Available commands in KFRONT 68
Table 4.5: Results of clustering of food data 69
Table 4.6: Important problems from the review of systems 70
Table 4.7: Features describing a patient. 71
Table 4.8: Effects of problems. 72
Table 4.9: List of features used in the analysis. 73
Table 4.10: Results of PCA 75
Table 4.11: Correlation matrix of the data set. 75
Table 4.12: Results of PCA after elimination of two features. 76

Table 4.13: Averages and standard deviation of 5916 patients. 77
Table 4.14: Clustering results for K=8, normalized input data. 77
Table 4.15: Description of cluster groups. 77
Table 4.16: Normalized cluster center distance table. 78
Table 4.17: Standard deviations per group, per parameter. 78
Table 5.1: Methods for the collection of data on limit setting. 83
Table 5.2: Limits assigned by anesthesiologists and residents 85
Table 5.3: Description of cluster groups. 85
Table 5.4: Data tables of the database. 90
Table 5.5a: Minimal demographic data 90
Table 5.5b: Preoperative summary 91
Table 5.5c: Limit database 91
Table 5.5d: Patient groups database 91
Table 5.5e: Scaling data 92
Table 6.1: Preoperative information of the 49 patients, and the total population. 97
Table 6.2: Power of the t-test for $p<0.05$, $n=45$ 99
Table 6.3: Results of paired two-tailed t-test for limits based on limited and full preoperative evaluations. 99
Table 6.4 SNK test results for systolic blood pressure 102
Table 6.5 SNK test results for diastolic blood pressure 103
Table 6.6 SNK test results for heart rate. 103
Table 6.7 SNK test results for EtCO2 103
Table 6.8 SNK test results for SpO2 104
Table 6.9 SNK test results for all the signals 104
Table 6.10 SNK test results for all the signals after merging groups 105
Table 6.11: Limits assigned by anesthesiologists after merging groups 105
Table 6.12: Performance results 106
Table 6.13: Percentage of correct limits per signal for a simple (one group) system 106
Table 6.14: Percentage of correct limits per signal for the factory default limits of the HP Merlin 107
Table 6.15: Number of COR limits within or outside the limits suggested by our system 108
Table A.1: Contents of foods per 100 gram. 117

Introduction

This chapter is an introduction to the past and present practice of anesthesiologists and describes the present state of physiologic monitoring of patients. This introduction gives a description of the objectives of this thesis, and outlines the succeeding chapters.

1.1. Anesthesia

anesthesia (ăn'ĭs - thē'zhə) **2.** Artificially induced unconsciousness or local or general insensibility of pain. (Am. Heritage Dictionary 1981)

As the dictionary description indicates, anesthesia is induced to abolish or minimize pain. Before the invention of anesthetic drugs like ether and chloroform it was not uncommon to give a patient a good bottle of rum to make him drunk and thus sedate him somewhat. In the year 1798 Sir Humphry Davy discovered the anesthetic properties of nitrous oxide, which was used in 1844 by Dr. Horace Wells to extract a tooth from a patient. Ether was first used in 1842 by Dr. Crawford Long in Georgia, and in 1846 by Dr. Morton of Boston (Davidson 1965; Poore 1872; Raftery 1968; Green 1979). In a short time the news of the powerful potential of ether spread to Europe. Since the first reported anesthetic, anesthesia has rapidly evolved with introductions of new drugs and new equipment for administering drugs, gases, and anesthetic vapors. Although the goal of the early anesthetics was just to abolish pain, anesthetic and adjuvant drugs also provide for anxiolysis, amnesia (loss of memory), unconsciousness, and muscle relaxation (aiding both anesthetist and surgeon).

From an anesthetic perspective, a modern surgical procedure can be divided into three major parts:

- Pre-anesthetic period
- Anesthetic period
- Post-anesthetic period

This thesis will focus on those patients that require anesthesia for a scheduled surgical procedure, excluding emergency and trauma patients (5-10% of patients have anesthesia not for an operation, but for diagnostic or non-surgical therapeutic procedures). The pre-anesthetic period starts after the patient has been scheduled for

the spinal cord); and local anesthesia where only part of the body, local to the surgery, is anesthetized.

The *ASA classification* as defined by the American Society of Anesthesiologists is used to indicate the patient's physical status. Table 1.2. lists five possible classes.

Table 1.2: Definition of ASA classification (ASA 1963; Julian 1984)

Class	Description
I	no organic, physiologic, biochemical, or psychiatric disturbance.
II	mild to moderate systemic disturbance that may or may not be related to the reason for surgery.
III	severe systemic disturbance that may or may not be related to the reason for surgery. Such diseases include heart disease that limits activity, poorly controlled essential hypertension, diabetes mellitus with vascular complications, chronic pulmonary disease that limits activity, angina pectoris, history of prior myocardial infarction.
IV	severe systemic disturbance that is life-threatening with or without surgery. Examples include congestive heart failure, persistent angina pectoris, advanced pulmonary, renal, or hepatic dysfunction.
V	moribund patient who has little chance of survival but is submitted to surgery as a last resort (resuscitative effort). Examples include uncontrolled hemorrhage as from a ruptured abdominal aneurysm, cerebral trauma, pulmonary embolus.

The preoperative evaluation includes two outside information sources: 1) the patient's chart for review of previous operations, possible complications, and diagnostic tests and 2) recent laboratory data. Traditionally preoperative information is gathered with paper forms completed by anesthesiologists. Because handwritten forms are hard to read, and often incomplete, computerized versions have been studied and reported as early as 1969 (Chodoff and Helrich, 1969; Chodoff and Ginaris 1973). Chodoff and Ginaris describe a Clinical Decision Support System for the gathering and storing of patient data and for clinical decision making based on the preoperative data. A computerized preoperative evaluation has been designed and used at Shands Hospital at the University of Florida (Gibby et al. 1991a; personal communications Dr. GL Gibby 1992; Gibby et al. 1992).

1.1.2. The anesthetic period

Several types of anesthetic procedures have been described in the previous paragraph. We will describe the general anesthetic procedure in more detail.

On the day of the surgery (or the day before), the patient will typically receive pre-medication if prescribed by the anesthesiologist to relieve anxiety. When the patient arrives in the operating room physiologic monitors are connected and some baseline

readings are obtained. Anesthesia is induced by administrating intravenous drugs, volatile anesthetics (halothane, enflurane, isoflurane or others), and anesthetic gases (nitrous oxide) mixed with oxygen. Often an endotracheal tube (ET tube) is inserted into the patient's trachea to assure an adequate airway that can be used for spontaneous or mechanical ventilation. When muscle relaxant drugs are given, the respiratory muscles are relaxed and it becomes necessary to mechanically ventilate the patient. The anesthesiologist can use a combination of inhalation agents and intravenous drugs to titrate the level of anesthesia. An anesthesia machine with a breathing circuit connected to a mask or to the ET-tube is used when the patient's lungs are mechanically ventilated.

The anesthesia machine with breathing circuit

The major tasks of an anesthesia machine are: the delivery of oxygen to the patient's lungs via the breathing circuit, to deliver volatile anesthetics and anesthetic gases, to

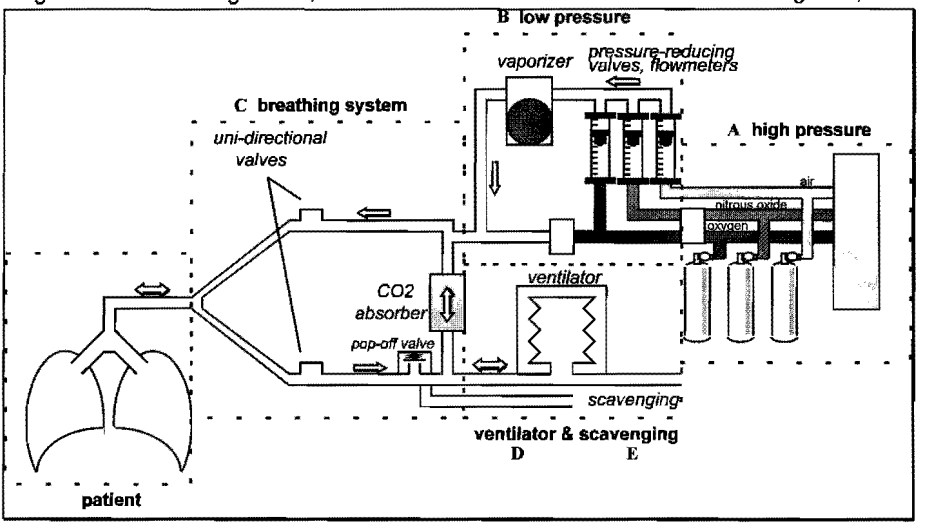


Figure 1.1: The five major components of the anesthesia machine with breathing circuit:

- *A) The high pressure section provides gases a central gas supply or, alternatively, from bottled gases.*
- *B) The low pressure section mixes the gases, meters and controls the gases and volatile anesthetics.*
- *C) The breathing circuit is the pathway for gas to move from the ventilator and low pressure circuit to the patient, and for exhaled gas to be removed from the system. The circle breathing system is the most common anesthesia breathing system in the U.S., but others like the Bain circuit are also used.*
- *D) The ventilator moves gas into the patient's lungs at controlled intervals, with controlled amounts.*
- *E) The scavenging system ensures that anesthetic gases are not vented into the operating room.*

remove CO_2, and to scavenge waste gases. Five major parts can be identified in the anesthesia machine with breathing circuit, see Figure 1.1.

First the gas enters from a high pressure central gas supply or (as a backup) from tanks (A). Typically these gases are O_2, N_2O, and air. The pressure-reducing valves of the individual gases can be adjusted, creating a lower pressure (B). The gases are then blended together and enter a vaporizer where an anesthetic agent is added. This gas combination (called the "fresh gas") enters the breathing circuit (shown in Figure 1.1 as the circle breathing circuit).

The ventilator compresses gas in a bellows (D), forcing gas through the breathing circuit (C) into the patient's lungs. A set of unidirectional valves in the circle breathing circuit causes the gas to flow in only one direction (in Figure 1.1. in counter clockwise direction). After the ventilator stops compressing, the patient exhales passively. The exhaled gas will fill up the ventilator bellows (D), and subsequently excessive gas will be scavenged by the scavenging system (E). During controlled inspiration, gas from the ventilator passes through a CO_2 absorber before entering the patient, so that the patient always breathes CO_2-free gas. An adjustable pressure limiting (APL or pop off) valve is installed in the breathing circuit to prevent the pressure in the circuit (and thus in the lungs) to exceed a pre-set limit. When this valve opens, the gas flows to the scavenging system.

Monitoring

Monitoring of the patient and the anesthesia equipment is done during the anesthetic period. Physical signs (assessment of skin color and temperature, the pupil size and reactivity to light, the nail beds, capillary filling and color, and the arterial pulse strength and heart rhythm) can be observed, or monitoring devices can be used. Monitoring devices are instruments that measures a particular entity (flow rate, pressure, voltage, etc.) and that display this entity or a derived patient variable. For example, the ECG is a measurement of voltage, but the heart rate can be derived from it. Monitoring devices can be subdivided into instruments using invasive and non-invasive methods. Non-invasive methods have the advantage of being non-intrusive to the patient (thus safer), and are usually easily applied. They can be less precise than invasive methods, because the sensor is typically some distance away from the signal source (e.g. ECG sensors are at a distance from the heart, a better measurement could be obtained by placing the sensors on the heart, which is not practical). Instruments using non-invasive methods include blood pressure measurement with a cuff, Doppler blood flow measurements, ECG, body temperature measurement, heart and lung sounds analysis

with a stethoscope, neurophysiologic monitors, respiratory monitors, and blood oxygen saturation monitoring with pulse-oximetry. Invasive monitoring methods include Foley catheter for urine output measurements, arterial blood pressure catheters, central venous and pulmonary artery catheters (Swan-Ganz catheter), and intracranial pressure transducers.
Monitoring is discussed in more detail in paragraph 1.2.

1.1.3. The postoperative period

After the operation is completed and the patient is awake and breathing spontaneously, the patient is moved to the post anesthetic care unit (PACU), where he is monitored for any adverse effects of the anesthetic and surgery. When the patient's vital signs are stabilized the patient is moved to the ward. There the anesthesiologist will perform a postoperative visit within 48 hours upon completion of the procedure, to make sure there are no adverse effects of the anesthetic. A postoperative note is made indicating the patient's physical status, and any complications that may have occurred post-operatively. Some patients go home on the day of surgery. If there are no anesthesia related problems the patient is released from the care of the anesthesiologist.

1.2. Current status of Alarms in Monitoring of Patients

The anesthesiologist's task of abolishing pain, and providing sufficient oxygen to the patient, is facilitated by physiologic monitoring of patients. Three main objectives of monitoring of patients can be defined: monitoring the correct function of the anesthesia equipment, the titration of drug effects, and monitoring the safety of the patient. Monitoring the effects of drugs and vital functions is required to assure the detection of adverse situations and to minimize the risk to the patient. Because of increasingly complex operations, and the availability of more powerful anesthetic drugs the anesthesiologist must assess the physiologic state of the patient in increasing detail.

1.2.1. Literature

Hug identifies three levels of physiologic monitoring of patients during anesthesia and surgery (Hug 1981):

1. Routine Monitoring - applicable to all patients regardless of their physiologic status.
2. Specialized monitoring - for a particular pathologic problem (e.g. serum glucose determinations in the diabetic patient) or for the use of a specialized technique. (e.g. controlled hypotension).
3. Extensive monitoring of all major systems in the critically ill patient and in those undergoing extensive surgery potentially affecting all organ and tissue functions (e.g. cardiac surgery with cardiopulmonary bypass).

Standards for routine monitoring have been defined that are currently considered appropriate and accepted as minimal standards for all patients undergoing surgery (Eichhorn 1989).

Monitoring standards

Standards for monitoring have evolved in the last five to ten years. They have usually been drafted to increase patient safety in one institution, but they have become a national standard. In the United States the minimal monitoring guidelines were originally drafted by the Harvard Medical School (Eichhorn et al. 1986), were adapted by the American Society of Anesthesiologists in 1986, and last amended in 1990 (see Table 1.3).

Table 1.3: ASA monitoring standard (ASA 1993).

monitor	interval
Oxygenation	continuous
Inspired gas oxygen concentration	
Blood oxygenation	
Ventilation	continuous
Breathing system disconnect	
End tidal CO_2 (recommended)	
Circulation	
Electrocardiogram	continuous
Blood pressure	every 5 minutes
Heart rate	every 5 minutes
Body temperature	continuous

Studies have measured the impact of monitoring standards on the number of critical incidents and overall patient safety. Eichhorn analyzed 1,001,000 ASA status I and II patients for anesthesia related intraoperative accidents at the hospitals of the Harvard Department of Anesthesia before the monitoring standards were in effect. They report 11 major intraoperative accidents, 8 of which (73%) could have been prevented by

applying the standards presented in Table 1.3. Unrecognized hypoventilation was the most common accident (Eichhorn 1989). This study suggests (although not statistically significant) that the Harvard Monitoring Standard increases patient safety, but since accidents are so rare, a study with millions of cases is needed, and that is currently not feasible. Tinker et al. report on an ongoing Closed Claims Study of the ASA Professional Liability Committee (Tinker et al. 1989). Anesthesiologists reviewed 1,097 incident cases. The reviewers report that 31.5% of the incidents could have been prevented by the use of one or more additional monitoring devices. The monitors deemed most useful in the cases of preventable injuries or deaths were pulse oximetry (40%) and capnography (2%).
In an editorial in the British Journal of Anaesthesia the editors summarize the activities towards increased patient safety by comparing monitoring standards. There is indication that the number of critical incidents is decreasing, but it is not clear that this is caused by the implementation of monitoring standards (Editorial. British Journal of Anaesthesia, March 1990; Witcher et al. 1988). It is expected that with the advent of new monitors (e.g. an Anesthetic Depth Monitor (Cluitmans 1990)) and more elaborate studies about the usefulness of particular monitors, monitoring standards will change.

Besides the definition of monitoring standards and the formation of patient safety organizations, other factors like ergonomics and alarms play a role in patient safety. We will elaborate on alarms in the following section.

Alarms

To maintain vigilance is of critical importance in anesthesia monitoring of patients. Vigilance is negatively affected by the sometimes repetitive and monotonous nature of the task of anesthetizing a patient (especially in the maintenance phase of the anesthetic procedure). Damage to the patient is possible if problems are not anticipated and recognized early. Alarm systems have been developed to aid the anesthesiologist in the task of maintenance of vigilance. But there are drawbacks to the use of these systems. They are subject to artifacts and transients, and can produce many false positives that distract the anesthesiologist from more important clinical information (Berry and Katz 1989). Kestin et al. describe the frequency of auditory alarms in a group of 50 patients and found that only 3% of the alarms represented a risk to the patient, and that 75% of the alarms did not originate from a change in physiologic variables (Kestin, Miller, and Lockhart 1988; Schaaf and Block 1989). Another study reports that from 1455 alarms recorded from 26 patients in the intensive care unit, only

1.1% were relevant and required action by the nursing staff (O'Carrol 1986). A large number of alarms (58%) were due to artifacts and minor malfunctions.
Setting a too narrow range between upper and lower limits can also produce frequent false positives, while a wide range setting may produce undesired false negatives.
A multitude of physiologic variables are routinely measured and the anesthesiologist has trouble keeping track of all the values. With so many measured variables the selection of patient specific alarms becomes a problem and is therefore frequently omitted.

Current solutions
Beneken and van der Aa analyzed the need for smart alarms in increasingly complex monitoring systems (Beneken and van der Aa 1989). They argue that simple limit based alarms are not well suited for this task, and that a knowledge based expert system (smart alarm system) would be more appropriate. The authors present five approaches to establish alarm limits: 1) organize consensus meetings of experienced anesthesiologists, 2) use *a priori* knowledge and models, 3) introduce a learning period when the conditions change, 4) make use of statistical data from categories of patients, and 5) use information derived from combining different signals.
Other authors describe the delivery of anesthesia as an engineering closed loop control system that needs monitoring of its three components: the anesthesia machine, the patient, and the controller (the anesthesiologist) (Schreiber and Schreiber 1989). They describe a critical incident as the following sequence: 1) adverse condition begins, 2) alarm generated, 3) alarm identified, 4) problem identified, 5) problem corrected, and 6) return to the safe state. The time required to complete this sequence depends in part on the alarm threshold of the alarm system (how soon does it sound the alarm), how easily the alarm can be identified (clear indication of alarm condition, complexity of the alarm system), the specificity of the alarm (is there an explanatory alarm text), and the problem itself (how long does it take to treat the problem, and for the actions to take effect). Alarm systems should therefore facilitate the rapid enunciation of the alarm and the identification of the problem.
Studies have been performed to improve alarm systems by decreasing the false alarm rate, and by making an intelligent assessment of an alarm.
Mäkivirta et al. evaluated the quality of limit alarms when variations in the monitored data were filtered by median filtering (Mäkivirta et al. 1991). Although false alarms decreased with this technique (compared to unfiltered data false alarms decreased 80% for 1 min. delay, and 93% for 2.5 min. delay) some correct alarms were missed (19% for

1 min. delay, and 53% for 2.5 min. delay). A combination of wide limits for unfiltered data and tight limits for filtered data was used to ensure that no correct alarms were missed. The number of false alarms in this dual system decreased by 63%. The authors suggest that with this method priority alarms systems can be built based on alarm, alert, and advisory categories.

Intelligent solutions to this problem have been proposed. Van der Aa (1990) described an expert system based system that monitors the integrity of the circle breathing system. This system was able to correctly detect anesthesia machine malfunctions like disconnects, leaks, stuck valves etc. in 96% of the incidents. With the inclusion of physiologic patient oriented problems like main stem intubation, hypoxic mixture, low O_2 delivery etc. the system was able to respond with a correct intelligent message in 88% of the incidents. The author recognizes that in order for alarm systems to work in the operating room, research on how alarm limits are set is required. (van der Aa 1990, p. 137). Another intelligent alarm system has been presented by Orr and Westenskow (1990). Their system, based on a neural net, was able to generate specific alarms in 95% of 746 events during controlled ventilation, and had a low false positive alarm rate of 1.7 false alarms per hour during clinical trials.

1.2.2. Overview of commercially available systems

Many integrated monitoring systems are now commercially available (see Table 1.4.) The problem of configuring the systems and setting them up with proper alarm limits has been addressed in some of these systems. At the 1992 Annual Meeting of the American Society of Anesthesiologists in New Orleans, different vendors of integrated systems were interviewed by the author with the intent to document how alarm technology was implemented in their systems. The techniques used in these systems are not always documented in the literature. Interviewing sales people does not always result in an in depth understanding of their systems, because detailed information is not always available to them. Sometimes the way alarm technology is implemented is proprietary information (e.g. when an alarm is based on several measurements that are averaged, information on how this averaging is done is often not available). Additional information comes from brochures available from the vendors of the systems and from personal observations. Table 1.4. shows vendors with the type of systems they sell.

Table 1.4: Vendors and integrated monitoring systems. See Appendix B for abbreviations.

Vendor	System type	System features
Ohmeda	CD anesthesia machine	Anesthesia machine with integrated monitoring of flows and gases (disconnect), CO_2, anesthetic agent, S_nO_2.
Spacelabs	Integrated monitoring system	Monitors for ECG, BP, CO_2, anesthetic agent, S_nO_2, PA.
Marquette	TramScope 12	Monitors for ECG, NIBP, S_nO_2, PA, CO_2
Datascope	Integrated monitor	NIBP, S_nO_2, CO_2, Agent.
Dräger	anesthesia machine	NIBP, PA, S_nO_2, CO_2, Agent
HP Merlin	Integrated monitoring system	NIBP, PA, S_nO_2, CO_2, ECG

All alarm system implementations rely on fixed limit alarms. When a previously set limit is exceeded an alarm will sound, or will be visible on the display. All monitoring systems provide for a setup mode of alarm limits by pressing keys and following menu driven instructions on the screen. The Spacelabs system defaults to fixed amount (20%) above and below baseline data, while systems from the other vendors default to factory set values. The Datascope system has an auto-set button that sets all the alarm limits to +/- 20% of the current values. The Dräger system has a similar feature, but in addition settings allow for a wide or narrow range setting. It is not clear how these settings are achieved. The Ohmeda CD provides three types of alarms: 1) advisory - a monitor is on standby, 2) alerts - alert zones can be set to 2-4-6% of the current baseline, and 3) alarms - fixed alarm limits that have no default setting.

Currently no system has the capability of adjusting for different types of patients. Some systems provide alert zones as a percentage around a baseline, but no methods are available for establishing a true baseline. In some systems a set of alarm limits can be saved and retrieved for later use, but the selection of which set of limits to choose is up to the anesthesiologist.

1.3. Project Objective

1.3.1. Problem definition

Alarms are designed to warn or alert the anesthesiologist of an untoward event. These systems do not work if the alarm limits are not set. Intraoperative alarm limits are often not set by the anesthesiologists because many monitors are used and the setting of limits is a time consuming task. Many times they are left at the factory defaults or at the values of the previous case.

Anesthesiologists frequently have a mental picture of what ranges are acceptable for a specific patient. Intelligent monitoring systems (van der Aa 1990; Westenskow 1992), Quality Assurance (QA) systems, automatic record keepers, and others also need to know these ranges of the patient variables. These systems usually compare patient variables to a 'normal' value or an acceptable range of values, but the definition of these values is primitive at best. In this study we will examine what the clinical operating ranges (COR) are for individual patients.

1.3.2. Problem solution

Alarm limits are individualized per patient based on the information obtained from the patient, primarily the preoperative evaluation (preop), and the anesthesiologist's experience. We will first analyze what data is currently used intraoperatively and what the clinical operating range (COR) of a patient looks like.

The concept of COR was defined as the range that is clinically acceptable to the anesthesiologist, with the purpose of deriving alarm limits. The COR limit is not the same as the alarm limit because transgressions of COR limits are expected to be frequent. A transgression of a COR limit can be thought of as a transgression from a green zone into a yellow zone. When an alarm limit is crossed, this is a transgression from a yellow zone into a red zone (the traffic light paradigm).

Most hospitals are moving to integrated networked computer systems that interconnect clinics, labs, and operating rooms, which makes electronic patient information available in the operating room, and provides a physical platform for an alarm system based on patient information. Current monitoring and alarm systems do not use any *a-priori* patient information. Monitoring equipment is not configured differently for different types of patients or different types of operations. This dissertation will link preoperative information with knowledge from anesthesiologists to select patient specific alarm limits. By automatically selecting alarm limits, the time consuming task of setting alarm limits is eliminated.

1.4. Description of the chapters

This chapter has given an introduction to anesthesia, and how alarm technology is currently used in the operating room. Chapter two will describe the flow of patient data in the operating room. It describes what is recorded and how, and discusses methods for automatic recording. These methods can be used to document the variability and

ranges of physiologic signals, which is related to how alarm limits should be set. Results of a review study of the use of intraoperative alarms concludes chapter two. Chapter three discusses different suggestions of implementing a system for the automatic setting of alarm limits. The selected method is presented in depth in chapter four, and includes selection of input data, data analysis, and clustering of patients into groups. Methods for the collection of the anesthesiologist's knowledge on alarm limit setting, and the design and implementation of a database for the storage and retrieval of all this information are presented in chapter five. Chapter six discusses hypotheses that need evaluation for validation of the system, and presents results of the evaluations. This thesis is completed with conclusions and recommendations in chapter seven.

Intraoperative data

Many data are being measured in the Operating Room (O.R.). These data are recorded for clinical, operational, and legal reasons (Gravenstein 1989; Ream 1989; Jackson 1989; Gibbs 1989; Peters 1989; Linnarsson and Hallén 1987). Clinical reasons include 1) the display of trends, 2) to remember the anesthetic management during the case (e.g. how much of which drug was given), 3) to share intra-operative information between colleagues, or from the O.R. to the recovery room, intensive care unit, or postsurgical ward, 4) peer review to ensure quality of care, and 5) for review of previous anesthetics. Operational reasons include 1) generation of billing information for patient billing, 2) generation of operational statistics, and 3) generation of statistics on resident training. Legal reasons are 1) documentation of the quality of care, and 2) documentation for legal defense. In an educational institution like a teaching hospital, there are also scientific reasons to record intraoperative data for studies of intraoperative incidents, effects of drugs, variability of physiologic patient variables, etc. Intraoperative data are recorded on an anesthesia record.

To determine if we can define alarm limits based on a detailed log of what happens intraoperatively, combined with preoperative information about a patient, we proposed to record intraoperative real-time data. A data collection tool was designed and built, based on a method we developed to record data transparently from any physiologic monitor. This is described in paragraph 2.2.
To determine what the clinical operating ranges of physiologic patient parameters are, we studied 50 patients intraoperatively. This study is presented in paragraph 2.3.

2.1. The anesthesia record

The data currently recorded onto the intraoperative record come from different sources, and are in different formats. The different data items of a typical anesthetic record, their format, and source are listed in Table 2.1. It is obvious that there are many departments involved, as well as different data formats used. Table 2.2. lists the basic data formats on which the intraoperative record is based. With this list of basic formats a whole anesthetic record can be constructed.

Table 2.1: Anesthetic record.

Data	Format	Source
Demographic data	Text	Patient Information System
Case data	Text	O.R. scheduling
Diagnosis/Procedure	Text	Surgery Department
Preop Information (brief)	Text / Numerical	Anesthesia Department
Pre-medication	Drug Entry	Anesthesiologist
Induction drugs	Drug Entry	Anesthesiologist
Maintenance drugs	Drug Entry	Anesthesiologist
Monitors	Check List	Anesthesiologist
Physiologic data	Numerical	Monitors / Anesthesiologist
Events	Time-Text	Anesthesiologist
Notes on Anesth. Management	Time-Text	Anesthesiologist
Recovery Room Report	Numerical	Recovery Room Personnel

Table 2.2: Data formats of an intraoperative record.

Data format	Description	
Text	Regular text; phrases, words, abbreviation.	
Numerical	Measured number, integer or real	
Drug Entry	`[Drugname][time]<[Amount][Unit][route]>`	
	Drugname	Name of drug or gas
	Time	Time the drug was administered
	Amount	Amount of drug (optional)
	Unit	Unit of the drug (cc/mg...) (optional)
	Route	Way drug was administered (optional)
Check List	List of items that can be checked off	
Time-Text	`[Time][Text]`	
	Time	time mark in real time (hh:mm:ss) or offset time
	Text	see above

To collect these data most hospitals use a one or two page paper form. Record keepers that record many of these data items automatically are now commercially available. These two methods of recording intraoperative data are explained in the next paragraphs.

2.1.1. Manual recording

Paper forms for recording intraoperative information have been around since at least 1894, when Cushing and Codman started to keep an anesthetic chart for their patients, which was adopted by others in subsequent years (Beecher 1940). Devices to record intraoperative pulse traces of heart rate exist since 1860 (Schneider and Redford 1979). Codman used a piece of paper to record heart rate and respiratory rate every 5 minutes on what was called an "ether chart". That methodology is basically still in use today. The manually created anesthetic records of today contain much more information than

its predecessors, but they are still written with pen and paper. Many problems associated with this method of record-keeping can be identified. The main problems are legibility, incorrectness, and incompleteness of the anesthetic record (Meijler 1987, p. 133; Feldman and Good 1993). Incorrectness and incompleteness are caused by the inadequate time the clinician has for record-keeping after he takes care of the patient's needs. Because of lack of time, and a place to write on (table, desk, etc.), often no record entries are made during a critical time period like induction or emergence of anesthesia, or when a problem arises, and the record has to be recreated from memory at a later time. It is during these critical periods that record keeping is needed the most (Whitcher 1987). Incorrectness can also be caused by a conscious or unconscious bias towards recording a less controversial value in the record (Feldman and Good 1993). Incorrectness of systolic blood pressure measurements has been documented in a study by Cook et al. (Cook et al. 1989). They showed that in 33 instances of 50 patients the automatically recorded systolic blood pressure was higher than 110 mm Hg, while in none of the manual records a pressure higher than 110 mm Hg was recorded. In a study of 48 manually recorded anesthesia records at Shands hospital, we observed 27 incidents[1] with a systolic blood pressure of 160 mm Hg or higher. For 14 incidents a systolic blood pressure of more than 5% lower was put on the record than recorded by an independent observer. The observer used the same monitor data as the clinician. The largest recorded discrepancy was a measured value of 213 mm Hg, while the record showed 155 mm Hg; this high pressure, which occurred during intubation, was corrected by deepening the anesthesia. Of the 13 incidents, where the data were correctly recorded, the highest systolic blood pressure on the record was 229 mm Hg, showing that high values do get recorded properly in other cases.

A typical anesthetic record of the Shands hospital at the University of Florida illustrates illegibility (Figure 2.1). Illegibility is caused by poor handwriting and the limited space for comments and events. There is no space for the recovery room report, and therefore it is written wherever space is available, usually in the graphical part of the record (as indicated).

[1]incident is defined here as a period when the physiologic value exceeds a preset upper limit. That period ends when the value returns to a value below that limit, or when the limit is reset.

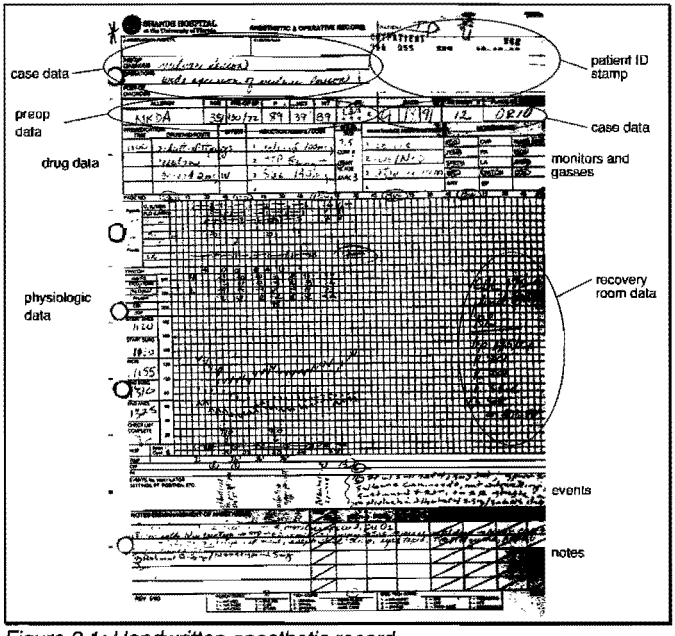


Figure 2.1: Handwritten anesthetic record.

To address the problems of incompleteness, incorrectness, and illegibility, automatic record-keepers have been designed.

2.1.2. Automatic recording

The ideal automatic anesthesia record-keeper would provide the anesthesiologist with automatic recording of all physiologic and other parameters from any monitoring device, and present this information in a clear, obvious way, in printed form and/or on a display. In addition it should be extremely easy to add annotated notes and drug entries to the record (Ream 1989), and the data should be available forever in an easy to retrieve way (Block 1989; King and Smith 1991; Frazier 1987).
This ideal automatic anesthesia record keeper is hard to build. Although the automatic data capture from the monitors can be implemented, despite hardware and software differences between monitors (see par. 2.2.), the user interface between the anesthesiologist and the automatic record-keeper is the most difficult part. The ease with which any anesthesiologist handles a pen and paper is hard to duplicate in an automatic (computerized) device. Nevertheless several automatic record-keepers are

currently commercially available, and some inventive methods are used to approach the ideal record-keeper. Table 2.3. lists the record-keepers that are currently commercially available on the U.S. market with a short description of their user interface.

Table 2.3: Currently available automatic record-keepers.

name	manufacturer	user interface
Arkive	Diatek, San Diego CA	touch screen input, configurable event and drug entries.
Co-Writer	NA Dräger, Telford PA	handwriting input (with a pen) on a plotter.
Compu-record	PPG, Lexena	touch screen and keyboard input, capable of networking.
LifeLog	Modular Instruments, Malvern PA	touch screen, keyboard, and mouse input. Artifacts can be tagged and modified.
OR Data Manager	NA Dräger, Telford PA	keyboard input with function keys.

2.2. Measuring real time data

One component that all automatic anesthesia record-keepers have in common is the ability to read data from intraoperative monitoring devices. There is a wide variety of data formats and hardware implementations, which makes it difficult to make a device that can read from all of the monitors. Some standards have been proposed, but none have been implemented on a large scale in intraoperative monitors yet (Phillips, Gordon, and Cousins 1982; Clemmer and Gardner 1992). One of these standards is the definition of the Medical Information Bus (MIB) by the IEEE Engineering in Medicine and Biology Society (Figler and Stead 1990; Nolan-Avila, Paganelli, and Norden-Paul 1988; Gardner et al. 1992). The MIB standard was defined with the following objectives (adapted from Figler and Stead 1990):

- Interface medical devices with host computers in a compatible, vendor-independent fashion.
- Provide a network that is appropriate for the acute patient care setting.
- Ensure high reliability for accuracy of transmission and delivery of data and for availability and fault tolerance of the network.
- Accommodate frequent reconfiguration and changes in equipment location.
- Provide a simple, non-technical user interface.
- Provide support for a wide range of network topologies.
- Remain cost effective.

Intraoperative monitoring has long been a "stand alone" business, meaning that one manufacturer provides one monitor that measures one or more physiologic parameters.

Monitors were never designed to be an integral part of a data communications network. Redesign of monitors currently on the market, and implementation of complicated networking hardware into new monitors is costly. Therefore manufacturers of monitoring equipment watch the development of new standards closely, but only implement them when they become a golden standard.
Currently intraoperative monitors use serial RS-232 communications, analog signals, or proprietary data buses to communicate with the outside world.

To be able to record intraoperative patient data, we developed a method that can be used to collect data from many different monitors, that is cost-efficient, and easy to use. Based on this method an easy to configure intraoperative data collection tool was built to record serial RS-232 data in the operating room. Paragraph 2.2.2. describes this tool. An analysis of the different types of data is presented in the next paragraph.

2.2.1. Definition of data types

In daily life, and also in computers, we make a distinction between different types of data. The sentence "This patient has a blood pressure of 120 over 80 mmHg" can be classified as a string, and the numbers 120 and 80 can be classified as integers. In computers data types are very important, because they define which operations can be applied to a data set. The standard C computer language defines some basic data types, listed in Table 2.4.

Table 2.4: C language basic data types (Kernighan and Ritchie 1988, p. 36).

Name	Description
char	one character
int	integer number
short	short integer number
long	long integer number
float	real number
double	double precision real number

Other data types can be derived from the basic types e.g. a string can be defined as an array of char (`char str[80]`).
Intraoperative data formats can also be defined in the basic data types listed in Table 2.4. New data types can be defined as combinations of the basic data types. Table 2.5 lists the basic data formats of Table 2.2. in their computerized form.

Table 2.5: Computerized basic data formats of intraoperative data.

Data format	Computer basic data type
Text	array of char
Numerical	int, long, float or double
Drug Entry	*Text+long+Numerical+Text*
CheckList	*Text*+int
Time-Text	long+*Text*

Once data are organized in this way it becomes easier to define structures to store these data and to define operations to manipulate them. As an addition to the data format a meaning can be given to the data. For example data can be of the numerical data format and the meaning defines that it is a systolic blood pressure. An implementation of this concept was made and is outlined in the next paragraph.

2.2.2. A real time data collection tool

With the increase in complexity of equipment, and the increase in computer use, the need to exchange data has also increased. Computers communicate with other computers to exchange data which are stored or collected in one place and need to be viewed in another. The user of this information should not have to go to the location where the data originated, nor have to arrange for the data to be moved manually via courier or mail. An example of this increased communication is the world-wide internet computer network that enables users to send electronic mail, move data files all over the world, and use remotely located information services. Other examples include the availability of laboratory data throughout a hospital, or the access of a library system throughout a university. For communications to work, standards that both sides of the communication agreed upon have to be defined. If the analogy of the telephone is used, this definition ranges from the level of which voltages are sent over the wires, to the high level as which language is spoken by the user through the receiver.

In the definition of network communication standards often the Open Systems Interconnection (OSI) reference model of the International Organization for Standardization (ISO) is used (ISO 1977). This model specifies a frame work of seven layers for connecting computer devices. Each layer is responsible for a specific task, and for exchanging data between adjacent levels. This enables developers of computerized communication devices to implement their systems in a hardware independent way, and define their systems layer by layer. The layers are listed in Figure 2.2.

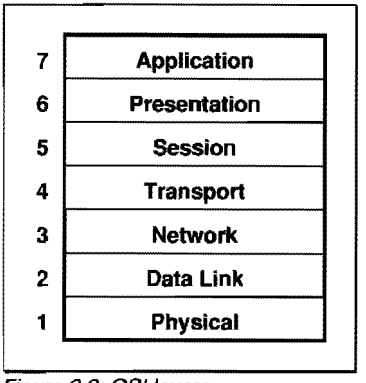


Figure 2.2: OSI layers.

Physical Layer

This layer is responsible for the physical link. It defines the transmission of bits over the link and provides functionality to establish, maintain and deactivate the link. This layer deals with voltages, currents, timing etc.

Data Link Layer

This layer provides reliable transfer of data over the physical link. Synchronization, error control and flow control (e.g. XON/XOFF) of the physical link is handled in this layer.

Network

This layer is responsible for methods of connecting, maintaining a connection, and disconnecting in a hardware independent way.

Transport Layer

This layer provides for the reliable transport of data between two end points. It provides for error detection, and error correction of the data. This may involve the retransmission of data in case of an error.

Session Layer

This layer manages a connection session, by providing for different types of connections. The lower layers make sure that the connections are error-free regardless of connection type.

Presentation Layer

Presents the application program with a standardized way of data representation. Data management like compression, conversion, etc. can be handled here.

Application Layer.

Implementation of standard protocols for data exchange. The user program merely has to call functions in the application layer to have the data transferred.

The OSI model is attractive for applications that transfer data between devices and where dissimilarities exists in the way communication is handled. Similarities between the devices only need to be implemented once in this layered approach. Intraoperative monitoring is such an environment. Similarities and dissimilarities between the data communication parts of intraoperative monitors are shown in Table 2.6.

Table 2.6: Data collection similarities and dissimilarities of intraoperative monitors.

Similarities	**Dissimilarities**
Common low level hardware (RS-232, analog signals)	Non-standard data streams from the monitors (every vendor defines its own)
Common data format (time related physiologic data)	Non-standard control and setup data communication between the monitors
Common data items measured (e.g. blood pressures, heart rates, etc.)	Different hardware communication settings per monitor

The dissimilarities between the monitors can be accounted for in one or more layers in the OSI model. When the differences can be easily modified or configured, all these layers can be implemented in one computer program, enabling uniform data capture from intra-operative monitors. We have made an implementation of a system that can read from intraoperative monitors using the serial RS-232 hardware protocol. The differences between the monitors are stored in a monitor definition file. This file contains the configuration settings that are used for several layers. The next sections describe how the different layers of this system are implemented.

Physical Layer

Many intraoperative monitors are provided with standard RS-232 (EIA Standard RS-232-C 1981) serial ports. This hardware communication protocol is well defined, and if

both devices (the monitor and the PC) adhere to this standard, the physical layer is taken care of.

Data Link Layer

The data link layer needs to provide the physical layer with the appropriate parameters, and provide some simple error detection and connection maintenance. Several computer operating systems provide this functionality, and only the passing and determination of the communication parameters is needed. Some examples of serial communication functions built into Microsoft Windows are shown:

Table 2.7: MS Windows communication functions.

function	Description
CloseComm	Closes a communications device
FlushComm	Flushes a transmission or receiving queue
GetCommError	Retrieves the communications-device status
OpenComm	Opens a communications device
ReadComm	Reads from a communications device
WriteComm	Writes to a communications device

Using these functions a serial port can be opened (`OpenComm`), data can be read from and written to the serial port (`ReadComm` and `WriteComm`), and a serial port can be closed (`CloseComm`).

The parameters that are needed to setup the serial RS-232 port are listed in Table 2.8.

Table 2.8: Parameters of RS-232 protocol.

parameter	description
baud rate	transmitting and receiving speed of the connection
data bits	number of data bits per byte of data
parity	simple error checking scheme
stop bit	number of trailing bits after a byte

These parameters are fixed per monitor, and are stored in the monitor definition file. After these parameters are set, basic communication tasks can be performed on the interface with standard operating system functions that read and write to/from communication ports. Raw data communication has been established.

Network Layer

The network layer maintains the connection and communication with the monitors, by passing additional parameters specific for the monitor to the data link layer, and by assuring that data is being moved until the connection is closed down.

Table 2.9: Additional communication parameters.

parameter	description
name	Name of the monitor
EOTchar	End of Transmission Character
Request	Request strings to prompt the monitor for data

The request strings are used to request the monitors for data if the monitor doesn't provide data in *printer mode*[2]. These request strings are sent out whenever the application program wants to receive data. The EOTchar is the End-Of-Transmission character that indicates the end of a message or data packet from the monitor to the computer. When this character is received the data is passed on to higher levels for processing. The name of the monitor can be used to simplify setup.

The network layer and the data link layer are very similar in passing parameters to the communication interface. We merged these two layers so that they can both read from the same monitor definition file. This file stores the parameters needed to connect and stay connected to a monitor. These data are stored as shown in Figure 2.3.

```
[Monitor1]
Name=Critikon Dinamap NIBP
baud=600,N,8,1
EOTchar=0D

Request1=B*C
```

Figure 2.3: Monitor communication data in monitor data definition file.

Transport Layer

The transport layer provides reliable communication between the monitor and the computer. The data sent out by the monitor usually does not provide for any checking of reliability of the data communication. The reliability checking has to be implemented with the available data. Usually data strings have some identification information in them that labels the signal values (e.g. the string "`BP=090`"). We will call this

[2]In *printer mode* a monitor automatically sends out data from the moment it is turned on.

identification information a *search key*. Search keys can be specified as ASCII (plain text) characters, or as binary characters (some monitors use a binary identification). If the data from a monitor do not contain any search keys given for the monitor, the data are discarded. The search keys are also used to extract data from the monitor data stream and are explained in detail in the discussion of the presentation layer.

Session Layer

A session in this implementation is defined as one application program connected to one or more monitors. The program needs to know which monitors are connected to which port. On a standard PC four serial ports (COM1-COM4) can be used simultaneously. The setup file for the application program includes the port name, and the name of the monitor connected to it.

```
[Samples]
Data=500
Request=500

[Ports]
Port1=com1
Device1=Monitor1

Port2=com3
Device2=Monitor3
```

Figure 2.4: Setup file of the application program.

The example setup file shown in Listing 2.4 shows the setup of two monitors. Additional data as the sample rate (`Data`) and the data request rate (`Request`) are also shown. The sample rate can be adjusted to meet monitoring needs (Gravenstein et al. 1989).

Presentation layer

The session layer offers the presentation layer a reliable stream of data. It is the task of the presentation layer to extract the data and represent it in a uniform way, independent of how the data arrived, so that an application program can use it.

Data sent out by the monitors can be integer or real numbers, strings, etc. For every data type several representations can occur in the data stream from the monitors. Numerical integer data (e.g. the numbers `3` and `122`) can be send as `003122` (ASCII fixed length integer) or `3,122` (ASCII variable length integer). Figure 2.5. shows the different ways data are represented in a data stream.

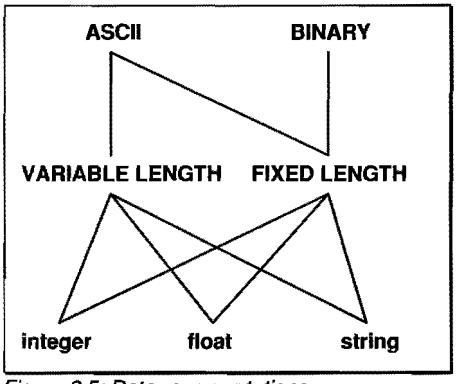


Figure 2.5: Data representations.

An implementation for extracting ASCII, fixed length, integer data from the data stream was made. Extractions for other data representations can be designed in a similar fashion. ASCII, fixed length integer data is most frequently available on the serial ports of physiologic monitors.

ASCII fixed length integer data

Because the data, sent on the serial connection of a monitor, are usually intended for a printed log, the data appear in fixed positions in the data packet. A typical data packet might be: `HR=088;BP=120↵`[3] In this example the heart rate data start at the 4th position in the packet and is three characters long. In addition a program can look in the packet to search for HR to make sure it received a correct packet and a name can be added to identify the signal. This leads to the following data needed per signal on each monitor:

Table 2.10: Signal identification parameters.

item	description
Key	Search Key for data reliability
Start	Start position of the data item
Length	Number of bytes of the data item
Name	Name of the signal

In the monitor definition file these parameters are stored as follows:

```
Key1=72 82 61          (HR=)
Signal1=4 3            (position length)
n1=Heart rate
```

Figure 2.6: Signal data in monitor definition file.

[3]The ↵ character denotes the carriage return (CR) character.

The search key is stored in hexadecimal format, because some monitors (e.g. Datascope monitors) require binary search keys. The position of the beginning of the data item is relative to the beginning of the search key. If no search key is indicated (specify `00 00 00`) no validation is done, and the position is relative to the beginning of the packet.

After extracting the data from the data stream they have to be stored in a uniform manner. For each data item an identification (or meaning), the time of measurement, the format type of the data, and the value of the data item have to be stored. One approach is outlined in Tables 2.11. and 2.12.

Table 2.11: Data needed per signal.

section	size in bytes
id	1 byte
time	4 bytes
type	1 byte
data	depends on type

Table 2.12: Defined signal identifiers.

id	description	type
0	Systolic Blood Pressure	INT
1	Diastolic Blood Pressure	INT
2	Heart Rate	INT
3	Intra-op event	STRING
4	Drug administration	STRING
5	$P_{et}CO_2$	FLOAT
6	S_nO_2	INT
7	Insp. O_2	FLOAT
8	Mean Arterial Pressure	INT

The four bytes that store the time contain the number of seconds since 1/1/70 00:00:00, a standard way of handling time in a computer (on Unix and DOS systems). Currently there are three format types of the data:

Table 2.13: Data format types.

id	id define	description
0	INT	signed integer: 2 bytes
1	FLOAT	floating point: 4 bytes
2	STRING	string: first byte indicates the length, string data follows

Application layer

The application layer consists of user defined programs that use the underlying layers, and that manage and further process the data. The most simple program would be one that initiates monitor communications, reads the data, stores or displays the data, and closes monitor communications. We defined and implemented these basic functions as follows:

```
IO_Init()
```

Opens all the communication ports and associates monitor data with them.

```
IO_CloseDown()
```

Terminates the communications that were initiated with the IO_Init call.

```
IO_GetData(struct DATASTRUCT *alldata)
```

Reads all the data from the ports for processing in the application. The data are returned in the structure DATASTRUCT (Table 2.14).

```
void IO_RequestData()
```

Sends request strings to the connected monitors to query for new data.

The data structure used in the IO_GetData is defined as follows:

Table 2.14: Data structure to store signal data.

DATASTRUCT		
name	**type (size)**	**description**
id	byte (1)	identification of data
time	long (4)	time the measurement was taken
type	byte (1)	type of the data
data	pointer	pointer to the data

Listing 2.1 shows the framework of a data collection program based on the functions of the application layer.

```
#include "osimon.h";

void main()
{
        struct DATASTRUCT alldata[MAXITEMS];

        IO_Init( );

        while (!the_end)
        {
                if (time_to_query)
                        IO_Request( );
                if (time_to_read)
                {
                        IO_GetData(alldata);
                        ProccessData(alldata);
                }
        }

        IO_CloseDown( );
}
```

Listing 2.1: Minimal data collection program.

An implementation of the techniques presented in this paragraph was used to record intraoperative data (van Oostrom et al. 1992). The implementation was made in Microsoft Windows as a multiple document interface (MDI). The MDI monitoring program consists of one main window with a menubar, and multiple smaller windows on top of that, representing the connected monitors. The monitoring program displays the data coming from the intraoperative monitors in tabular form and the data are also written to file. The display of the program is shown in figure 2.7.

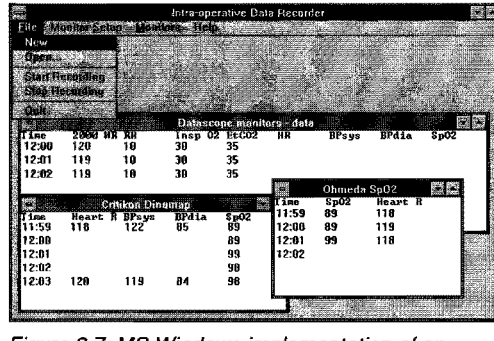


Figure 2.7: MS Windows implementation of an intraoperative data collection tool.

This system is able to read the data listed in Table 2.15, which accounts for many of the intraoperative monitoring devices at Shands Hospital. Other monitors can be added by entering their communication data in a monitor definition file.

Table 2.15: Monitors compatible with the intraoperative data recorder with available data.

Monitor name	Measured variable(s)
Datascope	heart rate
	respiratory rate
	inspired O_2
	end tidal pCO_2
	non invasive systolic blood pressure
	non invasive diastolic blood pressure
	S_nO_2
Nellcor S_nO_2	heart rate
	S_nO_2
Critikon Dinamap	heart rate
	non invasive systolic blood pressure
	non invasive diastolic blood pressure
	S_nO_2
Ohmeda S_nO_2	heart rate
	S_nO_2
HP Snoopy	end tidal pCO_2
Ohmeda 5250	end tidal pCO_2

2.3. Current practice of alarm settings

This paragraph is a reprint of the paper "Acceptable Ranges for Vital Signs during General Anesthesia" to be published in the Journal of Clinical Monitoring 1993.

Acceptable Ranges for Vital Signs during General Anesthesia

J.H. van Oostrom, C. Gravenstein, and J.S. Gravenstein, M.D.

van Oostrom JH, Gravenstein C, Gravenstein JS. Acceptable Ranges for Vital Signs during General Anesthesia.
J Clin Monit 1993;in press

ABSTRACT

Objective: Define the ranges for normal vital signs during general anesthesia

Methods: We studied 50 patients undergoing general anesthesia. We asked residents to state desirable ranges for each patient's systolic and diastolic blood pressure (BP), heart rate (HR), SpO_2 and $P_{ET}CO_2$ during induction, intubation, maintenance, and emergence from anesthesia. We called these ranges the clinical operating range (COR) and observed the frequency, duration and magnitude of transgressions of these CORs. We also recorded whether the transgressions were treated, tolerated, or whether the COR values were changed.

Results: Upper COR values in the maintenance phase for systolic BP were 38% ± 20% above the preoperative values and 30% ± 20% above the values recorded just before induction ot anesthesia. Lower COR values in the maintenance phase for systolic BP were 27% ± 9% below preoperative, and 31% ± 11% below pre-induction values. For HR, upper and lower COR values in the maintenance phase were 53% ± 44% above and 38% ± 17% below preinduction values, respectively. Transgressions of COR values for BP and HR were common, treatment frequent, and redefinition of COR values rare.

Conclusion: Clinicians recognize ranges for vital signs during uneventful anesthesia. These CORs may differ from one stage of anesthesia to the next. Transgressions of these ranges are common. Not all transgressions are treated.

KEY WORDS. Equipment: alarms. Monitoring.

Neither in daily life nor during anesthesia are vital signs immutably stable. Instead they rise and fall: under circadian influences, in response to activities, and during anesthesia as a consequence of drugs, mechanical ventilation, body position, and surgical perturbations. Anesthesiologists expect to see such changes; for instance, in a healthy patient being anesthetized with nitrous oxide and isoflurane in oxygen, we expect a rise in arterial blood pressure (BP) and heart rate (HR) during intubation of the trachea and a reduction of these values during maintenance of anesthesia. Anesthesiologists anticipate not only the direction of changes, but also their magnitude. Thus, when we give anesthesia, we work with (often unstated) ranges of acceptable values for a given patient and for given phases of a specific anesthetic. In this study we call these ranges the clinical operating range (COR) for a given, monitored physiologic variable. In contra-distinction we define alarm limits as thresholds where an imminent danger is recognized and a response is required.

When a patient's vital signs fall outside the COR during anesthesia, the anesthesiologist can intervene by deepening or lightening anesthesia, giving fluids, or administering drugs. Or, the clinician can elect to tolerate, for a limited time, values that lie outside the established COR. For example, if the incision is expected to reverse an unacceptable hypotension, a temporary transgression of the COR for BP might be deemed permissible.

Clinical operating range values are not the same as alarm limits. We expect frequent crossings of COR values during uneventful anesthesia, but no transgression of alarms limits. For a population of 50 patients, we set out to collect COR values, their ranges, and the frequency of transgression of these values during routine anesthetic care.

METHODS

Over a span of four weeks, we studied 50 patients (23 men and 27 women) undergoing short, elective procedures under general anesthesia. The patients' ages ranged from

13 to 85 years. The study was approved by the institutional Review Board for Human Experimentation; informed consent was not required, as the study was limited to observations of routine events. Seventeen residents in their 4th to 24th month of clinical anesthesia training anesthetized 50 patients under staff supervision. We informed the resident of the study and explained that transgressions of these COR values were expected. In response to a transgression, the resident could (1) take action to bring the signal back into the desired range; (2) change the COR value; or (3) ignore the transgression. After the resident had obtained baseline data for SpO_2, systolic and diastolic BP, and HR, he or she announced to the investigator a set of physiological limits (COR) for systolic and diastolic BP, HR, SpO_2, and $P_{ET}CO_2$ within which he or she wished to keep the patient for each phase--induction, intubation, maintenance, and emergence--of anesthesia.

One patient had halothane anesthesia by mask; all others were intubated. One of the intubated patients received propofol and narcotic, while the rest received isoflurane as the major anesthetic. The induction phase began with the first medication in the operating room. The intubation phase began when the face mask was removed and the laryngoscope inserted and ended when the patient had been positioned for surgery. The maintenance phase ended when the inhalant anesthetic was turned off or the last suture was set, and the emergence phase was completed when the patient was extubated or left the operating room.

One observer (CG) recorded all data. Preoperative values for BP were copied from the pre-anesthesia evaluation form; pre-induction baseline data were read from the monitors in the operating room. The observer watched the monitors and recorded all transgressions of COR limits. End-tidal PCO_2 was measured with a mass spectrometer (Perkin Elmer, Pomona, CA) or infrared analyzer (AccuCap, Datascope, Paramus NJ, or Ohmeda 5250, Ohmeda, Boulder, CO), and SpO_2 was measured with a Biox 3700 (Ohmeda, Boulder, CO) or N100 (Nellcor, Haywood, CA) pulse oximeter. Values for BP were obtained with an oscillometric BP sphygmomanometer (Ohmeda 2120, Ohmeda, Englewood, CA; Dinamap, Critikon, Tampa, FL; or Datascope, Paramus, NJ), and those for HR from either the ECG (Datascope, Paramus, NJ) or the pulse oximeter. We recorded all transgressions, their values, and their duration on an IBM-compatible personal computer that used Windows 3.0. For each transgression, the time, the name of the signal, the COR limit (upper or lower) transgressed, the current value of the signal, and the reaction to the transgression were noted. If the resident did not react to an exceeded limit after about a minute's delay, the observer asked about the given value and whether the resident wished to ignore it, treat it, or change the limit.

We collected all preoperative evaluation forms, all anesthesia records and, where available, the printout of the automated BP monitors.

Standard deviations and averages of the COR limits were calculated. Statistical tests included the Whitney-Mann test for duration and magnitude of the treated and untreated transgressions, and the Wilcoxon signed rank test for pairs for the pre-induction and pre-operative systolic BP.

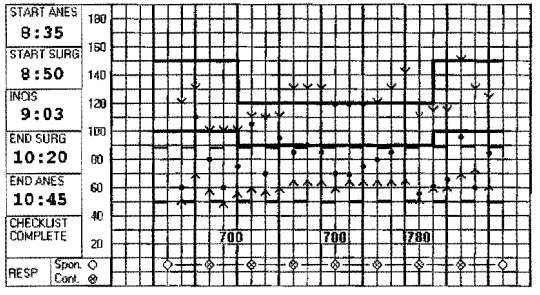


Fig 1. Part of the anesthesia record of one of the patients. The clinical operating ranges (CORs) for systolic blood pressure (solid line) and heart rate (dotted line) are indicated. In this case, the COR values for induction and intubation were identical. One episode of tachycardia was treated with esmolol, another one remained untreated. One episode of hypertension was treated by increasing the isoflurane concentration.

RESULTS

We counted 305 transgressions in the 50 cases that we recorded. An excerpt from a typical anesthesia record shows the wide variation in the signals that cause these transgressions (Figure 1). The COR limits for heart rate and systolic BP were the ones most often involved (Figures 2 and 3).

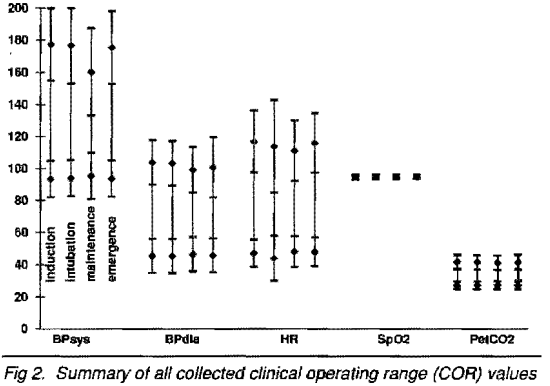


Fig 2. Summary of all collected clinical operating range (COR) values and their standard deviations. The COR values for SpO_2 and for end-tidal PCO_2 changed only minimally over the 4 phases of anesthesia. Abbreviations: BPsys = systolic blood pressure (mmHg); BPdia = diastolic blood pressure (mmHg); HR = heart rate (beats/min); SpO_2 = oxygen saturation as measured by pulse oximetry (%); $P_{ET}CO_2$ = end-tidal PCO_2 (mmHg). For all the signals, the order of the phases is as indicated for BPsys.

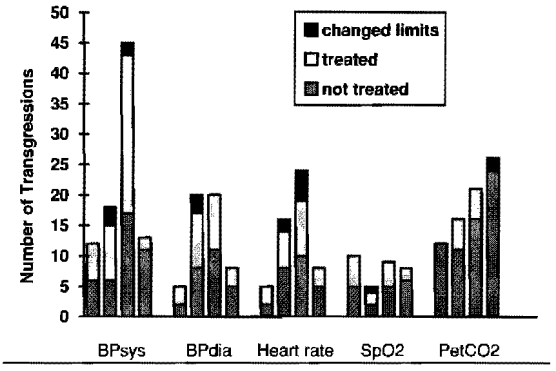


Fig 3. The graph shows all transgressions of clinical operating ranges and indicates the clinician's response to each transgression. Each bar represents, from left to right, the induction, intubation, maintenance, and emergence phases for each variable.

Eleven patients accounted for 50% of the transgressions. With one patient, COR limits were transgressed 23 times, while, with 5 other patients, COR limits were never reached.

The 17 residents who participated in this study all set similar COR limits for SpO_2 and $P_{ET}CO_2$, and they maintained these limits throughout the 4 phases of the operation. There was much greater variation, however, in the limits the residents set for HR and systolic and diastolic BP. Generally, though, the range of acceptable values was more narrow for the maintenance phase than for the other phases (Figure 2).

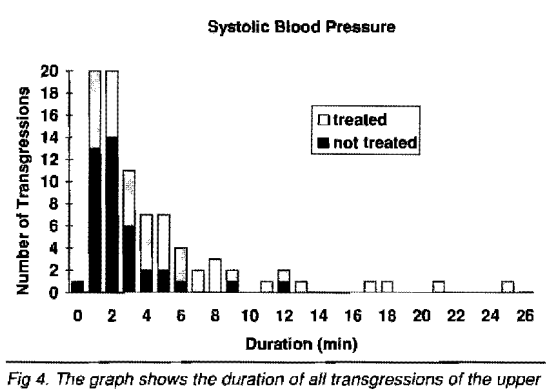


Fig 4. The graph shows the duration of all transgressions of the upper clinical operating range limit established for systolic blood pressure, with an indication of whether or not the transgression was treated.

The duration of transgressions of COR limits for BP and HR was skewed (Figures 4 and 5). Generally, the transgressions that lasted the longest were smallest in magnitude, and the briefest transgressions were largest in magnitude. Of all transgressions, 76% lasted less than 10 minutes and were within 20% of the designated COR limit, 12% lasted less than 10 minutes and were larger than 20% of the COR limit, and 9% lasted more than 10 minutes and were larger than 20% of the COR limit (Figure 6). There was

no statistically significant difference between treated and untreated transgressions with respect to the duration or magnitude of the transgressions.

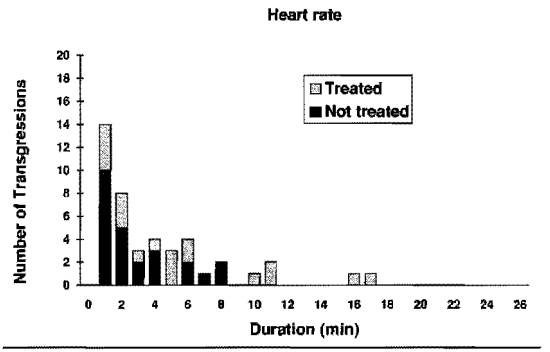


Fig 5. The graph shows the duration of all transgressions of the upper clinical operating range limit established for heart rate, with an indication of whether or not the transgression was treated.

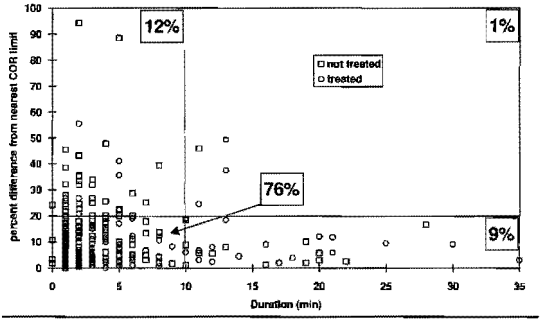


Fig 6. The duration and magnitude of all transgressions of clinical operating range (COR) limits for systolic blood pressure. The magnitude of each transgression is plotted as the percentage of difference from the nearest COR limit.

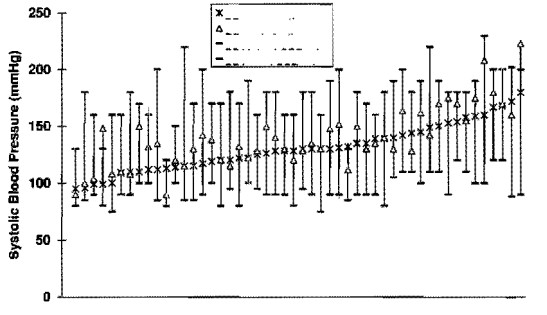


Fig 7. For each patient, values for systolic blood pressure obtained preoperatively (BP preoperative) and just before induction of anesthesia (BP pre-induction) are shown in relation to the upper and lower clinical operating range limits established by the resident. The patients were ranked according to their preoperative blood pressure.

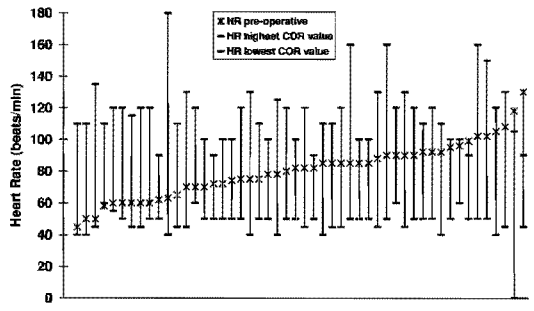


Fig 8. For each patient, the value for heart rate obtained just before induction of anesthesia (HR pre-induction) is shown in relation to the upper and lower limit of the clinical operating range established by the resident. The patients were ranked according to their pre-induction heart rate. In one patient, no lower limit was set.

Preoperative values for systolic and diastolic BP (but not HR) for each patient were recorded 1 to 2 days before anesthesia. The Wilcoxon signed rank test for pairs showed that preoperative systolic BP was significantly lower ($p = 0.0026$) than pre-induction systolic BP, a finding that confirms previous studies [1]. On average, the upper maintenance phase COR values for systolic BP chosen by the residents were 38% $\pm$ 20% greater than the preoperative values and 30% $\pm$ 20% greater than the values recorded just before induction of anesthesia. The lower maintenance phase COR values for systolic BP were 27% $\pm$ 9% less than values recorded preoperatively and 31% $\pm$ 11% less than values recorded before induction of anesthesia (Figure 7). Compared to values obtained before induction, the upper maintenance phase COR values for HR were an average of 53% $\pm$ 44% greater and lower maintenance phase COR values were an average of 38% $\pm$ 17% less (Figure 8). The COR values for BP and HR showed no statistically significant correlation with pre-operative or pre-anesthesia control values.

A statistical significant difference (two-tailed Wilcoxon signed-rank test for pairs) was found between the COR limits of the maintenance phase and the COR limits of the other phases for BP and HR.

DISCUSSION

That vital signs vary considerably during uncomplicated anesthesia is well recognized and described [1-3]. Clinicians are aware of this variability, and our residents and attending anesthesiologists were willing to predict ranges for vital signs to be expected for individual patients, for the phases of specific, uncomplicated anesthetic regimens. We have called these ranges "clinical operating ranges." To use a common analogy, COR values may be likened to a green zone where we hope the patient's vital signs will remain; the range lying just outside this green zone might be pictured as a yellow zone where the clinician begins to consider interventions. Outside the yellow zone lies the red zone of imminent danger. We suggest that alarms should be sounded when the patient's vital sign cross from the yellow into the red zone. The utility of and the problems surrounding alarm limits and alarm technology have been discussed

repeatedly [4]. A major problem is false alarms, which lead many clinicians to disable alarms in frustration. False alarms occur either because of artifacts or because the limits were set inappropriately.

Transgression of COR values might occur when some unexpected events, such as a hemorrhage, causes the arterial pressure to fall. Recognition of the problem would trigger a response and correction long before the patient's vital signs would drift into the red danger zone. The boundaries of CORs may also be crossed when artifacts are recognized, as was the case in 16 of 172 transgression that were not treated. At other times, transgressions of COR values are not treated because the clinician expects the variable to return to the COR spontaneously, as with the reversal of hypotension by the surgical incision. We have recorded 7 such instances out of 172 untreated transgressions. Finally, the COR values themselves may be changed when the clinician decides that the chosen limits are too narrow. This occurrence is of considerable interest as it indicates that the clinician is guided by several signals rather than focusing on the value of only one. For instance, in one case the patient's arterial BP fell, but all other indications suggested that the patient was adequately perfused (heart rate, skin color, skin moisture, pupil size, urine output). Readjustment of COR values occurred a total of 16 times in 13 patients (7 for HR, 5 for systolic BP, 3 for diastolic BP, and 1 for S_pO_2). This sample is too small, however, to allow us to draw conclusions about the variables that influenced the decision to readjust the COR values of these signals.

In this study we asked clinicians to define CORs for their patients. Consequently, the CORs presented here are subjective and are based on a small sample. Another, better way to obtain COR values would be to collect vital signs from many uncomplicated anesthetics with a proven uncomplicated outcome. Such measurements would provide the data for the generation of objective COR values through statistical analysis. Lacking such statistics, clinicians call on their experience to define the CORs.

REFERENCES

1. Berger JJ, Donchin M, Morgan LS, van der Aa J, Gravenstein JS: Perioperative changes in blood pressure and heart rate. Anesth Analg. 1984;63:647-652
2. Block FE Jr: Normal fluctuation of physiologic cardiovascular parameters during anesthesia and the phenomenon of "smoothing." J Clin Monit 1991;7:141-145
3. Donchin M, van der Aa JJ, Berger JJ, Gravenstein JS: Perioperative cardiovascular patterns, in Fleming WW, et (eds): Neuronal and Extraneuronal Events in Anatomic Pharmacology. New York, Raven Press, 1984, pp 241-249.
4. Kestin IG, Miller BR, Lockhart CH: Auditory alarms during anesthesia monitoring. Anesthesiology 1988;69:106-109

2.4 Summary

In this chapter we have presented the data flow in the operating room to get a better understanding what data is available and how it is used. We presented a system that can record intra-operative data from different monitors. This system was used to collect data from monitors to record the variability of the intra-operative data. It proved to be difficult to relate the real-time data with intra-operative events, because they need to be entered by an anesthesiologist. To observe intraoperative events and to document the use of intraoperative limits, a study was done with 50 operative cases where event/action data were recorded by an observer. The results of this study give a basic idea of how intra-operative limits are used and manipulated, and provided an assessment of the clinical operating ranges for specific patients. The study shows that anesthesiologists can decide on acceptable ranges for vital signs, based on information about the patient, and that these ranges are not changed frequently during the operation. We can conclude from this study that limits differ between the maintenance phase and the other three phases, and are set similarly for intubation, induction and emergence phases. This dissertation will focus on the maintenance phase of anesthesia because it is the phase where patient variables are fairly stable and the setting of alarm limits is not affected by artifacts or events that are occurring (insertion of the endotracheal tube for example).

The next chapter will present different approaches on how available patient information can be combined with this knowledge to automatically set alarm limits.

3 Approaches to Automatic Alarm Limit Setting

3.1. Introduction

Anesthesiologists use limits for physiologic patient variables that are based on their clinical knowledge and experience, information about the patient, and case related data (see *What kind of knowledge is involved* later in this chapter). Information about the patient is collected on preoperative assessment forms or with the help of computer programs like the Preoperative Assessment Program.

To set alarm limits, anesthesiologists decide which values are desired or acceptable for physiologic patient variables, and which ones are not. Anesthesiologists have gained knowledge and experience about how physiologic patient variables fluctuate during their clinical practice, from literature, and from presentation and discussion of other cases.

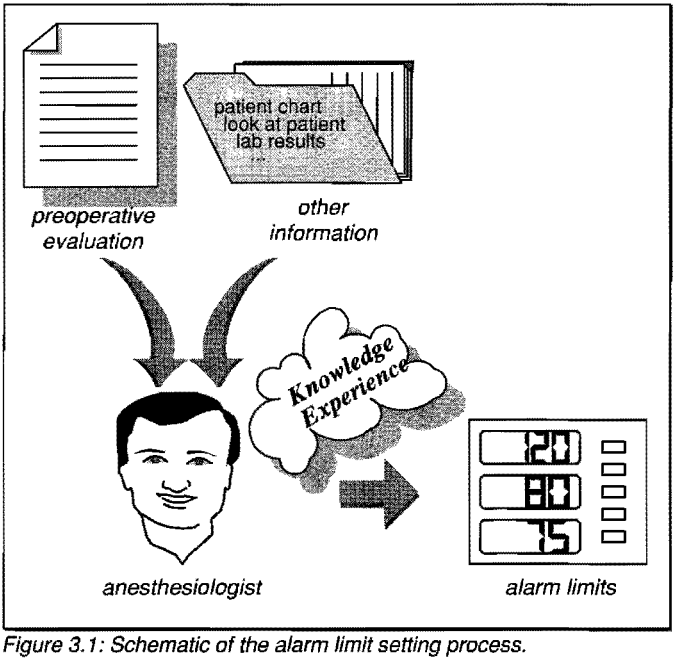


Figure 3.1: Schematic of the alarm limit setting process.

A system that can set alarm limits automatically should have the same information available as the anesthesiologist. It also must have access to the same knowledge the anesthesiologist has. In computer science, the field of Artificial Intelligence (AI) is devoted to store, retrieve, and manage knowledge and data to make computers solve problems in an intelligent manner. A system (man or machine) that shows intelligent behavior is assumed to have a structure called a knowledge base (Winston 1984). This is the structure that is used to store the knowledge. A separate mechanism is implemented to manipulate and retrieve knowledge, this is called the inference engine (Reichgelt 1991).

Before an AI solution is attempted for a problem, there are some criteria that require evaluation. These criteria are (Winston 1984):

1. Is the task clearly defined?
2. Is there an implemented procedure performing the defined task?
3. Is there a set of identifiable regularities or constraints from which the implemented procedure gets its power?

1. Definition of the task

The task is defined as the automatic setting of intraoperative alarm limits for patients undergoing surgery, based on information about the patient and the case.

2. Implemented procedure

The procedure implemented for this task is 1) gather all the information available about the patient and the intended surgery, 2) decide on which ranges of physiologic values are acceptable for this patient based on this information, and based on general knowledge, and 3) decide which alarm limits to set.

3. Regularities and constraints

Several constraints can be identified: 1) an upper limit always has a higher numerical value than a lower limit, 2) an upper systolic blood pressure limit always has a higher numerical value than an upper diastolic blood pressure limit, 3) a lower systolic blood pressure limit always has a higher numerical value than a lower diastolic blood pressure limit, and 4) the highest value for a S_pO_2 limit is 100%. There are also some physical constraints like a blood pressure limit of 0/0 can never occur (Block 1989), and similarly

a lower heart rate limit of 0 is evidently wrong (except during special cases like cardio-pulmonary bypass).
Some regularities of this problem are 1) if the same patient comes in for the same operation, and the patient has the same physical state as the previous visit, and the limits used for that visit were satisfactory (not many false positive alarms, no missed alarms), then those limits can be used for the current operation, and 2) patients with a similar physical state, the same type of problems, and undergoing the same type of operation, will require similar alarm limit settings.

These answers suggest that an AI solution is feasible. The main question that now remains is which implementation to choose and how the knowledge should be gathered, stored, managed and retrieved.
Some basic questions can be formulated to assist in the task of knowledge engineering (Winston 1984):

- What kind of knowledge is involved?
- How should knowledge be represented?
- How much knowledge is required?
- What exactly is the knowledge needed?

This chapter will address the first two questions, and the next chapter will answer the last two.

What kind of knowledge is involved?
The analysis of the type of knowledge that is involved for an AI solution to a problem can be made independently of the representation of the knowledge. To find out what types of knowledge are involved in the setting of intra-operative alarm limits, a questionnaire was designed and distributed during one of the research meetings, held monthly at the department of Anesthesiology at the University of Florida for residents and faculty. Eighteen residents and two faculty members returned the questionnaire. One of the questions was "Where do you get the information that you use to define your alarm limits?". Ten answers included the preoperative evaluation, six included the patient record, and one the vital signs. Six of the questioned did not answer or did not know the answer to this question. The three most important parts of the preoperative evaluation could be indicated on the questionnaire. The importance for each answer was rated as 3 for the most important, 2 for medium importance and 1 for the least

important. The results of the answers of the 20 residents and faculty were added up and are presented in Figure 3.2.

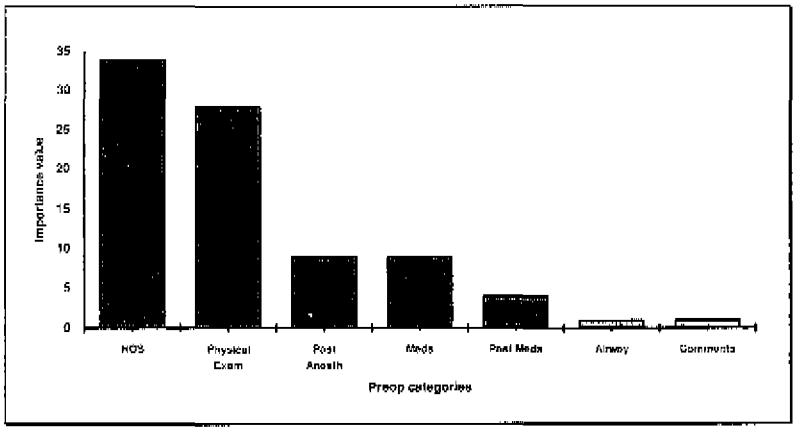


Figure 3.2: Importance of preoperative data items.

The preoperative evaluation form includes the following measurements in the physical exam: weight, height, non-invasive systolic and diastolic blood pressure, heart rate and body temperature. The Review of Systems (ROS) includes the following physiologic systems: central nervous system, cardiovascular system, respiratory system, gastrointestinal system, genito-urinary, endocrine system, hematological system, and skeletal system. In addition the ROS includes notes about allergies, infectious diseases, and miscellaneous notes.

Other knowledge that is used to set alarm limits is knowledge obtained by the anesthesiologist over years of studying and experience. Three methods of knowledge representation and knowledge acquisition are outlined in the following paragraphs.

3.2. Expert system approach

Expert systems typically maintain their knowledge in rule bases. A rule base is defined as a collection of rules. A rule is a description of a fact or other small piece of knowledge (e.g. 0 < patient age < 120). Knowledge in a rule base is built up by first defining basic rules, and then expand those with more complex ones. Expert system rules are sometimes called *if-then* rules because they usually have an *if* part and a *then* part:

```
if ([not]) <condition1> [and/or [not]] <condition2> [and/or [not]] ...
      <conditionn>
then <action1> <action2> ... <actionm>
```

A rule base can contain many rules. To find a solution to a problem, evaluation of the rules that are needed to solve this problem is necessary. An example is shown in Figure 3.3. The rules *no boot hard disk* and *no boot floppy* depend on other, more basic rules. By dividing the rule base this way, the knowledge contained in the rules becomes more manageable.

```
1: no boot hard disk rule:
if machine doesn't boot and machine has hard disk
      and hard disk works
then reinstall DOS on the hard disk

2: no boot floppy rule:
if machine doesn't boot and not machine has hard disk
then use a boot floppy

3: machine doesn't boot rule:
if ask the user
then this rule is true

4: machine has hard disk rule:
if call dos HasHardDisk function or ask the user
then this rule is true

5: hard disk works rule:
if call dos AccessHardDisk function or ask the user
then this rule is true
```

Figure 3.3: Expert system rule base example. Rules are in italics, and actions are underlined.

The process of evaluation of the rules and finding the rules that are needed is the task of the inference engine. The goal of the example of Figure 3.3. is to evaluate why a machine does not boot . The inference engine should first evaluate the *no boot hard disk* rule and then evaluate the rules that it includes. The most efficient inference engine only evaluates the rules that are needed. The most simple inference engine evaluates all rules, which includes the rule that defines the goal of the evaluation. Two types of knowledge can be separated in a rule base: permanent knowledge and temporary knowledge. Permanent knowledge is stored in the rules and is fixed. Temporary knowledge reflects the current state of the system. Table 3.1. shows the different types of knowledge stored in the example rule base.

Table 3.1: Permanent and temporary knowledge in the example rule base.

permanent knowledge	temporary knowledge
the machine doesn't boot from the hard disk if it 1) doesn't boot, 2) has a hard disk and 3) the hard disk works (rule 1)	the user says something is true *(ask the user)*
the machine doesn't boot from the floppy disk if it 1) doesn't boot and 2) doesn't have a hard disk (rule 2)	DOS says there is a hard disk `(call dos HasHardDisk)`
the machine doesn't boot if the user tells me so (rule 3)	DOS can access the hard disk `(call dos AccessHardDisk)`
the machine has a hard disk if DOS or the user tells me it does (rule 4)	
the hard disk works if DOS can access it or the user says it works (rule 5)	

Rules are based upon knowledge of experts, who have knowledge of the specific domain that the expert system will work in. This knowledge is usually obtained by interviewing the experts, but some other methods have been tried successfully such as heuristic methods (Winston 1984) or analysis of simulated data (van Oostrom 1989, p. 20).

Several expert systems have been developed for medicine (Rennels and Miller 1988; Holman and Cookson 1987). Examples include an intelligent alarm system that gives descriptive messages about the status of the anesthesia circle breathing system based on an expert system rule base of about 200 rules (van der Aa 1990; van Oostrom 1989), MYCIN, a medical consultation system, ATTENDING a system that critiques an anesthesiologist's preoperative plan for anesthetic management (Miller 1983, 1984), and Resac a system providing decision support for control of depth of anesthesia (Greenhow SG et al. 1992).

There are some disadvantages associated with using expert systems. Before knowledge can be stored in a rule base the knowledge has to have a structured form. Domain experts have to be able to identify rules and procedures they use. This is not always possible, especially if their knowledge is partially based on their experience. Retrieving this knowledge is no small task, and once the knowledge has been collected it is fixed. Rule bases are usually large and hard to maintain, which is problematic if procedures or opinions change. In our case of the setting of alarm limits, the knowledge

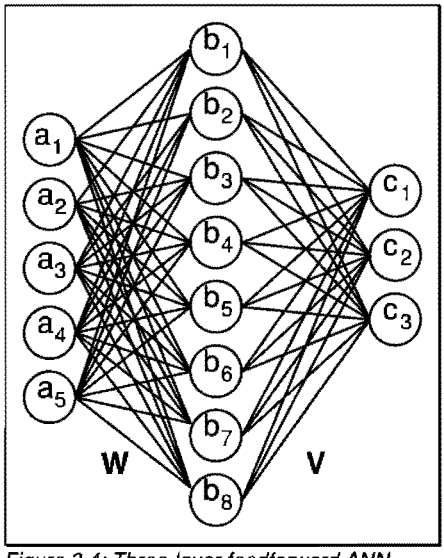
is based on current practice and experience. An expert system implementation for alarm limit setting would only be useful for knowledge that is known to be static (e.g. an upper limit has a higher numerical value than a lower limit). For the other knowledge, the expert system is not the best approach.

3.3. Neural net approach

Artificial Neural Network Systems (ANNs) are mathematical models of theorized mind and brain activity. The main idea behind ANNs is that if the brain can be modeled mathematically and if these mathematics can be implemented in a computer, that this is a good approach for solving problems or capturing knowledge.
The brain is thought to be a network of neurons and connections between neurons. An artificial neural system is built the same way. The neurons in an artificial neuron net are called Processing Elements (PE). A PE takes inputs from the outside world, or from another PE, and can be represented as a vector $\mathbf{A} = (a_1, ..., a_n)$. Each element of **A** represents the activity level of the input. Weights are applied to the inputs and form a vector $\mathbf{W_j} = (w_{1j}, ..., w_{1j})$ for the *jth* PE. By applying these weights it is determined which inputs have more effect than others. The values of the weights are optimized in a learning process of the training of the neural net (see the next section). The output of the PE is calculated with the following formula (Simpson 1990):

$$b_j = f(\mathbf{A} \cdot \mathbf{W_j})$$

The function f() is a threshold function which determines how the PE propagates it's inputs (in real neurons this function determines when a neuron fires). These threshold functions are usually linear, ramp, step or other functions. An example of this is a neuron with a step threshold function, that only fires if the voltage at the input exceeds 1mV. Voltages of <1mV have no effect on this neuron.
Networks of PEs are built to form an artificial neural network system. Different topologies can be built, but one frequently used one is shown in Figure 3.4.



In this example there are 5 input nodes ($a_1...a_5$), 8 hidden layer nodes ($b_1...b_8$), and 3 output nodes ($c_1...c_3$). The lines between the nodes represent the weight factors; Matrix **W** between **a** and **b**, and matrix **V** between **b** and **c**. v_{11} is the weight factor between a_1 and b_1 and w_{83} is the weight factor between b_8 and c_3.

Figure 3.4: Three-layer feedforward ANN.

Learning

Just like a human brain an ANN needs to be trained with knowledge. This is called the learning process. The memory of this artificial brain are the weight matrices **V** and **W**. When the network is stimulated with an input vector, the output vector is calculated by applying the weights. The output vector represents the reaction to the stimulus (input). When the network is first created the weights have an initial state. This could be randomly assigned or assigned to be zero. The reactions to stimuli at this point will also be random or zero. The ANN in this state is useless unless it learns which action goes with which input. This learning process is performed by adjusting the weights such that for a selection of inputs, the output predictions by the ANN matches the actual output. By feeding the network many inputs together with the correct outputs, the weights can be adjusted (Simpson 1990). Once the ANN is trained it can be used.
Implementations of neural network systems in medicine include an intelligent alarm system for the detection of faults in the anesthesia breathing circuit (Farrel et al. 1992), EEG analysis to determine a patient's sedation level (Veselis et al. 1991), and OP^2, a patient outcome prediction system for anesthesia (Jackson 1992).

An artificial neural network implementation of a problem is appropriate when knowledge cannot easily be deducted from the experts, but when input and output data are available for a specific problem. ANNs could be used to implement an automatic limit

setting system, if a large number of input and output vectors were available. The input vectors are constructed from data of preoperative evaluations, of which many are available. Output vectors would be alarm limits that match the inputs. These output data are difficult to get, because they have to be assigned by an anesthesiologist. When these data are available, a ANN based system could be constructed. Another disadvantage is that neural networks are completely empirical, and that a huge network is required to solve a complex problem like ours. Neural nets can represent the grouping that is present in the inputs (see next chapter), but it is better to derive them under more control. Patients have traditionally been put into groups (healthy, old, sick, etc.), and a grouping technique matches the thinking process of the anesthesiologist more closely. The next paragraph explains pattern recognition and clustering techniques with the purpose of selecting different groups of patients.

3.4. Patient clustering approach

Clustering is the grouping of similar objects. The objects can be anything: characters, planets, cars, foods, patients, etc. Clustering, grouping or classification is something we do on a daily basis, to bring structure to the world around us.
In clustering techniques two states can be identified: 1) the state where we have n objects and we want to divide them into groups (*group creation*), and 2) the state where we have m groups and we want to assign an object to a group (*object mapping*).
Two types of measurements are needed for group creation and object mapping:

1) measurements that characterize the objects
   These are the features that characterize the object. The features have to be chosen so they fit the purpose of the grouping.
2) measurement of similarity between two objects
   This can be a measurement of distance if the objects can be described with a vector of features. Decision functions are used to assign objects to groups.

In the case of a mathematical or computer-readable description these measurements are numerical values. If we were to separate different foods with the goal of defining different food groups, we might measure fat contents and energy contents. Example data of measurements on the objects (foods in this case) are listed in Table A.1 in appendix A.

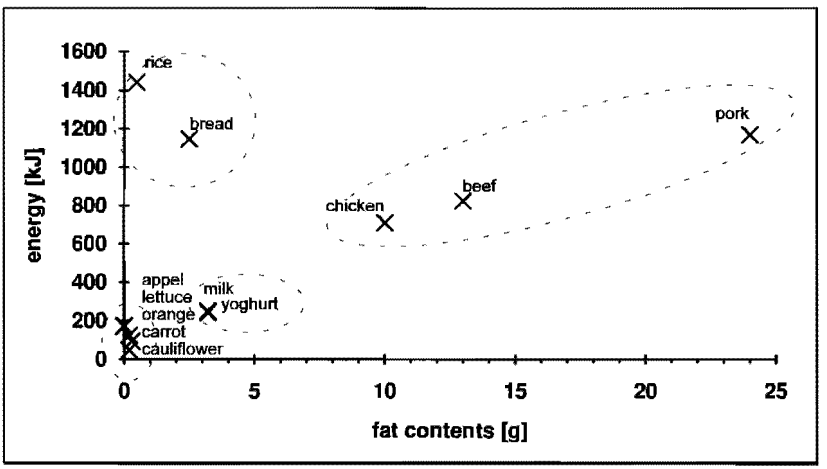


Figure 3.5: Graph of fat contents versus energy.

A graph of the two features (fat contents and energy) of the foods is shown in Figure 3.5. Numerical mathematical techniques are available to do this clustering automatically. When the k-means clustering technique is applied to food data, the groups are formed as indicated in Figure 3.5 (for a detailed description of the k-means algorithm, and for more details on this example see the next chapter).

Once groups have been defined new objects can be assigned to a group. This *mapping* is done by assigning the object to the group that most closely matches the object. A decision function using a type of similarity measurement is used to decide which group is 'closest' to the object.

Clustering analysis has been used in medicine to analyze heart sounds of patients with a porcine bioprostetic heart valve to create two patient groups: patients with a normal valve, and patients with a degenerated valve (Durand et al. 1990). Another study used clustering techniques to create 16 classes, identifying depth of anesthesia (Thomson et al. 1991). Clustering was also used to form 6 groups of undergraduates with different eating patterns (Kristeller et al. 1989).

This method can be used to automatically assign intraoperative alarm limits if we divide patients into different groups and assign a set of alarm limits to each group.

3.5 Summary

Three different methods of knowledge representation have been suggested in this chapter. The expert system approach has as a disadvantage that rules for setting alarm limits are not easily obtained, and that it doesn't adjust for changes in the practice of setting alarm limits. If the knowledge and rules can be derived via other methods, and for fixed knowledge (physical limits on physiologic parameters for example), an expert system can be used.

The neural net approach could be feasible if a large number of outputs (alarm limits) were available, although this approach does not account easily for the grouping that is present amongst patients. This method does not easily adapt to changes in the practice of setting alarm limits since the neural net has to be re-trained.

For the automatic setting of alarm limits it seems appropriate to use a clustering method to measure similarity between patients, and to put patients that are 'similar enough' (in terms of available data of the patient) in the same patient group. The patient clustering approach accounts for patient grouping, and assigns alarm limits, indicated by experts, to each group.

Changes in the practice of setting the limits can be made by modifying the assigned limits for each patient group. These modifications can be made on a continuous basis, as explained in the next chapter. Modification to the patient grouping can be made when needed.

The next chapter explains selection of a clustering algorithm and selection of the input parameters.

4 Selection of patient groups

4.1. Introduction

The previous chapter presented three techniques for the design of a system for automatic intra-operative alarm limit settings: 1) using an expert system, 2) using neural nets, and 3) using clustering techniques on patient data. This chapter will describe the patient clustering approach in detail.

We propose the following approach to derive alarm limits for individual patients:

1) select features/parameters relevant for setting alarm limits
2) collect those selected data from the (computerized) preoperative evaluation for the particular patient
3) scale and normalize these data
4) select a matching pre-defined patient group
5) look up the limits assigned to the selected group
6) present the limits for the particular patient

Steps 1) through 4) are necessary clustering. Techniques for selecting data (needed for steps 1 and 2), for scaling and normalizing (step 3), and for clustering data (step 4) are presented in paragraph 4.2. For the lookup of assigned group limits (step 5) a database of limits assigned by experts based on previous cases, or based on other methods is required. The collection of the data for this database is described in paragraph 5.2. Paragraph 5.4. describes the complete system with the implementation of the different databases (containing patient data, preop data, group data, and limit data), the definition of the flow of data, and the presentation of the alarm limits. That paragraph also explains how to account for changes in the practice of setting alarm limits, or for differences between different institutions.

4.2. Grouping Patients

Our objective is to divide patients into different groups that require different intra-operative alarm limits. Clustering techniques to achieve this goal are examined in this paragraph.

4.2.1. An introduction to clustering techniques

Cluster analysis is the study of methods for grouping or classifying objects. These objects are described by a set (or vector) of measurements or by relationships between an object and other objects. The set of measurements is sometimes called the feature vector, because it is a vector of features that describes an object. Which features to measure depends on 1) what can physically be measured, 2) what is the goal of the grouping, and 3) how correlated are the measured features.

> As an example we use data on different foods from Appendix A. Some of the possible physical measurements of foods are shown in Table A.1. The goal of the clustering of the foods is to create food groups that can be used for advice on diets. We must look closely at inter-dependency of the features to decide which ones to eliminate. For example, there is a relation between the protein contents and phosphor contents of foods. Figure 4.1 shows the results of a simple linear regression between protein and phosphor. We could decide to delete one of these two features because they are related and, therefore, one of them does not add to the independent information of the data set.

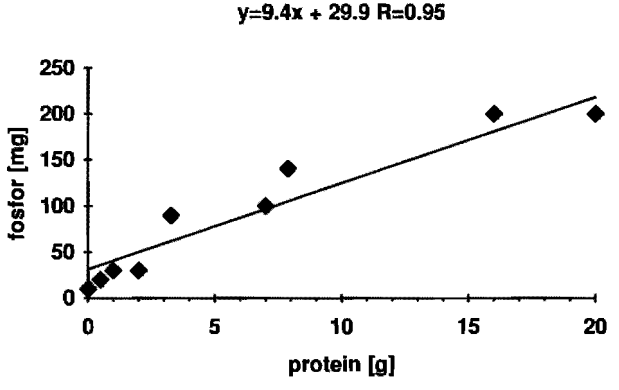


Figure 4.1: Protein contents versus phosphor contents for twelve foods.

A measurement of a quantity is made on a scale in order to relate it to another measurements of the same quantity. Before different measurement quantities can be compared, data representation and scale needs to be examined.

Data representation

Three basic data types can be distinguished: binary, discrete, and continuous. A measure of the binary type can have only two values (yes/no, 1/0), a discrete measure can have a finite, usually small, number of possible values, and a continuous measure can have any number in the real ($\mathbb{R}$) space.

A measurement scale defines how different measures relate to each other. How we should interpret the measures, depends on the type of scale. Four different scales are defined (Anderberg 1973, see Table 4.1). The nominal scale is the weakest scale because it only distinguishes between measures, and the ratio scale is the strongest scale (Jain and Dubes 1988). A stronger scale defines more about a measure than a weak scale (see "Meaningful statement" column in Table 4.1).

Table 4.1: Definitions of different data scales. Parameter x is measured for two objects (A and B) resulting in measures x_a and x_b.

Name	Description	Meaningful statement	Example	Type
Nominal	This scale only distinguishes between measures.	$x_a=x_b$, $x_a\neq x_b$	color measurement as red, blue, yellow, etc.	Qualitative
Ordinal	This scale specifies the order between the measures.	$x_a=x_b$, $x_a\neq x_b$, $x_a<x_b$, $x_a>x_b$	smallest, small, big, biggest.	Qualitative
Interval	On this scale the separation between the measures have a meaning.	$x_a=x_b$, $x_a\neq x_b$, $x_a<x_b$, $x_a>x_b$, x_a-x_b is defined.	Temperature measurement in °F.	Quantitative
Ratio	On this scale the numbers have an absolute meaning. (there is an absolute zero)	$x_a=x_b$, $x_a\neq x_b$, $x_a<x_b$, $x_a>x_b$, x_a-x_b is defined. $X = x_b/x_a Y$	Temperature measurement in °K.	Quantitative

A measure on a nominal scale can only be identified as equal or not equal to another measure. An ordinal scale measurement adds ordering to the nominal scale, and a statement can be made specifying that one measure is bigger or smaller than another measure. For the interval scale the difference of two measures is defined (e.g. a measure of 60°F is 10° different from a measure of 70°F). An absolute zero is required

for a ratio scale. For example the Fahrenheit measurement does not have a meaningful absolute zero, and the statement 40°F is twice as high as 20°F is not valid. It is true if they were measures on the temperature scale in Kelvin (40°K is twice as high as 20°K).

Numerical data analysis techniques require that data measurements are on quantitative scales of the same type. Data conversion may be necessary when essential data are only available on qualitative scales.
Qualitative scales have to be converted to quantitative scales because clustering analysis only uses quantitative measurements of proximity. When most data are measured on the *interval* scale, other measures need conversion to that scale. No conversion is necessary to represent a *ratio* scale variable on an *interval* scale, when the absolute zero is ignored. To represent a *nominal* scaled value on an *interval* scale, first a representation on the *ordinal* scale is required. This means that the measures have to be put in an order (ranking). For example, a measurement of color, where we can only decide if two colors are equal or not equal, we can define an order by using the ordering of colors in the spectrum by wavelength. We can assign numbers that represent that ordering on the ordinal scale. In the spectrum scale of colors the color violet (value 1) < yellow (value 2) < red (value 3). To represent a variable on the *ordinal* scale on an *interval* scale the difference between two measures has to be meaningful. In the color example this means that we need to use a meaningful measure of color (wavelength). On the *interval* scale violet=420 nm, yellow=570 nm, and red = 670 nm.

Using numerical matrix algebra, we can represent *m* features measured on *n* objects in a *n* x *m* matrix $\mathbf{X}$, where x_{ij} is the *j*th feature of the *i*th object. A measurement of proximity or proximity index (distance) *d(i,k)* between the *i*th and *k*th object can be calculated in different ways. A common measurement is the Euclidean distance:

$$d(i,k) = \left[\sum_{j=1}^{m} (x_{ij} - x_{kj})^2 \right]^{1/2} \qquad (4.1)$$

where x_{ij} is the *j*th feature for the *i*th object stored in matrix $\mathbf{X}$ (Jain and Dubes 1988).

After the input data are converted to the same scale and a proximity index has been defined, it is imperative that the data are normalized to ensure that certain features do not numerically 'overpower' other features. A measurement in centimeters for example can easily overpower another measurement in meters. Normalization can be done in

several ways. One could scale each parameter such that all values are in the {-1,1} range. This can be done with formula 4.2 (Gill 1981).

$$x_{ij} = \frac{2x_{ij}^*}{b_i - a_i} - \frac{a_i + b_i}{b_i - a_i} \tag{4.2}$$

where x_{ij} is the new scaled variable
x^*_{ij} is the original variable
a_i is the measured lower bound on x^*_i
b_i is the measured upper bound on x^*_i

The disadvantage of this scaling method is that outlier data points caused by noise or other reasons cause a feature to be normalized to a very narrow range. A better technique makes use of the standard deviation and mean of the measured features:

$$x_{ij} = \frac{x_{ij}^* - m_j}{s_j} \tag{4.3a}$$

where $$m_j = (1/n)\sum_{i=1}^{n} x_{ij}^* \tag{4.3b}$$

and $$s_j^2 = (1/n)\sum_{i=1}^{n} (x_{ij}^* - m_j)^2 \tag{4.3c}$$

m_j is the mean value and s_j is the standard deviation of the jth feature. Normalization with this method converts the data to have zero mean and unit standard deviation. A disadvantage of this method is that it may disguise a grouping that was apparent in the original data (Jain and Dubes 1988).
We decided to use the normalization of equation 4.3 and used Principal Component Analysis to analyze if grouping is still present after normalization.

Principal Component Analysis

When data are organized in a *n* x *m* matrix of *n* objects and *m* features, and the features were selected based on availability and on selection by experts, principal component analysis (PCA) can be used to derive alternative axes, starting with the principal axis, and continuing with less important axes. PCA can be used to reduce the number of axes of the data set, and thereby reducing the number of features. In addition, when the data are converted to 2-dimensional space, an indication of the grouping of the data can be obtained with a graphic representation. An example is shown in Figure 4.2, where three objects having two features are plotted (a).

Original data: $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 4 \end{bmatrix}$ is converted with PCA into $\mathbf{B} = \begin{bmatrix} -0.94 & -0.064 \\ -0.11 & 0.11 \\ 1.05 & -0.05 \end{bmatrix}$

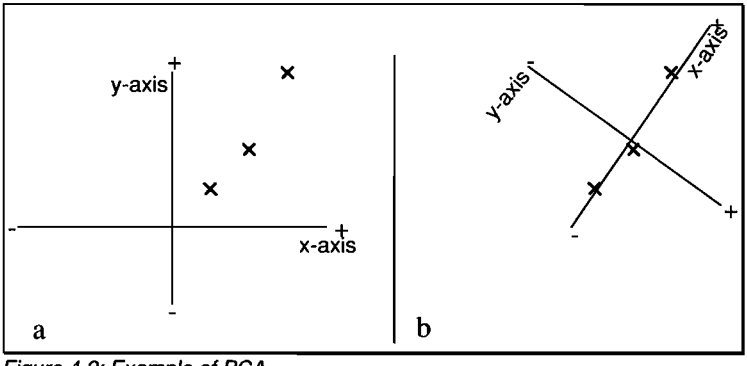


Figure 4.2: Example of PCA.

When the data are plotted on the new axes found by PCA, all the points have almost no y component. The y axis could in this case be eliminated, reducing these data to a one dimensional data set. PCA seeks the axes that the cloud of points are closest to, while maximizing the variance. This is done by minimizing the sum of distances b_i in Figure 4.3. Because $a^2=b^2+c^2$, and a is constant in this process, minimizing the sum of distances b_i is the same as maximizing the sum of c_i or maximizing the spread of the points (variance).

When a set of *n* objects with *m* features is represented by a *n* x *m* matrix $\mathbf{X}$, the objects can be regarded as row vectors in $\mathbb{R}^m$ and the features as column vectors in $\mathbb{R}^n$.
If the new axis is defined as vector $\mathbf{u}$ of unit length, then $\mathbf{Xu}$ gives the projection of the objects in $\mathbf{X}$ on the new axis. Maximizing the squared sum of c_i is equal to maximizing the squared projections of points on the new axes:

$$\begin{aligned} (\mathbf{Xu})^{\mathrm{T}}(\mathbf{Xu}) &\Leftrightarrow \\ (\mathbf{u}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}})(\mathbf{Xu}) &\Leftrightarrow \\ \mathbf{u}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{u} &\Leftrightarrow \\ \mathbf{u}^{\mathrm{T}}\mathbf{S}\mathbf{u} \text{ where } \mathbf{S} &= \mathbf{X}^{\mathrm{T}}\mathbf{X} \end{aligned} \tag{4.4}$$

Finding the maximum, while keeping in mind that $\mathbf{u}^{\mathrm{T}}\mathbf{u}=1$ is equal to finding where the derivative of

$$\mathbf{u}^{\mathrm{T}}\mathbf{S}\mathbf{u} - \lambda(\mathbf{u}^{\mathrm{T}}\mathbf{u} - 1) \tag{4.5}$$

is equal to zero:

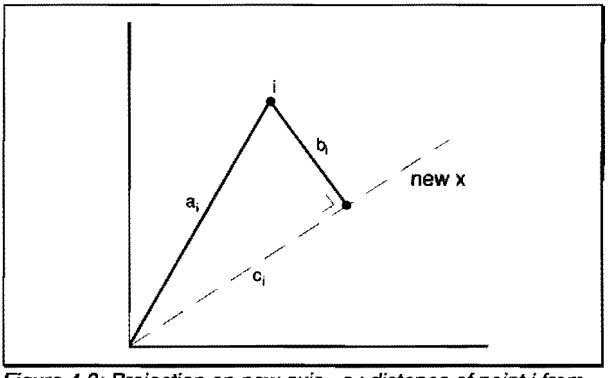


Figure 4.3: Projection on new axis. a_i: distance of point i from the origin, b_i: distance of point i from the new axis, c_i: distance of the projection of the point i from the origin. Adapted from Murtagh and Heck 1987.

$$\mathbf{2Su} - \mathbf{2\lambda u} = \mathbf{0} \Leftrightarrow$$
$$\mathbf{Su} = \lambda \mathbf{u} \tag{4.6}$$

The solution to this equation is that vector **u** is the eigenvector associated with eigenvalue λ of matrix **S** (Murtagh and Heck 1987).

When the calculated eigenvalues are labeled so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$$

(increasing importance) and the corresponding eigenvectors $\mathbf{u_1}$, $\mathbf{u_2}$, ..., $\mathbf{u_m}$ are labeled accordingly, a transformation matrix **H** can be defined as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{u_1^T} \\ \mathbf{u_2^T} \\ \vdots \\ \mathbf{u_m^T} \end{bmatrix}$$

The objects in matrix **X** can now be projected on the new axes with

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i \quad \text{for } i = 1, \cdots, n$$

where $\mathbf{x}_i$ is the object on the original axes, and $\mathbf{y}_i$ is the object on the projected axes. The object matrix **X** can be transformed with $\mathbf{Y} = \mathbf{X}\mathbf{H}^\mathbf{T}$ to form the transformed object matrix **Y**. This projection is called eigenvector transformation.

This projection has created a set of new features that are uncorrelated, which can be seen by calculating the covariance matrix of matrix **Y**:

$$(1/n)\mathbf{Y}^\mathbf{T}\mathbf{Y} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{bmatrix}$$

Eigenvector transformation can be used to reduce the number of axes (which is equal to the number of features). The eigenvectors (principal components) are in the order of importance because they were ordered by eigenvalue. When we only use the first *d* axes ($d<m$), by transforming with the first *d* eigenvectors the variance that is retained in the new space can be calculated with $\sum_{i=1}^{d}\lambda_i$. To decide how many axes to retain, *d* should be chosen so that:

$$r_m = \sum_{i=1}^{d}\lambda_i \Big/ \sum_{i=1}^{m}\lambda_i \geq 0.95 \qquad (4.7)$$

which assures that 95% of the variance is retained in the new space (Jain and Dubes 1988).

When we apply principal component analysis on the data on the contents of foods, and reduce the number of axes to two, a two-dimensional representation can be made (Figure 4.4). A separation into groups can be recognized from this graph, even though only 66% of the variance is retained (see Table 4.2). This two dimensional representation is not ideal if a large number of objects needs to be plotted. If several points have the same x- and y-coordinate they would show as only one point, which does not show the grouping of this data well. When we divide the two dimensional data space into squares, and count how many data points are contained in each square, we can create a three dimensional graph as shown in Figure 4.5. This graph shows the groups as peaks. A contour plot of the same data is shown in Figure 4.6. The contour plot shows cross sections of Figure 4.5, made at different levels of the z-axis. This figure clearly shows that grouping is present in the data.

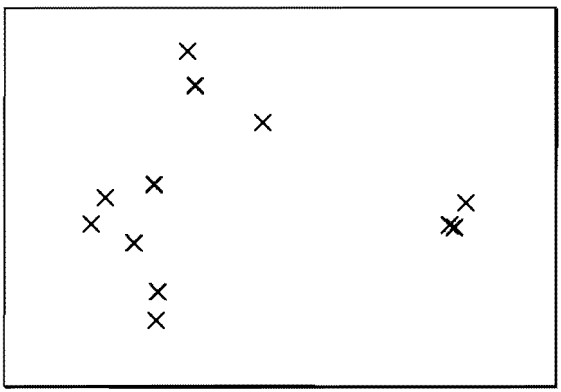


Figure 4.4: Two-dimensional projection of food data

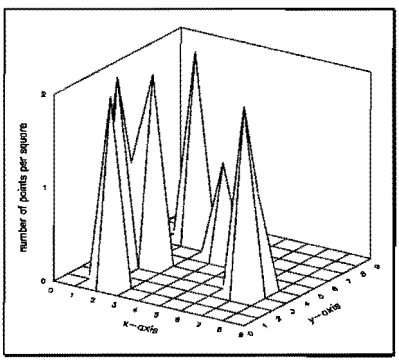


Figure 4.5: Surface plot of food data after application of a 10x10 grid.

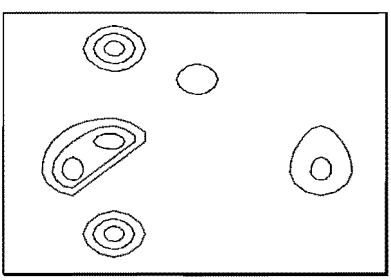


Figure 4.6: Contour plot of the food data

Table 4.2: Eigenvalues of the food example

Axis	Eigenvalue	% variance retained
1	5.42743	42
2	3.08837	66
3	1.78346	79
4	1.12069	88
5	0.89865	95
6	0.33747	97
7	0.15782	99
8	0.11619	99
9	0.05455	100
10	0.01426	100
11	0.00109	100
12	0.00003	100
13	0.00000	100

When we use the criterion that 95% of the variance should be retained when we want to reduce the number of axes, eight axes can be deleted, making the food example a 5-dimensional data set instead of a 13-dimensional set.
Principal component analysis provides a good way to reduce the number of axes in this example.

In addition to the principal component analysis, a correlation matrix can indicate measurements that are related. Based on the correlation table some features may be eliminated. This can be done instead of or in addition to principal component analysis.

Table 4.3: Correlation of food data.

	prot.	fat	sacc	Ca	P	Fe	sod.	K	vit A	vit B	vit C	H2O	E
prot.	1	0.31	-0.06	-0.43	0.95	0.92	0.73	0.49	-0.3	0.89	-0.37	-0.34	0.44
fat		1	-0.28	0.04	0.37	0.21	0.41	-0.02	-0.15	0.31	-0.32	-0.3	0.68
sacc			1	-0.26	0.02	-0.12	-0.49	-0.45	-0.14	-0.21	-0.16	-0.8	0.48
Ca				1	-0.25	-0.58	-0.02	-0.46	0.03	-0.39	-0.16	0.34	-0.25
P					1	0.84	0.7	0.36	-0.35	0.88	-0.49	-0.44	0.54
Fe						1	0.64	0.64	-0.17	0.8	-0.23	-0.23	0.32
sod.							1	0.54	0.16	0.71	-0.38	0.09	0.14
K								1	0.18	0.61	0.32	0.31	-0.22
vit A									1	-0.27	-0.09	0.28	-0.3
vit B										1	-0.21	-0.18	0.32
vit C											1	0.4	-0.46
H2O												1	-0.9
E													1

Clustering methods

Clustering is the classification of objects into groups. Properties of objects are defined as the measurement of *m* features represented in a feature vector, or points in a *m*-dimensional space. A definition of proximity or proximity index between objects is required to perform cluster analysis. Cluster analysis techniques can be divided into hierarchical clustering and partitional clustering. Hierarchical clustering methods divide the objects in a nested sequence of groups, partitional clustering methods divide the objects into single partitions (see Figure 4.7). Hierarchical clustering is impractical with more than a few hundred objects.

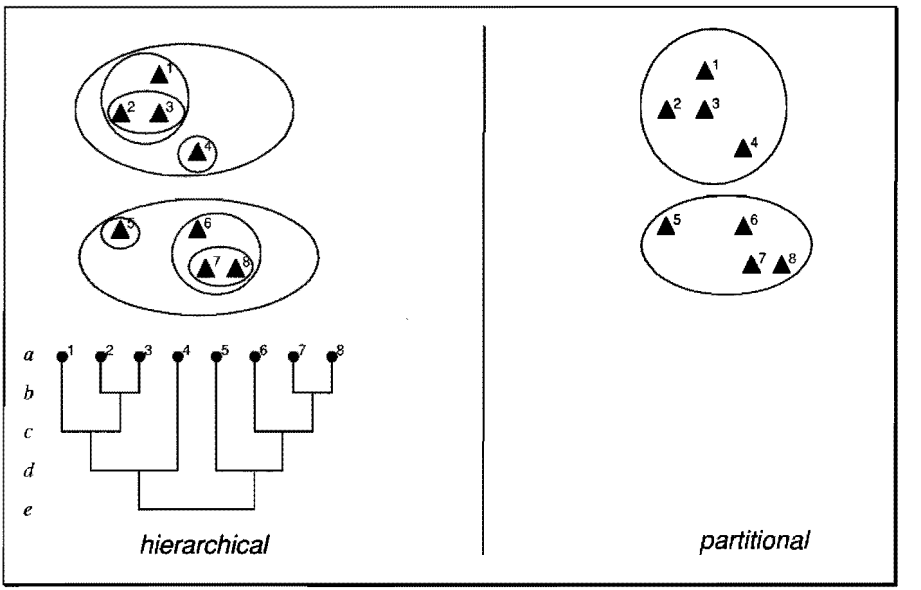


Figure 4.7:Two different types of clustering.

Hierarchical clustering methods try to find an hierarchy in the grouping. This is indicated by the tree in Figure 4.7. At level *a* in the tree no clustering has been done, and at level *e* two main groups were found. Hierarchical clustering is used if it is believed that the data being analyzed has a hierarchical structure, for example when grouping different animals (mammal→carnivore→tiger). Partitional clustering is done when all the objects have to be separated into different groups. The number of resulting groups may or may not be specified a-priori. Some clustering techniques adjust the number of groups dynamically, while others require an analysis of the data before the clustering is done. We will describe partitional clustering methods in more detail because it is the most common, and the most appropriate for our problem.

The basic procedure applied in partitional clustering is simply the selection of a criterion, the evaluation of that criterion for all possible divisions in K groups, and the selection of the division that optimizes the criterion. One example of a criterion is minimization of the sum of squared Euclidean distances between the objects and their cluster centers (minimizing E):

$$E = \sum_{j=1}^{K} \sum_{\mathbf{x} \in S_j} \left\| \mathbf{x} - \mathbf{m}_j \right\|^2 \qquad (4.8)$$

where K is the number of clusters, S_j is the set of objects belonging to the jth cluster and $\mathbf{m}$ is the sample mean vector of set S_j.

Clustering algorithms are devoted to efficiently optimize the criterion and find the best clustering. We will present some of the clustering algorithms that could be used to solve our problem. For all the algorithms we assume we have a data set of n m-dimensional vectors $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ (Tou and Gonzalez, 1974; Durand et al. 1990; Jain and Dubes 1988; Hartigan 1975).

Simple Cluster-Seeking Algorithm

This method groups together all the objects that are an arbitrary distance T from the cluster center. The following steps have to be taken:

1) Assign object $\mathbf{x_1}$ to be cluster center $\mathbf{z_1}$.
2) For each object $\mathbf{x_i}$ calculate the distance to the cluster center(s).
3a) If all the distances between $\mathbf{x_i}$ and the cluster centers are $> T$, create a new cluster with $\mathbf{x_i}$ as the cluster center.
3b) If there is one or more distance $< T$, assign $\mathbf{x_i}$ to the cluster with the smallest distance.
4) Continue with 2), until all the objects have been assigned.

In step 1) any of the objects can be assigned to $\mathbf{z_1}$.

Advantages: quick simple calculations.

Disadvantage: very difficult to establish a meaningful value for T in a multi-dimensional space. The clustering depends highly on the initial condition and the order of the objects. There is no measure of error to judge the quality of the clustering; it is a one-pass algorithm.

Maximin Distance Algorithm

This algorithm first identifies how many clusters there are in steps 1) through 4).

1) Assign $\mathbf{x_1}$ to be $\mathbf{z_1}$.
2) Find the object that is furthest from $\mathbf{x_1}$, and assign it to be $\mathbf{z_2}$.
3) For all the remaining objects, calculate their distances from the clusters centers, and save the minimum of these distances.
4) Select the object with the largest minimum distance that was saved in step 3), and assign it to be a new cluster center if this distance is at least a fraction (e.g. >½) of the minimum of the distances between the current cluster centers. If this is not the case, the algorithm exits.

After the cluster centers are found, each object is assigned to the closest center.

Advantage: Better than simple cluster seeking because it seeks appropriate cluster centers before assigning objects to them.

Disadvantage: Very dependent on initial condition, and the order of the objects. There is no measure of error to judge the quality of the clustering.

K-means algorithm

The K-means algorithm is based on the minimization of the sum of squared distances from all points to their cluster center (Equation 4.8). The number of clusters K is assigned before the clustering is done.

1) Choose an initial clustering by assigning objects to one of the cluster centers $\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_K\}$.
2) Calculate the cluster centers by calculating the sample mean vector of each cluster.
3) Re-assign all the objects to the closest (newly calculated) cluster centers.
4) If the new assignment of the objects was identical to the previous one, terminate the algorithm, otherwise continue with 2).

Advantage: Cluster centers are 'real' centers as opposed to one object from the cluster set, as in the previous two algorithms.

Disadvantage: The value for K has to be provided a-priori. A local minimum of the error may be found depending on the selection of the initial clustering.

4.2.2. Selection of a clustering method

For the clustering of patient groups a partitional clustering method was chosen, because we were not interested in the hierarchy of the groups, and only a separation between the patient groups was required. In addition, we have a large number of patients to be clustered (>5000), which is impractical with a hierarchical clustering technique.

The partitional clustering technique of choice was the K-means method, because this technique uses multiple passes to find an optimal clustering, which minimizes the impact of the initial assignment of clusters. When initial cluster centers are assigned randomly and multiple passes through the K-means algorithm are made, a global minimum of the optimization criterion is very probable. The number of groups, which

needs to be assigned a-priori in the K-means algorithm, was selected with an analysis which is described in the next paragraph.

4.2.3. Applying the K-means clustering technique

The K-means algorithm requires that the number of groups K is determined a-priori. It is important to select a valid number for K: suppose the data under investigation can be readily separated into 2 groups, but it may not be possible to separate that data into 3 groups (see Figure 4.8).

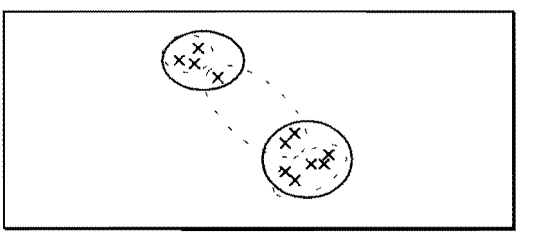


Figure 4.8: Data separation in 2 groups (solid line), and in 3 groups (dotted line).

K should be selected so that the clustering error (sum of squared distances from all points to their cluster center) is minimized when K is varied. It is expected that the error will decrease when K increases, and reach zero when K equals the number of objects. To determine a meaningful K the K-means algorithm could run starting at K=1 and incrementing K by 1 until K=10. The error is plotted versus the number of groups in Figure 4.9. The difference between the error for K=k+1 and K=k was calculated and is plotted as a bar in Figure 4.9. This difference is the improvement of the error when the number of groups is increased by 1. A good selection for K is when there is a minimum in the error (Hartigan 1975).

> For the example data a minimum can be found for K=8, but the selection of K=4 is appropriate because the improvement of higher values of K is minimal.

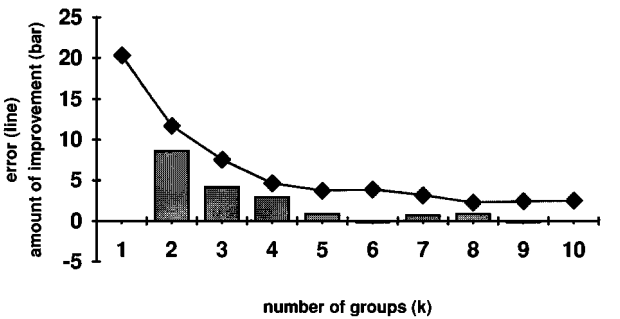


Figure 4.9: Analysis of the error vs. the number of groups for the food example data.

Clustering.

To perform K-means clustering, data normalization, and statistics, a computer program was written in C++. The C++ computer language was chosen, because of its ability to expand on the standard data type by defining classes. New classes for vector and matrix were defined. Once these classes were defined they can be used as any other data type, and operations like addition, multiplication, etc. can be performed (see Listing 4.1).

```
main()
{
      matrix a(10,10);  // a 10x10 matrix
      vector x(10);     // a 10 element vector
      vector b(10);

      ReadMatrixFile("data_a", 10, 10, a);
      ReadMatrixFile("data_x", 10, 1, x);
      b = a*x;
      b.print();
}
```

Listing 4.1: Example of the use of the vector and matrix classes

Based on these new classes the program KFRONT was written, which can read in data arrays, manipulate them, and perform K-means clustering. The following commands are supported in KFRONT:

Table 4.4: Available commands in KFRONT.

command	description
alias	show the aliases
analyze	analyze the data set for optimal K-means
close	unload the file opened by the open command
cls	Clears the screen
corr	displays correlation matrix
exit	exit the program (prompt for save)
help	displays the help screen
kmeans	kmeans clustering
norm	normalizes (MEAN=0, SD=1) the data set
open <filename> <row> <col>	loads a comma delimited file
quit	exit the program (no work saved)
save [Z|C|M|S] filename	save data
set [K|W] number(s)	sets parameters
show [Z|W]	shows a vector
stat	mean and SD of loaded array
sys <command>	execute a system command
unnorm	reverse of norm on Z and S
weight w1,w2,w3,....	Applies the weight vector

The names of the vectors and matrices are: Z the cluster center matrix, W the weight vector, S the set of matrices with the data separated in groups after running K-means. The input screen of KFRONT is shown in Figure 4.10. The program was compiled for MS-DOS computers and for HP-9000 machines running HP UNIX (hp-ux).

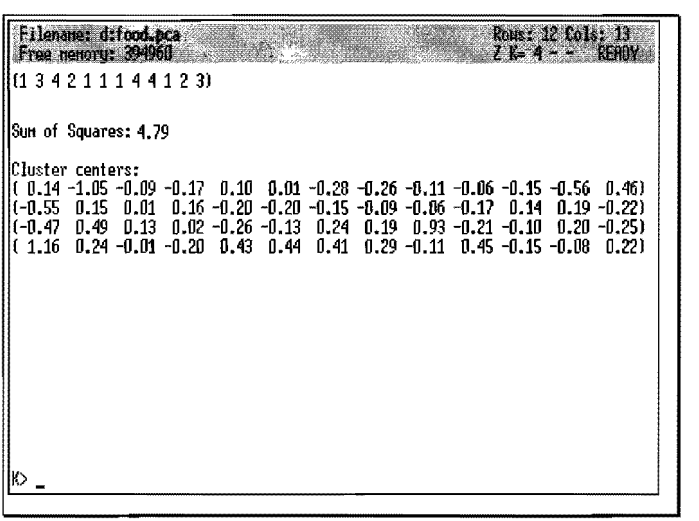


Figure 4.10: Result screen of the KFRONT program applied to the food data. The top bar indicates status information (filename, data set size, etc), and underneath it are the results of the clustering (group assignment per object, the error, and the cluster centers).

Although the program was originally designed on a DOS machine and compiled with the Borland C++ 3.1 compiler, it was ported to the UNIX environment (HP9000) and compiled with GNU's g++ compiler, because the UNIX platform has better memory management, and allows for memory blocks of virtually any size.
The K-means clustering of close to 6000 objects with 8 features takes approximately 14 minutes on the HP9000 with one single user. The DOS version is not able to use that many objects because of memory constraints.

When we perform kmeans clustering on the food data the groups of Table 4.5 result. We used the 5-dimensional PCA projection as found in the previous paragraph as input data for the clustering. The groups that were found are the expected four food groups indicated by the descriptions.

Table 4.5: Results of clustering of food data

Group 1	*fruits/vegetables*
apples	oranges
lettuce	cauliflower
carrot	
Group 2	*dairy*
milk	yogurt
Group 3	*grains*
bread	rice
Group 4	*meats*
chicken	beef
pork	

4.2.4 Applying clustering techniques to patient data

The methods outlined in the previous paragraphs will be applied to data from outpatients of Shands Hospital at the University of Florida that were scheduled for a surgical procedure. In order to separate our patients into separate groups, we have to go through the following steps (Kristeller et al. 1989):

1. Selection of the measured features.
2. Scaling and type conversion of the measured features.

3. Analysis of the principal components.
4. Selection of the number of groups.
5. Apply clustering techniques.

Selection of the measured features.

The data available to us are the data from the preoperative evaluation of the patient. Figure 3.2 in the previous chapter identifies the most important sections of the preop: the physical exam and the review of systems. All the data of the physical exam were used in the initial data set. The most important patient problems from the review of systems, in view of setting alarm limits, were selected by an expert anesthesiologist (personal communications J.S. Gravenstein, 1992) and from the literature (Balasaraswathi and El-Etr 1976; Charlson et al. 1990; Velanovich, 1991; Fleisher and Barash 1992), and are listed in Table 4.6.

Table 4.6: Important problems from the review of systems. The indicators are the words that are searched for in the preoperative evaluation to determine if a patient has the disease.

Problem	**Description** (Taber's Cyclopedic Medical Dictionary 1989)	**Indicators**
Stroke	Sudden loss of consciousness followed by paralysis caused by hemorrhage into brain, formation of an embolus or thrombus that occludes an artery, or rupture of an extracerebral blood vessel causing subarachnoid hemorrhage.	Stroke, cerebro-vascular accident (CVA).
Shock	A clinical syndrome in which the peripheral blood flow is inadequate to return sufficient blood to the heart for normal function, particularly transport of oxygen to all organs or tissues.	Shock
Coronary Artery Disease	Decreased flow of blood to the heart muscle to the extent that either basal needs of oxygen are unmet or the oxygen supply is insufficient when an increased demand for oxygen is made, as in work.	Angina, Exercise intolerance and chest pain
Hyperthyroidism	A condition caused by excessive secretion of thyroid hormone, which increases the basal metabolic rate, causing an increased demand of substrate to support this metabolic activity.	Hyperthyroid
Myocardial Infarction (MI)	Condition caused by occlusion of one or more of the coronary arteries.	MI
Heart Block	Condition in which the conductile tissue of the heart, the sinoatrial and atrioventricular nodes, bundle of His, Purkinje fibers, fails to conduct impulses normally from the atrium to ventricles, or within the ventricles	A-V Heart Block (1st, 2nd or 3rd degree), Left Bundle Branch Block (LBBB), Right Bundle Branch Block (RBBB)
Aneurysm	Localized abnormal dilatation of a blood vessel, usually an artery, due to congenital defect or weakness of the wall of the vessel.	Aneurysm

Scaling and type conversion of the measured features.

Table 4.7 lists the features, their range and scale that were used as an initial data set.

Table 4.7: Features describing a patient (for a description of the scale types see Table 4.1).

Feature	Range	Scale
Age	0-120	ratio
Sex	M/F	nominal
Weight	0-250 kg	ratio
Height	0-250 cm	ratio
Systolic BP	0-250 mmHg	ratio
Diastolic BP	0-250 mmHg	ratio
Heart Rate	0-250 bpm	ratio
ASA class	I,II,III,IV,V	ordinal
Procedure Code	0-9999	nominal
Stroke	0/1	ordinal
Shock	0/1	ordinal
Coron Art Disease	0/1	ordinal
MI	0/1	ordinal
Hyperthyroid	0/1	ordinal
Heart Block	0/1	ordinal
Aneurysm	0/1	ordinal

To analyze these data a homogeneity of scale types is required. We selected the interval scale, because conversions to this scale from other scales is feasible. Some information is lost when a conversion is done from the ratio to the interval scale (no absolute zero is necessary for the interval scale). The next section describes the scale conversions necessary on the data of table 4.7.

Scale conversions

For the conversion of sex on the nominal scale to the ordinal scale, it is necessary to decide if a male patient has a larger influence on setting alarm limit than a female patient, or if the opposite is true. We are not able to determine this, and since the sex of a patient is not one of the most important factors in setting alarm limits, this feature was omitted. ASA physical status is on a ordinal scale because a higher ASA status implies a larger influence on how alarm limits are set. With the help of an expert anesthesiologist we came up with a scaling of the ASA parameter to the interval scale as shown in Figure 4.11 (personal communications J.S. Gravenstein, 1993). We call this parameter ASA index (ASAi).

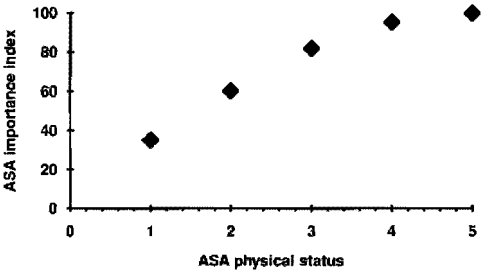


Figure 4.11: Assignment of ASA importance index (ASAi).

No scaling is necessary for the age, weight, height, systolic and diastolic blood pressure, and heart rate, because they are already measured on the interval or ratio scale.

The 7 problems (Table 4.6) on the ordinal scale (yes/no), require conversion to an interval scale. To achieve this we determined the risk range for each problem. If a problem would cause a high value of the related parameter the risk range was called high and given the value 1. Conversely, a low parameter value yielded a low risk range. Table 4.8 lists the problems, related parameter, and risk range (personal communications J.S. Gravenstein, 1992).

Table 4.8: Effects of problems.

Problem	Related parameter	Risk range
Stroke	Blood Pressure	high
Shock	Blood Pressure	low
Coron Art Disease	Heart Rate	high
Hyperthyroid	Heart Rate	high
Heart Block	Heart Rate	low
Aneurysm	Blood Pressure	high

MI was eliminated from the data set because it is an indication for coronary artery disease, and therefore not an independent parameter.

To convert the problem list to an interval scale, the problems with the same related parameter (e.g. blood pressure) were grouped together and the values were added. This method assigns an equal weight to each problem.

For the conversion of the procedure code (Table 4.7) to the interval scale the basic relative value of the procedure as proposed by the American Society of

Anesthesiologists was used. Basic relative values are proposed by the ASA to assist anesthesiologists in the development of their fee schedule, and are related to the complexity of the anesthetic service (ASA 1992). The basic relative value ranges from 3 (e.g. anesthesia for surgery on the knee) to 30 (anesthesia for liver transplant).

After the initial data set is scaled to the interval scale, the list of features in Table 4.9 results and is used in subsequent calculations.

Table 4.9: List of features used in the analysis.

feature	Unit	Symbol
Age	years	age
Weight	Lb.	wt
Height	cm	ht
Systolic Blood Pressure	mm Hg	BPsys
Diastolic Blood Pressure	mm Hg	BPdia
Heart Rate	beats per min.	HR
ASA physical status index		ASAi
Relative value of Procedure code		rel.val.
BP problem index		BPi
HR problem index		HRi

Analysis of the principal components.

We used principal component analysis to get an appreciation of the spread and shape of the groups, and to decide if PCA can be used to reduce the dimensionality of our data.

When data are reduced to two dimensions they can be plotted to get an appreciation of the spread of the data and the shape of the groups. The two principal axis are combinations of several features.

PCA was done on the scaled, normalized data and the two principal components are plotted in Figure 4.12. The same data are plotted as three dimensional graph after dividing the two dimensional space of Figure 4.12 into 50x50 squares and counting how many points are in each square. A contour plot at different levels of the z-axis of Figure 4.13 is presented in Figure 4.14. Grouping can be seen in this figure as is indicated by the dotted ellipses.

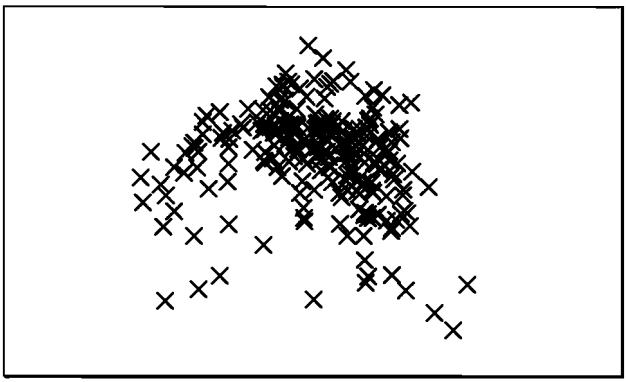


Figure 4.12: Two-dimensional projection (with PCA) of part of the patient data

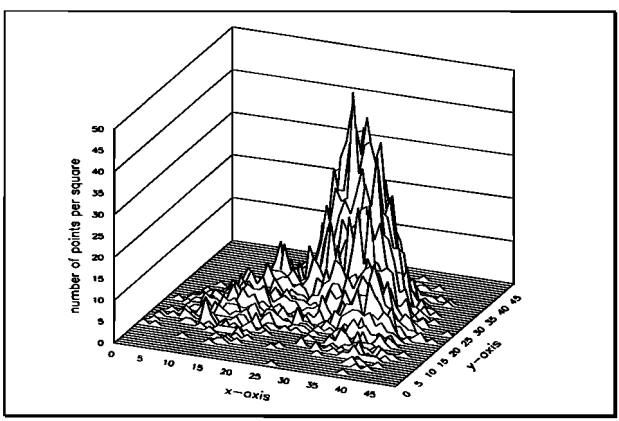


Figure 4.13: Three-dimensional plot of the patient data after a PCA projection on two axis.

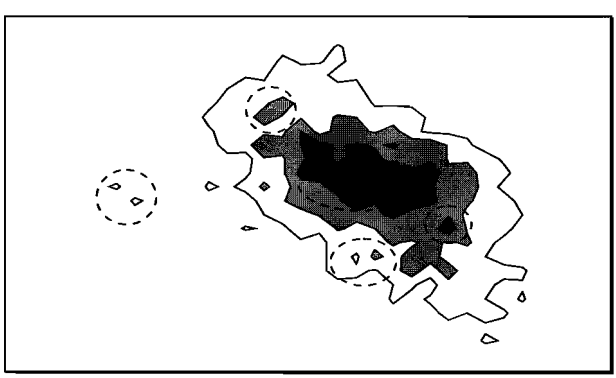


Figure 4.14: Contour plot of the patient data. The contours were made at the 10, 20, 30, and 40 level of the z-axis of the previous graph.

Table 4.10 lists the eigenvalues of the data with the percentage of variance retained. Figures 4.12 through 4.14 are made from the two-dimensional projection in which only 47% of the variance is retained (see Table 4.10), and they should be interpreted with this in mind. The two-dimensional projection gives an indication of the spread of the data, and shows some places with a higher density of points (grouping).

Table 4.10: Results of PCA

Axis	Eigenvalue	% variance retained
1	3.28	33
2	1.39	47
3	1.09	58
4	1.01	68
5	0.89	77
6	0.69	84
7	0.63	90
8	0.42	94
9	0.32	97
10	0.27	100

When we use the criterion of equation 4.7 that 95% percent of the variance should be retained, we can reduce the data set to a 8-dimensional data set.

In addition to PCA, the correlation matrix of the data can be used to analyze which features are most correlated. If one feature is highly correlated with another, one should be eliminated because it does not aid in the separation of groups.

Table 4.11: Correlation matrix of the data set.

	Age	weight	height	BPsys	BPdia	HR	ASAi	rel. val.	BPi	HRi
Age	1.00	0.45	0.48	0.54	0.40	-0.33	0.43	0.09	-0.13	0.27
weight		1.00	0.69	0.40	0.39	-0.34	0.19	0.07	-0.02	0.12
height			1.00	0.33	0.31	-0.49	0.14	0.05	-0.02	0.09
BPsys				1.00	0.64	-0.16	0.26	0.06	-0.09	0.17
BPdia					1.00	-0.07	0.22	0.05	-0.04	0.11
HR						1.00	-0.01	0.00	0.03	-0.06
ASAi							1.00	0.18	-0.11	0.27
rel. val.								1.00	-0.02	0.04
BPi									1.00	-0.12
HRi										1.00

The correlation matrix in Table 4.11 represents the correlation between two parameters. This value ranges from 0 (no correlation) to 1 (complete correlation). The two highest values are 0.69 between height and weight, and 0.64 between systolic and diastolic

blood pressures. These relationships are expected. We decided to eliminate the height measurement and the diastolic blood pressure measurement to obtain a less correlated data set. To support the decision of eliminating these two features, the PCA was run again on the 8 feature data set (Table 4.12). No further axis can be eliminated if 95% of the variance should be retained. We decided to use the unconverted data as opposed to the data converted with PCA, because analysis show that similar grouping is found with either data set. The parameters in the unconverted data set have a physiologic meaning, and is easier interpreted by physicians.

Table 4.12: Results of PCA after elimination of two features.

Axis	Eigenvalue	% variance retained
1	2.48	31
2	1.21	46
3	1.01	58
4	0.90	69
5	0.82	79
6	0.66	87
7	0.56	94
8	0.47	100

In order to determine the number of groups K, the clustering error was plotted versus the number of groups and the difference between the error for K=k+1 and K=k was calculated and is plotted as a bar in Figure 4.15. After analysis of the error for K between 1 and 20, we selected K=8, because the clustering error is minimized and the amount of improvement is the greatest for K=8.

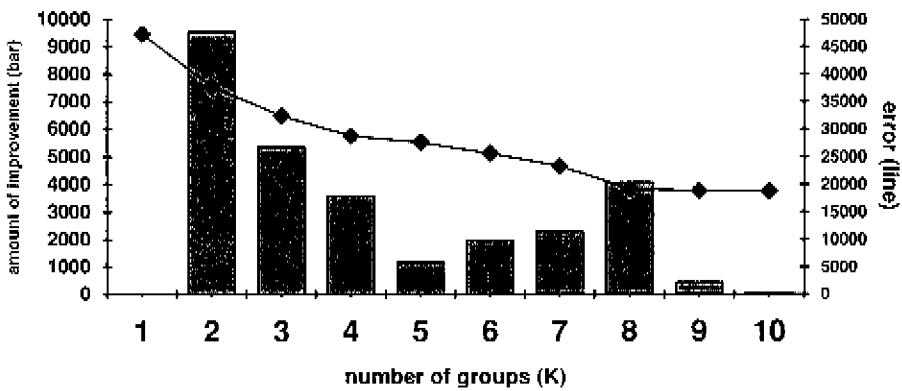


Figure 4.15: Analysis of the error vs. the number of groups.

4.2.4. Discussion of clustering results

Data of 5916 patients were collected from the preoperative evaluation system of the Department of Anesthesiology at the University of Florida. After conversion of ASA physical status to ASA index, and conversion of the problems to HR related, and BP related problems, the data had the following mean and standard deviation:

Table 4.13: Averages and standard deviation of 5916 patients.

	age	weight	BPsys	HR	ASAi	rel.val.	BPi	HRi
mean	40.39	150.76	130.98	80.19	57.64	5.78	-0.02	0.11
SD	23.11	58.51	21.81	15.92	15.59	2.15	0.17	0.33

K-means clustering of these data, after normalization, resulted in the groups of Table 4.14. The value for K was 8, and normalization of zero mean, and unit standard deviation were performed. The patients of each group can be described as shown in Table 4.15.

Table 4.14: Clustering results for K=8, normalized input data.

	age	weight	BPsys	HR	ASAi	rel. val.	BPi	HRi
1	36.27	162.54	121.87	77.93	63.05	5.44	0.00	0.00
2	6.36	48.98	109.56	102.70	50.63	5.30	0.00	0.01
3	27.42	153.16	124.58	74.79	35.00	5.06	0.00	0.00
4	62.18	176.09	151.27	76.93	64.07	5.48	0.00	-0.01
5	43.80	164.91	133.30	78.64	63.02	12.02	0.00	0.06
6	64.12	158.63	145.31	76.46	74.56	6.28	-1.00	0.43
7	60.03	174.36	142.39	77.01	70.26	5.74	0.00	1.03
8	43.50	147.94	131.35	79.91	72.53	6.74	1.00	0.18

Table 4.15: Description of cluster groups.

1	average age, ASA I or II patient
2	very young, healthy patient
3	young, healthy, ASA I patient
4	older healthy patient with high BP
5	average age, difficult procedure, potential problems with HR
6	older, sick, ASA II or III patient with multiple problems
7	older, sick ASA II or III patient with problems with HR, but not BP
8	average age, sick ASA II or III patient with multiple problems

In order to validate the clustering, a distance table between the cluster centers can be calculated:

Table 4.16: Normalized cluster center distance table.

Cluster Centers	z_1 n=1635	z_2 n=750	z_3 n=1068	z_4 n=1421	z_5 n=325	z_6 n=134	z_7 n=549	z_8 n=34
z_1	0.00	2.96	1.87	1.77	3.13	6.37	3.46	6.09
z_2		0.00	2.92	4.19	4.52	7.40	5.09	6.84
z_3			0.00	2.72	3.80	6.88	4.20	6.56
z_4				0.00	3.26	6.17	3.21	6.19
z_5					0.00	6.75	4.26	6.50
z_6						0.00	6.26	12.01
z_7							0.00	6.60
z_8								0.00

From the distance table (Table 4.16) it can be seen that cluster center z_8 is relatively removed from the other cluster centers. Because of the low *n* (34), one could consider grouping these patients into another group (probably group 1, the closest), but because this group is relatively far removed from the other groups, it is a separate entity. Groups 1, 3, and 4 are relatively close together, but because of the large number of patients in each group they are accepted as separate groups (Tou and Gonzalez, 1974).
The variances of a cluster about its mean can be used to get a measure for the relative distribution about the cluster center of the features (see Table 4.17). This shows us for example that ASA class is an important feature of cluster 3, because the standard deviation equals zero, showing that all the patients in that group have the same ASA classification.

Table 4.17: Standard deviations per group, per parameter.

	age	w	BPs	HR	ASAi	rel.v	BPi	HRi
cluster 1	14.41	44.85	12.41	12.29	7.66	1.33	0.00	0.04
cluster 2	6.05	27.69	15.99	17.63	16.16	1.58	0.00	0.10
cluster 3	13.16	44.16	15.50	11.71	0.00	1.41	0.00	0.05
cluster 4	13.60	45.03	17.09	13.18	10.16	1.44	0.00	0.10
cluster 5	19.82	50.00	19.43	12.72	12.81	1.88	0.00	0.23
cluster 6	15.71	39.79	23.25	11.94	11.95	2.15	0.00	0.56
cluster 7	15.10	43.61	22.45	13.39	12.58	1.77	0.00	0.17
cluster 8	22.26	43.61	22.37	19.72	13.12	2.80	0.00	0.38

4.3 Summary

In this chapter we have presented all the aspects of clustering techniques (data normalization, scaling, etc.). We applied those techniques to preoperative data of 5916 patients and formed eight groups. The next chapter will describe how limit data for the patient groups will be collected and how all the data fits together.

A System for Alarm Limit Setting in Anesthesia

5.1 Introduction

In the previous chapter we performed a clustering analysis on preoperative patient data, which yielded 8 patient groups. To be able to set alarm limits for each patient, we assign a limit set to each patient group. Individual patients are assigned the limits of the group they belong to. Data collection for setting alarm limits is presented in this chapter. Also a description of the complete system is given, showing how all the parts (patient groups, limit assignment, preoperative data summary, etc.) fit together, and how data flows between them.

5.2. Gathering Alarm Limits

Once the patient groups are defined, each group needs a set of limits assigned to it. To answer the question of what limits to select for each patient group, some definitions are required. First we have to define what type of limits we call 'the alarm limit'. Figure 5.1 shows the different levels that can be identified:

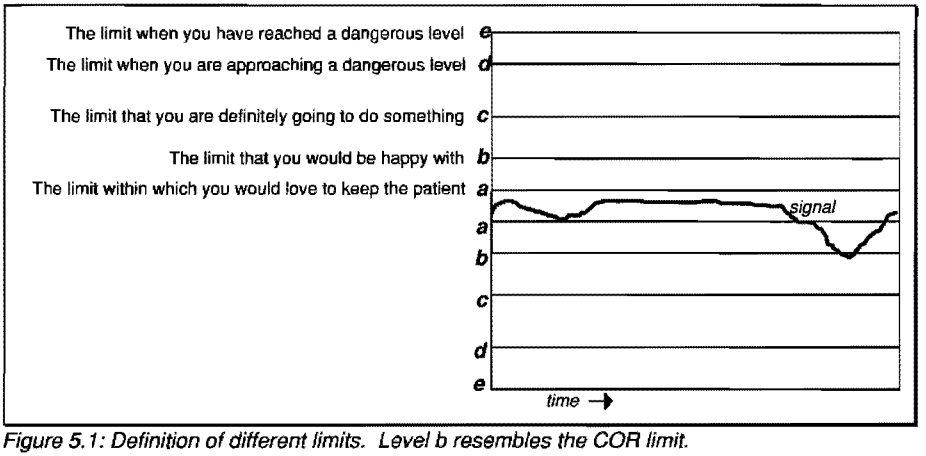


Figure 5.1: Definition of different limits. Level b resembles the COR limit.

We defined the clinical operating range (COR) earlier (paragraph 2.3) as the range that the parameter is expected to move within. In Figure 5.1 this would be the range b.

Beneken and van der Aa (1989) define three functions of alarms: 1) to assist the anesthesiologist in the detection of adverse or unexpected situations occurring in the patient and the equipment; 2) to step in when fatigue decreases the vigilance level of the clinician; and 3) to assist in those situations where the comprehension capabilities of the clinician are stretched to the limit either by an overload of signals and extracted features, or by lack of clinical knowledge or experience. A more complex alarm strategy is needed to achieve the last function by combining several signals and features to aid the physician in comprehending the situation (van der Aa 1990). To achieve the first two functions of an alarm system it would be most appropriate to set limits at level c in Figure 5.1.
Typically the tighter the limits are set, the more false alarms are generated. False alarms can be generated by the following causes:

- Artifact
- Expected variation
- Improperly set limits

There is very little an alarm system can do about artifacts. Artifacts are picked up by the sensors of the monitors and should be eliminated, if possible, before the signal is passed on to the alarm system[1]. The alarm system can only assume that the measured parameters it uses as inputs are correct. Intraoperative signals encounter variations that are sometimes expected by the clinician (expected variation). Expected variation is a problem for the alarm system. Variation can be expected by the clinician based on knowledge about what will happen next: e.g. awaiting incision, or spontaneous breathing. It can also be caused by current events: e.g. blood pressure increases at the moment of incision. An alarm can also be classified as an undesired alarm because the anesthesiologist was already aware of the situation and is treating it. Solutions to the expected variation problem include more intelligent alarm systems, that know more about what is going on around it (Schecke et al. 1992; Kahn et al. 1991).
The alarms that we are selecting for the patient groups are designed to reproduce the limits currently used by anesthesiologists (not always used on the monitors, but a least in their heads) in the operating room.

[1]we make a distinction here between monitors (the measuring instrument), and alarm system (the system that generates the alarm), although in most devices the two are combined.

5.2.1. Methods for limit selection

Several methods for the collection of limits were considered, and are listed in Table 5.1

Table 5.1: Methods for the collection of data on limit setting.

method	**description**	**advantage/disadvantage**
Automatic recording in the O.R.	Physiologic data from the monitors are automatically recorded. These data are used to derive the limits that would be statistically correct in terms of maximizing favorable outcome.	*advantage*: Limits are based on statistics to maximize outcome. *disadvantage*: Expensive to record in many O.R.'s. Impractical because there are many different monitors. If no precise record of what went on is available, interpretation is difficult.
Limit selection on the preoperative evaluation form.	Physicians indicate on the preoperative evaluation which limits they deem appropriate for the patient	*advantage*: The physician has a good impression of the patient because he just completed an evaluation. *disadvantage*: The resident that does the preoperative evaluation is not necessarily the one that takes care of the patient in the O.R.
Limit selection intra-operatively	Someone asks intra-operatively which limits are appropriate for that specific patient.	*advantage:* The case is underway, and there is a good understanding about the patient. *disadvantage*: The anesthesiologist is distracted during the operation.
Limit selection by expert anesthesiologists from data on patients	Expert anesthesiologists derive limits they deem appropriate for a patient based on the preoperative evaluation, or a summary of it.	*advantage*: Many patients can be quickly evaluated. *disadvantage*: The provided data may not be enough to make a decision.

The last two methods were implemented and the results of the data collection are presented in the next paragraph.

5.3 Group limit assignment

Data on how limits are set for different patients were collected from two sources: expert anesthesiologists, and anesthesia faculty and residents in the operating room. Anesthesia faculty of the University of Florida were given preoperative evaluations of a selection of patients that were operated on in Shands teaching hospital in the past 3 years. The patients were selected from the 5916 patients of the previous chapter that had complete preoperative evaluations captured electronically. Initially 80 patients were selected randomly. To assure that there were enough data for each patient group, additional patients were selected randomly per group.
Three expert anesthesiologists were presented with full evaluations (printout of the complete preoperative evaluation), and limited evaluations (printout showing only surgical procedure, age, sex, weight, height, preoperative systolic and diastolic blood pressure, heart rate, ASA physical status and whether the patient had any of the following problems: stroke, shock, coronary artery disease, hyperthyroidism, heart block, and aneurysm; see Table 4.7).
The expert anesthesiologists were asked to assign limits at which they would like to be alerted during the maintenance phase of anesthesia. They indicated upper and lower limits for systolic and diastolic blood pressure, heart rate, SpO_2, and end tidal CO_2. In order to test the hypothesis that the complete and limited evaluation produce the same limits several evaluations were presented in both complete and in limited form. The anesthesiologists were not informed about that fact.
A total of 106 cases were reviewed by the expert anesthesiologists.

In order to assign the 106 cases to one of eight groups, summary data were extracted from the preoperative evaluation, the ASA physical status was converted to the ASA index, and the problem list was converted to the blood pressure index (BPi), and the heart rate index (HRi), resulting in a patient vector. After normalization of this vector the Euclidean distances between each patient group (Table 4.14) and the patient vector were calculated, and the patient was assigned to the closest (minimum distance) cluster.
The limits assigned by the experts were averaged by group and are presented in Table 5.2. The description of the groups is listed again in Table 5.3 (same as Table 4.15).

Table 5.2: Limits assigned by anesthesiologists. Indicated as means per group with the standard deviation in small print.

group	**BPsys**		**BPdia**		**HR**		**$EtCO_2$**		**S_aO_2**	
(n)	upper	lower	upper	lower	upper	lower	upper	lower	upper	lower
1 (24)	162.9	91.5	97.1	50.0	103.3	55.1	48.4	31.3	100.0	92.0
	15.9	10.6	7.9	6.3	10.7	4.0	1.5	1.8	0.0	1.7
2 (13)	143.1	83.8	92.7	45.0	133.8	65.4	47.5	30.8	100.0	92.0
	17.7	9.6	8.5	6.2	15.5	8.4	2.1	1.8	0.0	2.3
3 (15)	152.7	87.3	99.0	50.0	108.7	55.7	48.1	31.1	100.0	92.1
	14.8	9.3	8.9	6.1	15.4	4.4	1.5	2.0	0.0	1.8
4 (17)	181.8	121.5	101.5	62.0	97.6	57.1	47.8	31.6	100.0	90.2
	8.6	11.0	4.8	6.0	7.1	6.2	1.8	2.0	0.0	6.8
5 (10)	162.0	104.0	97.0	60.0	94.5	54.0	43.5	32.0	100.0	93.6
	19.9	12.8	6.8	8.4	8.5	5.4	4.1	5.2	0.0	2.5
6 (9)	175.6	121.1	94.4	58.3	90.0	55.0	46.8	32.7	100.0	93.2
	24.6	22.0	8.3	14.1	6.7	7.8	2.9	3.0	0.0	2.9
7 (9)	164.4	102.2	98.9	47.8	93.3	48.9	47.6	30.3	100.0	91.9
	18.3	5.8	11.0	7.5	8.2	4.6	2.4	0.9	0.0	1.4
8 (9)	157.8	100.0	108.9	44.4	105.6	47.2	48.9	30.0	100.0	92.7
	26.8	12.9	11.0	5.0	13.4	4.8	3.1	0.0	0.0	0.9

Table 5.3: Description of cluster groups.

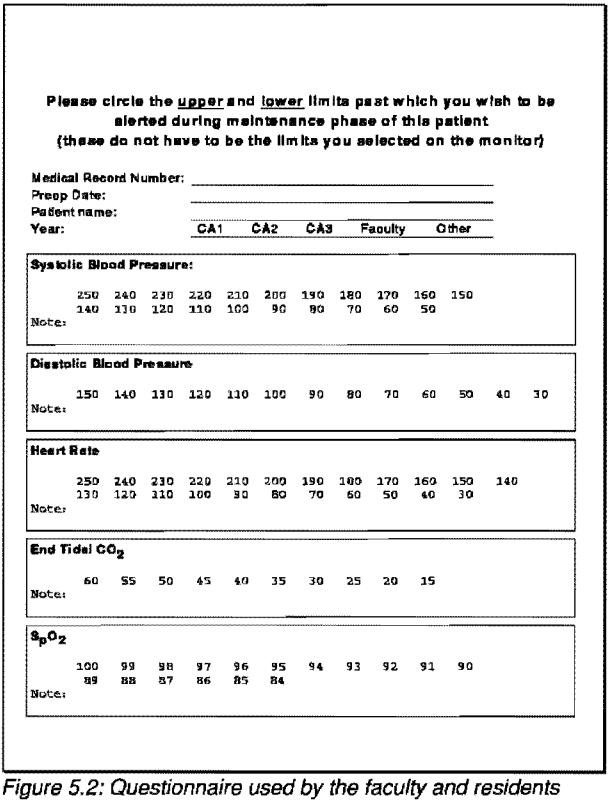
1	average age, ASA I or II patient
2	very young, healthy patient
3	young, healthy, ASA I patient
4	older healthy patient with high BP
5	average age, difficult procedure, potential problems with HR
6	older, sick, ASA II or III patient with multiple problems
7	older, sick ASA II or III patient with problems with HR, but not BP
8	average age, sick ASA II or III patient with multiple problems

The assigned limits show some results that we expected: the blood pressure limits for patients with high blood pressure (group 4) are set higher than in the other groups, and the heart rate limit is set the highest for very young patients (group 2). Limits for end tidal CO_2 and SpO_2 were set similar for all the groups. An evaluation of these results will be made in chapter 6.

In order to create an independent test population, faculty and resident anesthesiologists were asked in the operating room during the maintenance phase of anesthesia to indicate the limits that they were using in their mind during that phase for the current case (these are typically not the limits set on the monitors). We used a pre-printed form to take up as little of their time as possible (see Figure 5.2). When an electronic

preoperative evaluation was available for the patient, the date of the evaluation and the patient's medical record number were noted, as this is enough to retrieve the record from our preoperative evaluation system. When no electronic preoperative evaluation was present, the preoperative information was obtained from the handwritten evaluation.

Data of 49 cases were collected this way, and were used for the performance testing of the system, which is presented in chapter 6.



Please circle the upper and lower limits past which you wish to be alerted during maintenance phase of this patient (these do not have to be the limits you selected on the monitor)

Medical Record Number: _____
Preop Date: _____
Patient name: _____
Year: CA1 CA2 CA3 Faculty Other

Systolic Blood Pressure:

250 240 230 220 210 200 190 180 170 160 150
140 130 120 110 100 90 80 70 60 50
Note:

Diastolic Blood Pressure

150 140 130 120 110 100 90 80 70 60 50 40 30
Note:

Heart Rate

250 240 230 220 210 200 190 180 170 160 150 140
130 120 110 100 90 80 70 60 50 40 30
Note:

End Tidal CO_2

60 55 50 45 40 35 30 25 20 15
Note:

S_pO_2

100 99 98 97 96 95 94 93 92 91 90
89 88 87 86 85 84
Note:

Figure 5.2: Questionnaire used by the faculty and residents

All the data that are needed for the limit setting system are now complete. The next paragraph will describe how to manage these data and define how they are used when limits are to be assigned for a patient.

5.4 Database system design

5.4.1 Introduction to databases

A database system can be defined as a system that provides a structure for data in order to make information available on demand (Date 1986). Databases are typically used when large amounts of data need to be organized, easily and quickly retrieved, and made accessible to multiple users. Database system architecture can be divided into three general levels:

- *internal level*

  This level defines the physical storage structure of the database. Definitions of physical storage (hard disk etc.), methods of data compression, and indexing methods to facilitate quick searching are part of this level. It also defines the file format(s) of the database system.
- *external level*

  This level defines the way data are presented to the users. It defines the interactions between user and data (for example SQL (structured query language) or QBE (query by example) data queries). It also defines how data are viewed by the users (e.g. in tables, as data forms, etc.).
- *conceptual level*

  This level sits between the two previous levels and provides 'translation' services. It translates requests from the user into commands that can be understood by the internal level

If a user for example requested a record he may issue the command `GET RECORD #123`, which gets translated by the conceptual level into `GET INTEGER FROM #123; GET DATE FROM #123;` (assuming a record consists only of an integer number and a date). The internal level converts this commands into instructions that the hard disk can understand: `READ 10 BYTES START OFFSET 444555;`.
A graphical representation of these three levels is given in Figure 5.3.

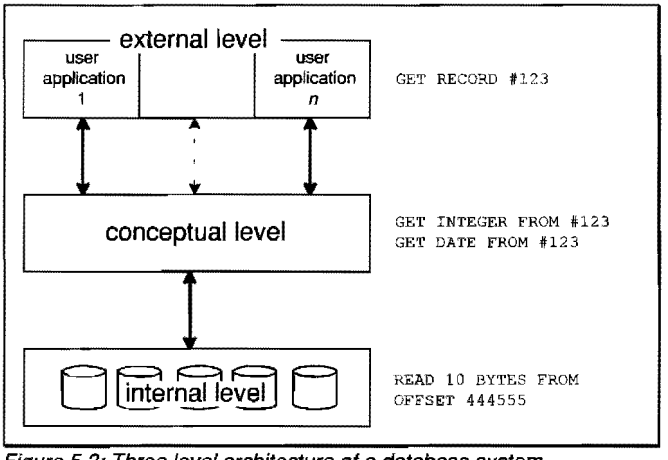
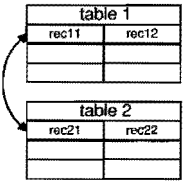


Figure 5.3: Three level architecture of a database system

There are three major data models that implement all three levels of the architecture of a database system: the relational data model, the hierarchical data model, and the network data model (Elmasri and Navathe, 1989):

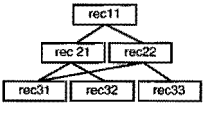
relational data model

In the relational data model the data in a database are represented as a collection of tables. Relations can be indicated by linking columns of different tables.



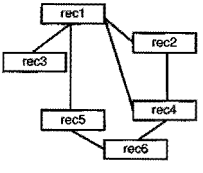
hierarchical data model

In the hierarchical data model the data are represented as a tree of records. The structure of the tree defines the links between the records.



network data model

In the network model the data are represented by a directed graph. The links of the graph define the links between the records.



When selecting a database system several factors play a role:

- which data model is most appropriate

- what is the hardware platform
- what is a reasonable cost

For our system we have several sets of data: patient demographic information, average preoperative summaries per patient group, and limits assigned to each group. These data sets do not stand alone, there are relations or links between them.
Relationships in our data can be M:N (we may find M patients that were assigned to group 4, and we may find N limits assigned by experts to group 4), and because of this the hierarchical data model is not appropriate because M:N relationships are difficult to implement in a hierarchical (tree) structure.
In the relational data model the data tables are linked by entries that have the same value (e.g. medical record number 12345 in Figure 5.4). In the network data model links have to be specified explicitly per record (Figure 5.5). It is easier to implement links between different data tables in a relational data model, and it is therefore currently the most popular data model for database systems (Elmasri and Navathe, 1989).

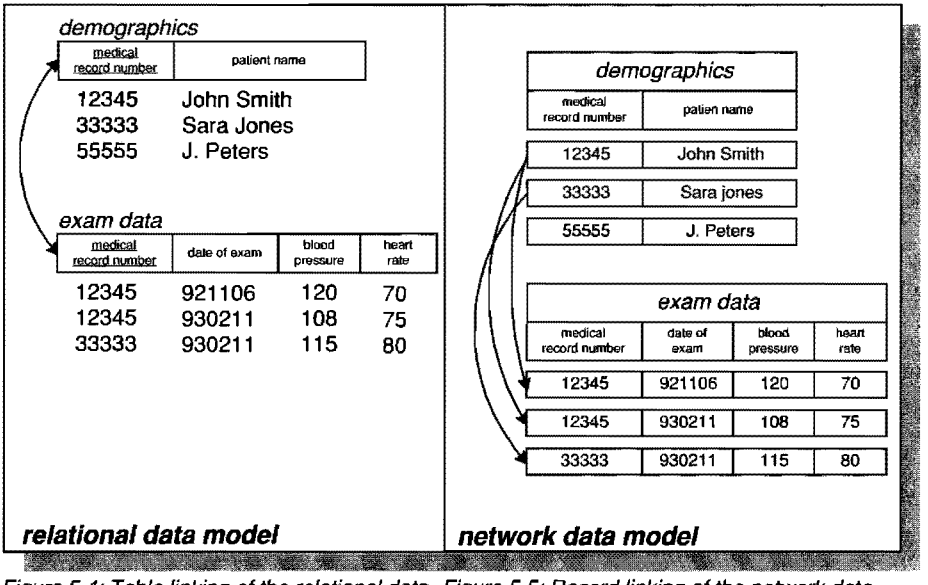


Figure 5.4: Table linking of the relational data model

Figure 5.5: Record linking of the network data model

We decided on the Paradox (Borland International Inc., Scotts Valley CA) relational database because it works well on the MS-DOS platform, and is cheaply available at educational discount price.

5.4.2 Definition of database

The system for alarm limit setting relies on many data: *demographic patient data* to identify the patient that limits are being set for, the *preoperative summary* of that patient with the assignment to one of the patient groups, the *limit data* that were assigned by experts for each patient, and the *results of the clustering algorithm* that is needed for the assignment of a patient to a group. In addition *scaling data* (the means and standard deviations of the 5916 cases) are needed to scale the preoperative summary data before a group assignment is made (scaling is done by subtracting the mean, and dividing by the standard deviation, see paragraph 4.2.1).

All the data outlined above was put in database tables. Table 5.4 lists the five data tables that were defined.

Table 5.4: Data tables of the database.

data table	**description**	**Table**
minimal demographic data	to store data about a patient that will not change once entered	5.5a
preoperative summary	excerpt of the preoperative evaluation per case.	5.5b
limit database	upper and lower alarm limits for a specific case assigned by expert or in the OR	5.5c
patient groups database	results of the clustering of the preoperative data of the patients	5.5d
scaling data	the means and standard deviations of the 5916 cases that were used for the initial clustering	5.5e

Tables 5.5 show the fields in the database tables, where fields in italics represent lookup keys into other tables. The next paragraph explains how these data tables are used. The data types are N: numerical, D: date, and A: text.

Table 5.5a: Minimal demographic data

MRN	*N*	*Medical Record Number*
DOB	D	Date of Birth
Sex	A1	Male/Female

Table 5.5b: Preoperative summary

MRN	*N*	*Medical Record Number*
OR Date	*D*	*Date of the procedure*
Preop Date	D	Date of the preoperative evaluation
Age	N	Age of the patient
Weight	N	Weight
BPsys	N	Systolic Blood Pressure
HR	N	Heart Rate
ASAi	N	ASA index
BPi	N	BP problem index
HRi	N	HR problem index
Rel.val	N	Relative value of procedure
Group ID	*N*	*Patient Groups assignment*

Table 5.5c: Limit database

MRN	*N*	*Medical Record Number*
OR Date	*N*	*Date of the procedure*
Group ID	*N*	*Patient Groups assignment*
BPsys upper	N	Upper limit of the systolic BP
BPsys lower	N	Lower limit of the systolic BP
BPdia upper	N	Upper limit of the diastolic BP
BPdia lower	N	Lower limit of the diastolic BP
HR upper	N	Upper limit of the heart rate
HR lower	N	Lower limit of the heart rate
SpO2 upper	N	Upper limit of the SpO2 (typically 100)
SpO2 lower	N	Lower limit of the SpO2
PetCO2 upper	N	Upper limit of the end tidal CO2
PetCO2 lower	N	Lower limit of the end tidal CO2

Table 5.5d: Patient groups database

Group ID	*N*	*Patient group assignment*
Age	N	Average group age
Weight	N	Average group weight
BPsys	N	Average group Systolic Blood Pressure
HR	N	Average group Heart Rate
ASAi	N	Average group ASA index
BPi	N	Average group BP problem index
HRi	N	Average group HR problem index
Rel. val	N	Average group Relative value of the procedure

Table 5.5e: Scaling data

m_Age	N	Age of the patient (mean)
sd_Age	N	Age of the patient (standard deviation)
m_Weight	N	Weight (mean)
sd_Weight	N	Weight (standard deviation)
m_BPsys	N	Systolic Blood Pressure (mean)
sd_BPsys	N	Systolic Blood Pressure (standard deviation)
m_HR	N	Heart Rate (mean)
sd_HR	N	Heart Rate (standard deviation)
m_ASAi	N	ASA index (mean)
sd_ASAi	N	ASA index (standard deviation)
m_BPi	N	BP problem index (mean)
sd_BPi	N	BP problem index (standard deviation)
m_HRi	N	HR problem index (mean)
sd_HRi	N	HR problem index (standard deviation)
m_Rel.val	N	Relative value (mean)
sd_Rel.val	N	Relative value (standard deviation)

5.4.3 Definition of data flow

Upon completion of the preoperative evaluation for a particular patient our system will start working to find appropriate alarm limits for that patient. The following steps are taken:

- A preoperative summary is created by extracting data from the preoperative evaluation (Table 5.5b)
- A lookup is performed in the demographic database to see if this patient has been operated on before (Table 5.5a). If this is the case, the preoperative summary of the previous procedure is displayed, and the physician is given the option to select the limits of the previous case, or to continue with the next step.
- The preoperative summary data are scaled (Table 5.5e), and the distances from the data in the patient groups database (Table 5.5d) are calculated. The patient is assigned the groups ID of the closest group.
- A lookup of all the limits for the selected group ID is performed in the limit database (Table 5.5c), and the user is presented with the average of those group limits, and with the number of patients the average is based on.

The physician can change the selected limit before or during the operation. At the end of the operation these updated limits should be fed back into the limits database, and these data then become a new entry for the patient group the patient was assigned to.

When this procedure continues over time, the calculated limits from the database can change. If alarm limits would be set closer over time, the limit database would also update to reflect this.

Over time it may be necessary to run the K-means algorithm again to make a new assignment of the groups, although it is not expected that the patient population will change rapidly. Figure 5.6 shows the steps: 1) extraction of a preoperative summary from the preoperative evaluation, 2) check to see if this patient has a set (or sets) of limits stored in the database of previous visits, present these limits; if no previous visit was found, or the physician was not pleased with the limits (for example if the type of operation is very different from the previous one), then 3) assign the patient to the closest cluster, and calculate the average limits for the cluster.; 4) this average is presented with the number of limit sets it is based upon. When during the operation the anesthesiologist determines that the selected limits were not appropriate 5), they can be changed. The changed limits can be fed back into the database after the operation 6), and may gradually change the average group limits to account for changes in clinical

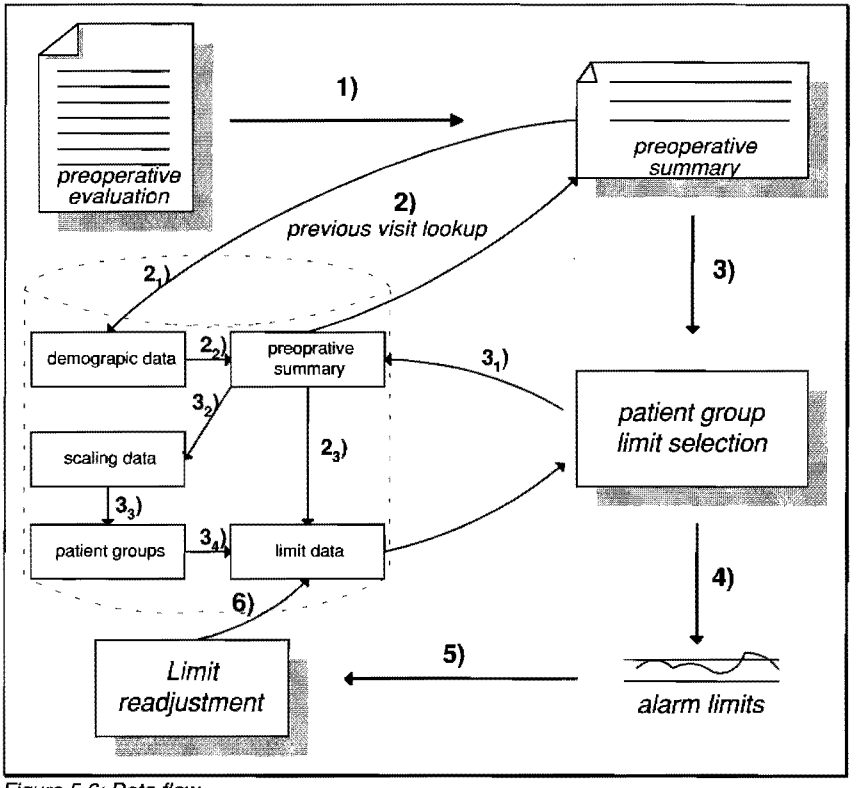


Figure 5.6: Data flow.

practice.
The previous visit lookup in step 2) consists of: 2_1) check if demographic data on the patient is present, 2_2) find preoperative summaries with 2_3) matching limit sets, and present them to the physician.
The patient group limit selection in set 3) consists of: 3_1) store the new preoperative summary, 3_2) scale the summary data for group matching, 3_3) find the closest matching patient group, and 3_4) look up the limits of patients with that group id, and calculate the average limits.

5.5 Summary

In this chapter the limits assignment for each group were presented. All the parts were put together to form the system for alarm limit setting in anesthesia. It explained how all the data are used, and how they are stored in the different database tables. Validation of the system is presented in the next chapter where we take a closer look at the assigned limits in relation to the patient clustering. Results of comparisons between the limits assigned by the system and the limits used in the O.R. for a group of patients are also presented.

Analysis of performance

6.1 Introduction

Before we can analyze the performance of our system, we first have to show that the design goals were met. The study was designed to achieve the following goals (see also paragraph 1.3: project objective):

1. Describe the patient with a summary of the available data.
2. Create different patient groups.
3. Assign different limits to the different patient groups.
4. The alarm limits set by the system based on the patient's preoperative information are similar to what a clinician would use intraoperatively.

The following hypotheses will be tested to assure that these goals are met:

A There is no difference between limits set by experts based on the full preoperative evaluation and based on a summary of the evaluation (goal 1).

B Different limits are set for different patient groups (goal 2, 3).

C The set of patients used to test the system is representative of the total patient population (goal 4).

D Limits set by the system are similar to the limits used intraoperatively by the clinicians (goal 4).

These hypotheses are evaluated in the next paragraphs.

6.2 Validation hypotheses

For the testing of the hypotheses we distinguish between three data sets: 1) the data of 5916 computerized preoperative evaluations that we used to create patient groups (see chapter 4); 2) data from expert anesthesiologists on how to set limits (106 cases, see chapter 5); and 3) operating room data from anesthesia residents and faculty of the limits they are using intraoperatively (49 cases). The second data set consisted of limited and complete preoperative evaluations (106 patients total), of which seventeen

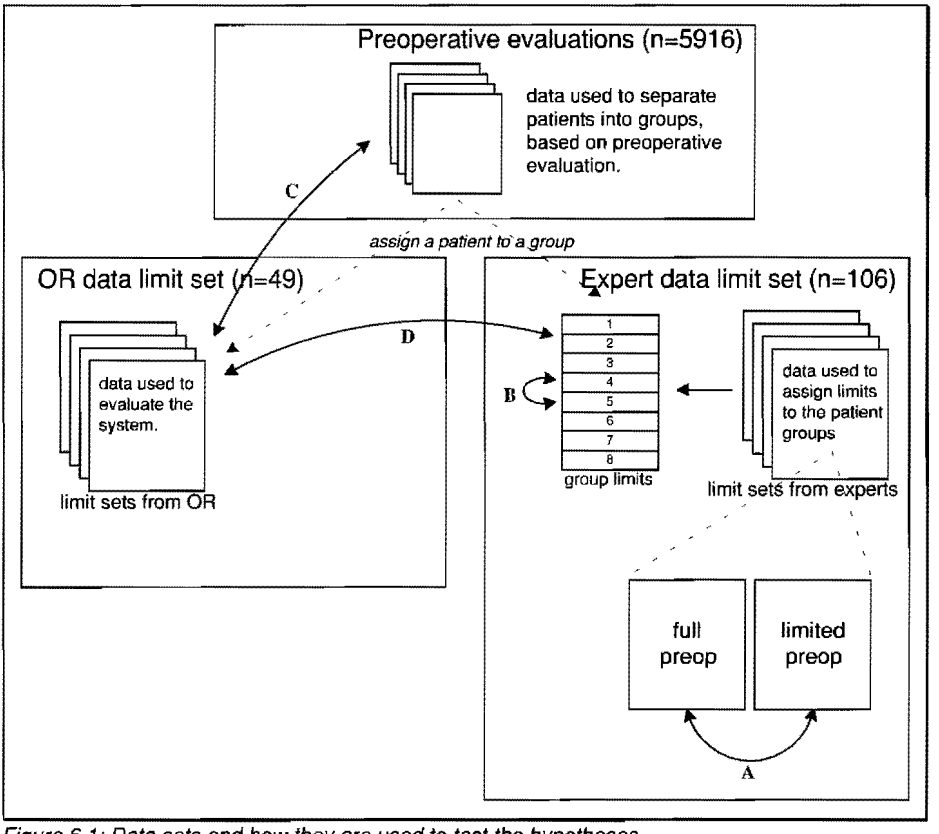


Figure 6.1: Data sets and how they are used to test the hypotheses.

Hypotheses tested:

- **A** *Difference between full and limited preops?*
- **B** *Different limits for different groups?*
- **C** *Test group representative?*
- **D** *Performance of the system?*

were presented in both full and limited form for the same patient (for a definition of full and limited preoperative evaluations see paragraph 5.3).

In Figure 6.1 the tests performed on the three data sets are indicated by letters A through D, according to the above mentioned hypotheses. The patients were assigned to groups based on the group separation of the K-means clustering algorithm of chapter 4 (indicated by the dotted lines in Figure 6.1). To test the hypothesis that there are no differences between limits set by experts based on the full preoperative evaluation and based on a summary of the evaluation the seventeen cases that were presented in both forms were compared (A). The validation of the resulting group limits, by determining if they were different for each group, was based on the evaluation of the grouped data

from the 106 patients that the experts assigned limits for (B). To evaluate if the test group was representative of the total patient population, the data of the 49 cases of the test were compared with the data of the 5916 preoperative evaluations (C).
Performance of the system was tested by comparing the limits assigned by the expert anesthesiologists to the limits desired by the residents and faculty in the operating room (D).

6.2.1 Patient sample is a representative sample of the population.

In order to interpret the performance of the system, we must show that the group of 49 patients is representative for the whole population (hypothesis C).
To test the hypothesis, we define the 5916 patients on which we recorded preoperative information as the whole population. After assuring that the distributions of the two data sets approximate the normal distribution, a two-tailed student t-test was performed to compare the distinguishing features of the 49 patients (our sample), with those of the total population (the 5916 patients). The results are tabulated in Table 6.1.

Table 6.1: Preoperative information of the 49 patients, and the total population. The p value is the probability of randomly obtaining a mean difference as large (or larger) than the one observed, when in fact there is no such difference in the overall population.

n		Age	weight	BPsys	HR	ASAi	rel.val.	BPi	HRi
49	**mean**	40.71	149.64	136.56	81.56	59.47	6.00	-0.022	0.067
	SD	22.26	53.05	22.68	17.96	14.19	1.78	0.15	0.25
5916	**mean**	40.39	150.76	130.98	80.19	57.64	5.78	-0.02	0.11
	SD	23.11	58.51	21.81	15.92	15.59	2.15	0.17	0.33
	p	0.92	0.89	0.07	0.55	0.41	0.93	0.36	0.48

Because we are attempting to show that there is no significant difference between the set of two features, we have to analyze what the power of the t-test is.
Suppose we are testing the hypothesis H_0: $P_1=P_2$ (population 1 is the same as population 2). The alternative hypothesis is: H_a: $P_1 \neq P_2$ (the two populations are different). The decision when to accept H_0 can be seen from Figure 6.2.

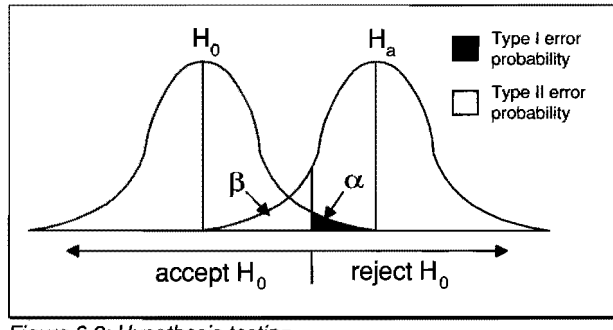


Figure 6.2: Hypothesis testing.

Two types of errors can be made when we decide to accept or reject hypothesis H_0: rejecting H_0 when it should have been accepted (Type I error, the probability is α, typically 0.05), and accepting H_0 when it should have been rejected (Type II error, the probability is β, in the medical field typically 0.10 to 0.20 (Shuster JJ 1990)). The power of the test is defined as the probability of correctly rejecting H_0 when it should have been rejected. This probability is defined as 1-β.
From Figure 6.2 we can also see that increasing α simultaneously decreases β. The sample size also has an impact on α and β: when the sample size increases, α and β both decrease.

An estimation of the type II error can be made if we decide the amount of difference we consider significant (δ) and estimate the population standard deviation (σ). When estimates for δ and σ are made, a lookup table or graph of the power function can be made to determine the power of a test for a specific number of samples (Glantz 1992). If we for example compare age of the total population (in our case the 5916 patients, with a variance of 23 years) with the age of the selected 49 patients, we may decide that a difference of 10 years is a significant difference. Lookup of the power of a t-test with these parameters gives a power of 90%.

We need to decide what would be a significant difference, and what is the standard deviation of the total population. An estimate of the standard deviation of the total population is the standard deviation of the 5916 patients.
Table 6.2 shows the difference we would call significant, and the power of the t-test if we had a measurement of 45 samples. The significance level was selected by considering the range of the variables and the number of groups (8) that we selected. For example

weight ranges from 1 Lb. to 300 Lb., which for eight groups gives a difference of at least 37.5 Lb.

Table 6.2: Power of the t-test for p<0.05, n=45

parameter	significance	power of test
Age	10 yr.	90%
wt	35 Lb.	95%
BPsys	10 mmHg	90%
HR	10 bpm	90%
ASAi	35	95%
rel.val.	1	90%
BPi	0.08	90%
HRi	0.1	90%

Table 6.1 shows that there are no statistical differences between the total population of patients and the sample that we took from this population when we consider the preoperative evaluation.

6.2.2 No differences between full and limited preoperative evaluation.

The seventeen cases that were presented to the expert anesthesiologists in both limited and full preoperative evaluations were compared to determine if limits are set differently (hypothesis A). For all the five signals (systolic and diastolic blood pressure, heart rate, SpO_2, and $EtCO_2$) a paired two-tailed t-test was performed on the upper and lower limits for the pairs of measurements of full and limited preoperative evaluations. The two data sets approximated a normal distribution. The results of these tests were all not significant with a p value > 0.05 (see Table 6.3).

Table 6.3: Results of paired two-tailed t-test for limits based on limited and full preoperative evaluations.

	BPsys		**BPdia**		**HR**		**$EtCO_2$**		**SpO_2**
	upper	lower	upper	lower	upper	lower	upper	lower	lower
limited preop									
mean	156.2	93.5	96.5	48.4	104.7	54.1	48.2	31.4	92.1
sd	25.1	10.7	7.0	6.3	13.7	6.7	1.7	1.5	1.5
full preop									
mean	150.9	87.9	100.6	48.5	111.2	54.7	47.4	30.5	91.8
sd	17.6	14.1	9.0	6.8	19.6	6.5	2.2	2.4	2.0
p	0.24	0.07	0.09	0.75	0.06	0.68	0.16	0.21	0.60

For seventeen samples and p<0.05, the t-test we performed has a power of 80% if we are trying to detect a difference of more than one standard deviation. This estimate of

the power is a worst-case estimate because it assumes $p<0.05$. For some of the tests the power is much higher because of the large p values. However, an estimate for the power of 80% is acceptable (Glantz 1992; Shuster JJ; Lemeshow et al. 1990).

Because there is no difference in the way limits are set between full and limited preoperative evaluations, we can conclude that the limited preoperative evaluation provided the anesthesiologists with enough information to set alarm limits. Our automatic system uses the same data as the limited preoperative evaluation, and we can conclude that we provide our system with a sufficient set of data.

6.2.3 Different limits are set for different patient groups

We must determine if the limits that were indicated by the experts (the 106 cases) for the different groups are different between the groups (hypothesis B). The automated assignment of patients to a group were based solely on the preoperative information (the limited preoperative evaluation), so it is possible that there are groups that require a similar limit set. If there are groups that require the same limit set, the limit data should be pooled in step 3_4 in Figure 5.6.

To evaluate differences between the limit sets of the groups, we used analysis of variance (ANOVA) to determine if the limit sets are generally different and the Student-Neuman-Keuls (SNK) to test statistical differences between specific groups. The ANOVA test is based on the F-test statistic (Glantz 1992):

$$F = \frac{\text{population variance estimated from sample means}}{\text{population variance estimated as averages of sample variances}}$$

The numerator is called the between-groups variance and the denominator is called the within-groups variance. This statistic can be calculated with:

$$F = \frac{s_{bet}^2}{s_{wit}^2}. \text{ where}$$

$$s_{bet}^2 = \frac{\sum_{}^{k} n_g \overline{X}_g^2 - \frac{\left(\sum^{k} n_g \overline{X}_g\right)^2}{N}}{k-1}$$

$$s_{wit}^2 = \frac{\sum^{k} (n_g - 1) s_g^2}{N-k}$$

For group g: n_g is the size of the sample in , $\overline{X}_g$ is the group mean, s_g is the group standard deviation. There are a total of N samples and k groups.
Since both the numerator and the denominator are estimates of the variance of the same population, F should approximate 1. The ANOVA test provides threshold values of F to decide if the means were different when we accept an error in this decision of smaller than 5% ($p<0.05$). These threshold values depend on the degrees of freedom in the numerator (k-1) and the degrees of freedom in the denominator (N-k).
As an example we calculate F for the upper limit of the systolic blood pressure. There are 8 groups (k=8), and a total of 106 samples (N=106), and the means and standard deviations are listed in Table 5.2. The within-groups variance s_{wit}^2 can be calculated with:

$$s_{wit}^2 = \frac{23(15.9)^2 + 12(17.7)^2 + 14(14.8)^2 + 16(8.6)^2 + 9(19.9)^2 + 8(24.6)^2 + 8(18.3)^2 + 8(26.8)^2}{106-8}$$

$$s_{wit}^2 = 312.8$$

The between-groups variance s_{bet}^2 can be calculated with:

$$s_{bet}^2 = \frac{24(162.9)^2 + 13(143.1)^2 + 15(152.7)^2 + 17(181.8)^2 + 10(162)^2 + 9(175.6)^2 + 9(164.4)^2 + 9(157.8)^2}{8-1} - \frac{\left((24 \cdot 162.9) + (13 \cdot 143.1) + (15 \cdot 152.7) + (17 \cdot 181.8) + (10 \cdot 162) + (9 \cdot 175.6) + (9 \cdot 164.4) + (9 \cdot 157.8)\right)^2 / 106}{8-1}$$

$$= 2063.1$$

Dividing s_{bet}^2 by s_{wit}^2 gives F:

$$F = \frac{2063.1}{312.8} = 6.60$$

A lookup table provided for the ANOVA test determines that for N=106, k=8, and F=6.60, the difference among the group means is statistically significant (Glantz 1992).

The ANOVA test also showed statistical significant differences between the group means for the upper and lower limits of all the other measurements, except for the lower limit of $EtCO_2$ and for both SpO_2 limits.

The Student-Neuman-Keuls (SNK) test is a multiple comparison test. The student t-test is designed to compare only two group means. If more than two groups are to be compared, a multiple comparison test like the Bonferroni test or the SNK test is required. We selected the SNK test because we are comparing 8 groups (28 comparisons[1]), which makes the Bonferroni test underestimate the difference between the groups because too many comparisons are done (Glantz 1992). The SNK test uses the statistic q to determine statistical difference between two groups :

$$q = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\frac{s^2_{wit}}{2}\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$

where $\overline{X}_A$ and $\overline{X}_B$ are the two means being compared, s^2_{wit} is the within-group variance estimated from the analysis of variance, and n_A and n_B are the number of samples in group A and B respectively.

The tables that follow (Tables 6.4 through 6.8) show the results of these tests. They indicate if the mean limit of one group differs (u for upper limit, l for lower limit) from another group. If there were no differences between two groups, the box is empty. For a description of the groups see paragraph 4.2.4.

Table 6.4 SNK test results for systolic blood pressure

groups	1	2	3	4	5	6	7	8
1		u		u l	l	l		
2				u l	l	u l	l	l
3				u l	l	u l	l	
4					u l		l	u l
5						l		
6							l	l
7								
8								

[1]To compare 8 values we need to make 7+6+5+4+3+2+1=28 comparisons

Table 6.5 SNK test results for diastolic blood pressure

groups	1	2	3	4	5	6	7	8
1				|	|	|		u
2				|	|	|		u
3				|	|	|		u
4								|	|
5								|	u |
6								|	u |
7										
8										

Table 6.6 SNK test results for heart rate.

groups	1	2	3	4	5	6	7	8
1		u				u		|
2			u |	u |	u |	u |	u |	u |
3					u	u	u |	|
4						u	|	|
5								
6								|
7								
8								

Table 6.7 SNK test results for $EtCO_2$

groups	1	2	3	4	5	6	7	8
1					u			
2					u			
3					u			
4					u			
5						u	u	u
6								
7								
8								

Table 6.8 SNK test results for SpO_2

groups	1	2	3	4	5	6	7	8
1								
2								
3								
4								
5								
6								
7								
8								

No significant differences were found for the lower limit of $EtCO_2$, and the lower limit of SpO_2 (note: the upper limit for SpO_2 is always 100%). We can conclude that anesthesiologists do not set these limits based on information about the individual patient. One approach is to assign these limits to be the average for the 106 cases. The following definition was used to determine if a set of limits for one group was different from all the other sets of limits of the other groups:

> *a set of limits for one group is different from the other groups if one or more limits of the set are statistically (SNK test) different from all the other groups.*

If for example all the limits for groups 1 and 5 are the same, except for the upper limit of systolic blood pressure, we conclude that the set of limits of group 1 is different from the set of limits of group 5.

When we apply that criterion Table 6.9 results, indicating x for difference and o for indifference. It can be seen that no differences were found between groups 1 and 3, groups 1 and 7, and groups 7 and 8.

Table 6.9 SNK test results for all the signals

groups	1	2	3	4	5	6	7	8
1		x	o	x	x	x	o	x
2			x	x	x	x	x	x
3				x	x	x	x	x
4					x	x	x	x
5						x	x	x
6							x	x
7								o
8								

We decided to merge the limits of groups 1 and 3 and also the limits of groups 7 and 8 because the limit sets are similar.
After recalculating the SNK statistics for the new groups and applying the criterion of difference, Table 6.10 results, where we can see that there are statistical differences between all the sets of limits of the new combination of groups.

Table 6.10 SNK test results for all the signals after merging groups.

groups	1/3	2	4	5	6	7/8
1/3		x	x	x	x	x
2			x	x	x	x
4				x	x	x
5					x	x
6						x
7/8						

After recalculating the averages of the limits set by the expert anesthesiologists when merging the groups the limits are assigned as shown in Table 6.11.

Table 6.11: Limits assigned by anesthesiologists after merging groups

group	BPsys		BPdia		HR		$EtCO_2$		S_nO_2	
(n)	upper	lower	upper	lower	upper	lower	upper	lower	upper	lower
1/3 (39)	159.0	89.9	97.8	49.4	105.4	54.1	48.3	31.26	100.0	92.0
2 (13)	143.1	83.8	92.7	50.0	133.8	65.4	47.5	30.8	100.0	92.0
4 (17)	181.8	121.5	101.5	60.0	97.6	57.1	47.8	31.6	100.0	90.2
5 (10)	162.0	104.0	97.0	60.0	94.5	54.0	43.5	32.0	100.0	93.6
6 (9)	175.6	121.1	94.4	58.3	90.0	55.0	46.8	32.7	100.0	93.2
7/8 (18)	161.1	101.1	103.9	46.1	99.4	48.0	48.2	30.2	100.0	93.2

We can conclude that, upon merging, 8 different patient groups require 6 different sets of limits during the maintenance phase of anesthesia.

6.2.4 Performance

To test the performance of the system we have to compare the limits the system suggests (based on the expert opinions) to the limits that are used in the operating room (hypothesis D, see Figure 6.1).
For each of the eight groups we compared the set of limits suggested for a group by the system to the set of limits for a particular patient belonging to that group, and indicated by residents and faculty in the operating room.

To determine how close the limits set by the system are to the limits used intraoperatively we determined the percentage of limits that fall within x percent above or below the limits used in the O.R. We varied x between 5 and 25 %. The results are tabulated in Table 6.12[2].

Table 6.12: Percentage of correct limits per signal (49 patients).

	BPsys		BPdia		HR		$EtCO_2$		SpO_2	Average
% from limit	upper	lower	upper	lower	upper	lower	upper	lower	lower	
5%	51	40	40	18	24	36	33	22	89	39
10%	58	55	62	33	56	56	51	40	98	56
15%	80	82	87	71	62	76	82	60	98	76
20%	84	84	91	73	80	82	89	84	98	85
25%	91	93	93	78	84	87	98	89	98	90

When we accept limits that are within 15% of the limits used intraoperatively the systems suggests 76% of the limits correctly. The system performs better for blood pressures than for heart rate or end tidal CO_2. We speculate that a reason may be that blood pressure limits are selected using a more standard method by physicians (there is more agreement on how these limits should be set), which makes it easier for our system to select correct limits.

We can also compare the results of Table 6.12 with a simple limit setting system that has only one group (see Table 6.13). When we use the means of the limits selected by the experts for the 106 cases, this would represent a system better than the current technology because the 'factory default' alarm limit setting is based on experienced anesthesiologists. We also evaluated the factory default limits of the HP 1176A Merlin monitor (Hewlett-Packard 1992). This is one of the systems currently in use in the operating room (Table 6.14).

Table 6.13: Percentage of correct limits per signal for a simple (one group) system (49 patients).

	BPsys		BPdia		HR		$EtCO_2$		SpO_2	Average
% from limit	upper	lower	upper	lower	upper	lower	upper	lower	lower	
5%	13	18	40	27	16	2	2	13	87	24
10%	27	31	60	27	27	67	47	18	98	44
15%	69	56	80	40	42	67	53	40	98	60
20%	80	67	91	49	73	82	84	44	98	74
25%	87	82	93	78	80	82	89	84	98	86

[2]We did not use the upper limit of the SpO_2 in the evaluation because it is always set to 100%.

Table 6.14: Percentage of correct limits per signal for the factory default limits of the HP Merlin. The module we evaluated did not allow for limits on the diastolic blood pressure. (49 patients).

	BPsys		HR		$EtCO_2$		SpO_2	Average
% from limit	upper	lower	upper	lower	upper	lower	lower	
5%	18	16	4	9	0	7	38	13
10%	51	24	33	9	9	13	98	34
15%	56	58	36	24	11	53	98	48
20%	67	62	53	24	47	58	98	58
25%	78	82	69	80	53	60	98	74

Figure 6.3 shows the difference between a conventional monitoring system (HP Merlin), a system that uses only one patient group, and our system that uses patient groups to assign specific alarm limits.

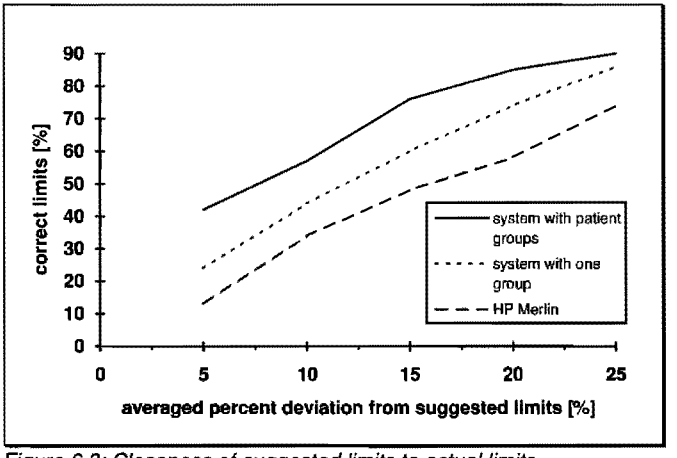


Figure 6.3: Closeness of suggested limits to actual limits

Our system scores higher than the one group system, which in turn scores higher than a conventional system. The signal that scores low is the lower limit of the diastolic blood pressure. A reason for this may be that the lower limit for diastolic blood pressure is not set as carefully as the other blood pressure limits. The lower diastolic blood pressure limit does not score well in the simple system either, which may indicate that this limit is partially based on information available intraoperatively (the evaluation of the simple system consists of comparing the limits set by experts with the limits used intra-operatively).

An additional test was done comparing the patients from the study in paragraph 2.3 with limits suggested for these patients by our system. Of the 50 patients we studied, 30 had a complete electronic preoperative evaluation, and could be used as inputs to our system (Table 4.9). We tested if the COR upper limit of the maintenance phase was lower than the limit suggested by our system, and if the COR lower limit was higher than the suggested lower limit. This is expected from the definition of the COR limits (see also Figure 5.1, level b). Table 6.15 shows the number of limits that were within and that were not within the range between the suggested upper and the lower limits.

Table 6.15: Number of COR limits within or outside the limits suggested by our system.

	BPsys		**BPdia**		**HR**		**$EtCO_2$**		**SpO_2**
	upper	lower	upper	lower	upper	lower	upper	lower	lower
within	29	30	27	25	24	22	30	21	29
not within	1	0	3	5	6	8	0	9	1
%	96.7	100.0	90.0	83.3	80.0	73.3	100.0	70.0	96.7

Of the 33 limits that were not within the range of suggested limits, 15 were less than 5% beyond the suggested limits. The other 17 (6% of the total) can be contributed to variability in the patient population and anesthesiologist.
These data confirm that in general our system suggests clinically useful limits.

6.3 Summary

In this chapter we have evaluated the system for automatic alarm limit setting. We have validated the assumptions we made and analyzed the performance. The analysis of the limits assigned to the different groups prompted us to merge two pairs of limit sets, resulting in a total of six limit sets. The next chapter will put the results of the analysis of our system in perspective with the way alarm limits are set in the operating room, it will draw general conclusions and propose recommendations.

Discussion and Conclusions

7.1 Discussion

System design

We described in this dissertation a system that is able to pre-select intraoperative alarm limits for blood pressure, heart rate, end tidal CO_2 and blood oxygen saturation. Our system assigns the patient to a pre-defined group and then presents the average of the limits as they were assigned by experts to each patient in that group. The patient groups were formed by the K-means clustering method, based on a summary of the preoperative evaluation.

The preoperative summary was formed with the help of experts to contain information needed to select alarm limits for specific patients. We showed that we initially could identify eight patient groups based on this information, and were able to find clinically useful descriptions for these groups.

Expert knowledge that was used for these groups consisted of the summary of the preoperative data, the selection of problems that have an impact on setting alarm limits, and the scaling of the data items.

Patient groups were created with clustering methods, and were evaluated to determine if they were different based on the limits assigned to them. There was no difference in how limits were set for groups 1 (average age, ASA I or II patient) and 3 (young healthy ASA I patient). Even though numerical differences were found by the clustering algorithm using the preoperative data, the experts did not use different limits for these patient groups. For the same reason the limits of groups 7 (older sick ASA II or III patient with problems with HR but not BP) and 8 (average age, sick ASA II or III patient with multiple problems) were combined. The fact that multiple groups mapped to the same set of alarm limits is not a problem for our system, because the limits are still assigned correctly. It is an indication that we defined enough patient groups.

The system was designed to allow for extensions and modifications. Because the limits assigned to a patient are updated if a clinician decides that the limits were not appropriate, changes in how alarm limits are set are accommodated. This also allows

for adapting to differences in the setting of alarm limits at different geographical locations, where clinical practice may be different from our test site.
Besides making changes to the limits assigned to each patient, the design methodology can be used to adapt the system to institutions where the patient population is different from our test site and includes many patients with specific problems, or many young patients. Following the same design philosophy as done in this dissertation, a system specific for one institution can be designed.

In paragraph 2.3 we documented the setting and use of alarm limits for 50 patients in the operating room. We defined four anesthetic phases (intubation, induction, maintenance, and emergence). From that study we concluded that limits differ between the maintenance phase and the other three phases, and are set similarly for intubation, induction and emergence phases. The limits that were used by anesthesiologists to design our system were specific for the maintenance phase. We showed that our technique (grouping patients based on the preoperative evaluation, and assigning limits per group) worked for the maintenance phase. The same technique can be used for the other phases, but some additional knowledge (e.g. intubation failed, trying to wake up the patient, or type of induction used) may be required. For example, during the beginning of the intubation phase there will be no end tidal CO_2 measurement because the endotracheal tube has not yet been inserted into the patient's trachea. An alarm system needs to be aware of this knowledge.

Human factors also play a role, as we found with the study of the 50 patients: Anesthesiologists did not set the upper limit of the end tidal CO_2 to a higher level during emergence, although it is expected that the end tidal CO_2 will rise because ventilation is turned down to stimulate the patient's respiratory center. We discovered by talking to the anesthesiologists that they found the alarm of high $EtCO_2$ reassuring during the emergence phase because they were expecting it.
Proper setting of alarm limits during intubation, induction, and emergence is very dependent on the context of what is happening to the patient (incision, inserting endotracheal tube, etc.). In addition, artifacts are a bigger problem during these phases than during the maintenance phase, because equipment is moved around, drapes are put in place, etc.

The road to 'better' alarm systems

In this thesis we have presented a system that can set alarm limits similar to the way anesthesiologists do. When a system is designed not only should we look at its performance according to the specification or goals, but also the overall usefulness has to be evaluated. The system for the automatic selection of alarm limits was designed to eliminate the need of clinicians to go through the time consuming task of setting alarm limits, to provide other automatic systems (like the Intelligent Alarm System (van der Aa 1990)) with an acceptable range of a patient's physiologic measurements, and to provide a reference set of patient specific alarm limits.

Once we are able to set alarm limits as anesthesiologists do, the question arises: "can we do better than that". To answer this question we first have to define what we mean by the word "better". One of the main goals of monitoring patients intra-operatively is to assure a safe condition of the patient and that neither the anesthetic nor the surgery cause any adverse effects. This is achieved by avoiding intraoperative incidents that have a negative effect on the outcome of anesthesia.

Even though monitoring standards have been developed and are currently used, few scientific studies have shown evidence that this monitoring practice has a favorable effect on outcome (Gravenstein 1986; Eichhorn 1989). In a more recent study Moller et al. studied 20,802 patients to determine if pulse oximetry has a favorable effect on outcome (Moller et al. 1993a, 1993b). Even though this study showed differences between the control (no pulse oximetry) group and the test (with pulse oximetry) group, no statistical differences could be shown in the outcome. This result shows that outcome studies need an extremely large number of patients to show a statistical difference in outcome (Eichhorn 1993). There are other factors that need consideration in the evaluation whether a pulse oximetry monitor should be used or not: anesthesiologists feel more secure when using pulse oximetry, and an impression (although not statistically significant) that pulse oximetry prevents problems that could cause an adverse effect on outcome.

These studies focused on the use of monitors, and did not address which values were acceptable for specific patients when these monitors were used.

Measurement of outcome related to anesthetic management is not easy. First, postoperative morbidity and mortality is far more (by an order of magnitude) likely to be affected by the patient's disease and the surgical procedure than by the anesthetic (Velanovich, 1991)). Secondly, no standard measure for outcome exists. Some attempts to standardize outcome measurements have been made. Cooper et al.

defined recovery room impact event (RRIE) as an unanticipated, undesirable, possibly anesthesia-related effect that required intervention, was pertinent to recovery room care, and did or could cause mortality or at least moderate morbidity (Cooper et al. 1987). Other outcome measures include unanticipated intensive care unit admission (UIA) associated with anesthesia (Cullen et al. 1992). Models for quality assurance have also been presented by Vitez (1990) and Edsall (1991) in an attempt to standardize outcome measurements. Even when a standard measure of outcome can be defined, it is still difficult to show statistical differences in outcome caused by differences in monitoring practice (Eichhorn 1993).
Because of the difficulty of outcome measurements, it is difficult to optimize an alarm system for optimal outcome.

Another criterion for better alarm limits could be an improved ratio of correct alarms versus false alarms. Alarms are frequently disabled in the operating room. Mostly this is caused by frustration of anesthesiologists with the current alarm system. False alarms are frequently indicated as the main problem. Anesthesiologists are not alone in their frustration with alarm systems. In other fields where alarms are used this same frustration exists, and the same tendency of disabling alarms is reported. Examples of these other fields include alarms in the locomotive cab of a train , where after investigations of some train crashes it turned out that some visual and auditory alarms signals were taped over (United States Congress 1987). Another example is the alarm system in the cockpit of an aircraft, where pilots wait with their finger over the 'alarm silence' button, expecting an alarm, and silencing it the moment it occurs without thinking why it occurred (Sorkin 1988). The same behavior has been observed in nuclear power plant control rooms where alarms were silenced and acknowledged without further concern or surveillance of the plant status (Sorkin 1988). In these engineering examples the 'system' is fairly well understood. Tests can be done to examine which values of measured parameters are extremes, and alarm limits can be assigned based on those tests (for example the maximum pressure limit of a boiler can be theoretically calculated, and tested in real life). There are not many differences between two engineering systems from the same series that were manufactured the same way. This is not the case with the patient 'system'. It is unethical to examine what the maximum blood pressure is that a human can sustain without injury.

Sorkin gives two reasons why alarms are being turned off: 1) The alarm signal can be very aversive and can interfere with important operator duties (e.g. interfere with tower-

cockpit communications, disturb the surgeon during an operation), 2) The perceived false alarm rate is excessively high (Sorkin 1988).

A solution to the first problem would be integration of alarms from different sources to maximize the content of presented information. When alarms from different monitors are brought together, and with a priority system, based on expert system technology, the alarms can be made less intrusive. By combining the signals, agreement can be made on the methods of conveying alarm information, and the alarm system can be managed from a central point.

Solutions to the second problem have been suggested by van der Aa (1990) and others. Alarms can be separated in different groups: 1) alarms that are helpful and have an effect on intraoperative anesthetic management, 2) alarms triggered by artifact, generated by the monitor (electrical interference for example) or generated by the patient/surgeon (e.g. patient was moved, surgeon leaning on chest), and 3) alarms that are expected (e.g. flushing catheters, high end-tidal CO_2 during the emergence phase of anesthesia, etc.). A truly helpful alarm system should supply alarms of the first type and perhaps a selection of the third type to the anesthesiologist. Artifacts should be eliminated as early as possible in the monitoring system. Artifact rejection should start at the sensor, and continue during the initial signal processing and during the determination of the measured parameter(s). After the values of the measured

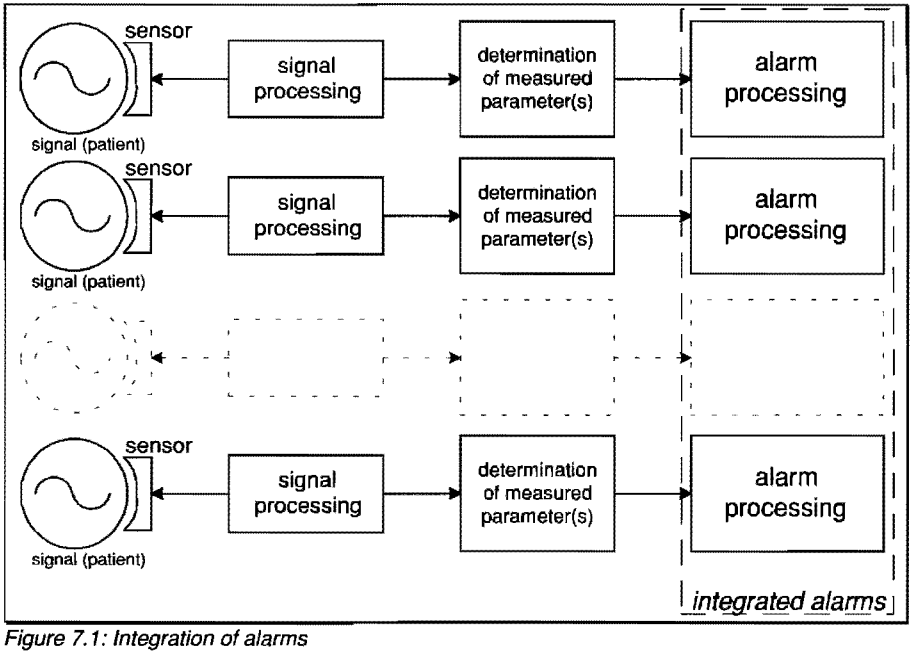


Figure 7.1: Integration of alarms

parameters have been determined by the front-end monitor, they are presented to the alarm processing unit to decide if an alarm condition is present (see Figure 7.1).

The decision to alarm or not to alarm is made based on the value of the parameter(s) measured by the front-end monitors, and on limits set by clinicians in a conventional (stand-alone) monitor. An integrated monitoring and alarm system also uses information from other front-end monitors to make that decision. In addition to this information, knowledge based (integrated) alarm systems decide to alarm based on knowledge specific for the anesthesia domain, and/or based on information about the current case, including context information like which drugs were administered, ventilator settings, etc.
Other approaches include the modeling patients, events, and alarm systems as proposed by Beneken and Gravenstein (1987). The authors present an approach that determines alarm conditions based on a model of an ideal patient, knowledge about events, and fault models related to these events. Patient models are still being designed, and are more accurate than seven years ago (van Meurs et al. 1993), but these patient models represent a hypothetical possible patient, and not a specific patient. If we could model a specific patient (the patient of the case we are monitoring), it then would be possible to look inside the systems of that patient, and determine the state of the (internal) patient systems (cardiovascular system, respiratory system, etc.). It would then also be possible to make predictions of the future, because (assuming no changes of the current settings), the future can simply be calculated. Some of these predictions have been made for concentration levels of anesthetics in patients to make a prediction of when they would wake up, or when the anesthetic would wear off (Gibby et al. 1991b). Unfortunately a complete, accurate model of specific patients does not exist, and a major amount of research would be necessary to create it.

7.2 Conclusions

Several conclusions can be drawn from the research presented in this thesis:

A summary of the preoperative evaluation can be used to separate patients into groups that require different alarm limits (paragraph 6.2.3).

Anesthesiologists can assign alarm limits for the maintenance phase, based on the preoperative evaluation, and set the same limits when they are presented with a summary of the preoperative evaluation (paragraph 6.2.2).

A system can be designed that assigns alarm limits based on expert opinion and grouping of patients that are similar to limits used intraoperatively (paragraph 6.3).

Variability of physiologic patient data can be documented, but no statement can be made whether the alarm, caused by the value of a physiologic parameter, was valid or not without knowing the context of the case (paragraph 2.3).

The method we presented to connect to different physiologic monitors can be used to create an intraoperative data collection tool. This method has great advantages over buying new integrated monitors, or monitors compatible with the Medical Information Bus (chapter 2).

7.3 Recommendations

Integration of alarm systems and systems with artificial intelligence (expert systems, neural nets, etc.) will become commonplace in the near future. It is likely that these systems perform better in terms of minimizing the false alarm rate. Integrated and knowledge based systems will have a need to know physiologically appropriate limits for specific patients, and this thesis has presented a method to arrive at those limits. An integration of our system with a knowledge-based alarm system can be made provided that the information needed to derive the limits (preoperative information) is available to the monitoring system.

The limit database created for this study can be used as a reference of standard practice. After many cases are inserted into the database, the database will reflect the current practice of how limits are used intraoperatively. These data can then be used in training, case analysis, or outcome evaluations.

With the arrival of new technologies that will connect the computers of many hospitals together, we will find that information on patients in different places can easily be made available. Outcome evaluations with large numbers of patients can be made more

easily because of the availability of data on more patients. The impact of limit setting on anesthesia outcome could be studied, when data from many places are combined.

An expansion to the system can be made to include other phases of the operation: e.g. induction, intubation, and emergence. More knowledge is needed to automatically detect the phase of the anesthetic and important events.

Table A.1: Contents of foods per 100 gram. From: Binas informatieboek vwo-havo voor het onderwijs in de natuurwetenschappen. Wolters-Noordhoff, 1977.

	protein	fat	saccharide	calcium	phosphor	iron	sodium	potassium	vit. A	vit. B	vit. C	water	energy
	g	g	g	mg	mg	mg	mg	mg	mg	mg	mg	g	kJ
apples	0	0	10	10	10	0.2	2	150	0	0.11	10	87	171
bread	7.9	2.5	43	20	140	1.5	0	200	0	0.41	0	40	1145
chicken	20	10	0	10	200	2	100	300	0	0.85	0	73	711
milk	3.3	3.2	4.6	120	90	0.03	50	150	0.05	0.24	1	88	251
lettuce	2	0.2	2	40	30	0.5	2	250	1.5	0.2	10	94	50
oranges	0.5	0	10	40	20	0.3	2	150	0.2	0.14	50	86	176
carrot	1	0.2	6	40	30	0.5	75	300	6	0.17	5	90	125
pork	16	24	0	10	200	2	100	350	0	0.84	0	59	1170
beef	20	13	0	10	200	3	100	350	0	0.51	0	68	824
cauliflower	2	0.3	3	20	30	0.5	15	400	0	0.32	80	93	96
yogurt	3.3	3.2	4	120	90	0	50	150	0.04	0.24	0	88	242
rice	7	0.5	78	10	100	0.4	2	100	0	0.2	0	13	1442

Appendix B: List of Abbreviations

Abbreviation	Description
AI	Artificial Intelligence
ANN	Artificial Neural Network
ASA	American Society of Anesthesiologists
ASAi	ASA physical status index
ASCII	American Standard Code for Information Interchange
BP	Blood Pressure
BPdia	Diastolic Blood Pressure
BPi	Blood Pressure problem index
BPsys	Systolic Blood Pressure
COR	Clinical Operating Range
ECG	Electrocardiogram
EEG	Electroencephalogram
EOT	End Of Transmission
$EtCO_2$	End tidal CO_2 partial pressure
HR	Heart Rate
HRi	Heart Rate problem index
ht	Height
ISO	International Organization for Standardization
MIB	Medical Information Bus
NIBP	Non-invasive blood pressure
OR	Operating Room
OSI	Open Systems Interconnection
PA	Invasive arterial pressure
PCA	Principle Component Analysis
PE	Processing Element
rel.val.	Relative Value
ROS	Review Of Systems
SpO_2	Pulse oximetry
wt	Weight

References

American Heritage Dictionary of the English Language. Morris W [Edt.]. 1981. Houghton Mifflin Co. Boston. ISBN 0-395-20360-0.

American Society of Anesthesiologists. New classification of physical status. *Anesthesiology.* 1963; **24**: 11.

American Society of Anesthesiologists. *Relative value guide 1992.* ASA, Park Ridge, IL, 1992.

American Society of Anesthesiologists. *Standards for basic intra-operative monitoring.* In: ASA 1993 directory of members, Park Ridge IL, 1993.

Anderberg, MR. *Cluster analysis for applications.* 1973. Academic Press Inc., New York. ISBN 0-12-057650-3.

Balasaraswathi K, and El-Etr AA. Cardiovascular preop evaluation. *AORN Journal.* Feb 1976; **23**(2): 209-211.

Beecher HK. The first anesthesia records (Codman and Cushing). *Surg. Gynecol. Obstet.* 1940; **71**: 689-693.

Beneken JEW and Gravenstein JS. *Sophisticated Alarms in Patient Monitoring: A methodology Based on Systems Engineering Concepts.* In: Gravenstein JS, Newbower RA, Ream AK, et al, editors: *The automated anesthesia record and alarm systems.* Boston, 1987, Butterworths.

Beneken JEW and van der Aa JJ. Alarms and their limits in monitoring. *Journal of Clinical Monitoring* 1989; **5**: 205-210

Berry AJ and Katz JD. Hazards of working in the operating room. In: *Clinical Anesthesia,* Barash PG, Cullen BF and Stoelting RK [eds.], Chapter 4. 1989. Lippincot Co. Philadelphia PA. ISBN 0-397-50836-0.

Block FE. Automatic anesthesia record keeping. *Journal of Clinical Monitoring.* 1989; **5**: 284-286.

Charlson ME, MacKenzie CR, Gold JP, Ales KL, Topkins M, and Shires GT. Intraoperative blood pressure. *Annals of surgery.* Nov 1990; **212**(5): 567-580.

Chodoff P and Helrich M. Construction of an Automated System for the Collection and Processing of Preoperative Data. *Anesthesia and Analgesia.* 1969; **48**(5): 870-876.

Chodoff P and Ginaris C,II. A Nurse-Computer Assisted Preoperative Anesthesia Management System. *Computers and Biomedical Research.* 1973; **6**: 371-392.

Clemmer TP and Gardner RM. Medical informatics in the intensive care unit: state of the art 1991. *International Journal of Clinical Monitoring and Computing.* 1992; **8**: 237-250.

Cluitmans PJM. *Neurophysiologic monitoring of anesthetic depth.* 1990. Ph.D. Thesis, Eindhoven University of Technology, The Netherlands. ISBN 90-9003453-6.

Cook RI, McDonald JD, and Nunzita E. Differences between handwritten and automatic blood pressure records. *Anesthesiology.* 1989; **71**: 385-390.

Cooper JB, Cullen DJ, Nemeskal R, Hoaglin DC, Gevirtz CC, Csete M, and Venable C. Effects of Information Feedback and Pulse Oximetry on the Incidence of Anesthesia Complications. *Anesthesiology.* 1987; **67**: 686-694.

Cullen DJ, Nemeskal AR, Cooper JB, Zaslavsky A, and Dwyer MJ. Effects of Pulse Oximetry, Age, and ASA Physical Status on the Frequency of Patients Admitted Unexpectedly to a

Postoperative Intensive Care Unit and the Severity of Their Anesthesia-Related Complications. *Anesthesia and Analgesia.* 1992; **74**: 181-188.

Date CJ. *An introduction to database systems.* Addison-Wesley, Reading MA. 1986.

Davidson MHA. *The evolution of Anaesthesia.* St. Ann's Press, Altrincham GB, 1965.

Durand LG, Blanchard M, Cloutier G, Sabbah, HN, and Stein PD. Comparison of pattern recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position. *IEEE transactions of biomedical engineering.* 1990; **37**(12): 1121-1129

Editorial. An international consensus on monitoring? *British Journal of Anaesthesia.* 1990 Mar; **64**(3): 263-266

Edsall DW. Quality assessment with a Computerized Anesthesia Information Management System (AIMS). *Quality Review Bulletin.* 1991 Jun; **17**(6): 182-193.

Eichhorn JH, Cooper JB, Cullen DJ, Maier WR, Philip JH, and Seeman RG. Standards for Patient Monitoring During Anesthesia at Harvard Medical School. *Journal of the American Medical Association.* 1986 Aug; **256**(8): 1017-1020

Eichhorn JH. Prevention of Intraoperative Anesthesia Accidents and Related Severe Injury through Safety Monitoring. *Anesthesiology.* 1989; **70**: 572-577

Eichhorn JH. Pulse Oximetry as a Standard of Practice in Anesthesia. *Anesthesiology.* 1993; **78**: 423-426.

Electronic Industries Association Engineering Dept. *EIA Standard RS-232-C: Interface between Data Terminal Equipment and Data Communication Equipment employing Serial Binary Data Interchange.* Washington DC, 1981.

Elmasri R and Navathe S. *Fundamentals of database systems.* Benjamin/Cummings Publishing Company Inc, Redwood City, CA. 1989.

Farrel RM, Orr JA, Kuck K, and Westenskow DR. Differential features for a neural network based anesthesia alarm system. *Biomed-Sci-Instrum.* 1992; **28**: 99-104

Feldman JM and Good ML. *Record-keeping and automatic record-keepers.* In: Ehrenwerth J and Eisenkraft JB, editors: *Anesthesia Equipment: principles and applications.* St. Louis, 1993, Mosby.

Figler AA, and Stead SW. The Medical Information Bus. *Biomedical Instrumentation & Technology.* 1990 Mar-Apr; **24**(2): 101-111.

Fleisher LA and Barash PG. Preoperative cardiac evaluation for noncardiac surgery: A functional approach. *Anesthesia and Analgesia.* 1992; **74**: 586-598.

Frazier WT. *Spatial Organization of a Computer-Assisted Anesthetic Record.* In: Gravenstein JS, Newbower RA, Ream AK, et al, editors: *The automated anesthesia record and alarm systems.* Boston, 1987, Butterworths.

Gardner RM, Hawley WL, East TD, Oniki TA, and Young HFW. Real time acquisition: recommendations for the Medical Information Bus (MIB). *International Journal of Clinical Monitoring and Computing.* 1992; **8**: 251-258.

Gibbs RF. The present and future medicolegal importance of record keeping in anesthesia and intensive care: the case for automation. *Journal of Clinical Monitoring.* 1989; **5**: 251-255.

Gibby GL, Castleberry JE, Gravenstein JS, and Jackson KI. Automated Preanesthetic Evaluation System Significantly Improves Completeness in Checking Laboratory Values. *International Journal of Clinical Monitoring and Computing.* 1991a; **8**(2): 136. Abstracted.

Gibby GL, Martinez EA, and Arroyo AA: Two models of uptake and distribution of volatile anesthetic agents to predict expired anesthetic values. *Journal of Clinical Monitoring* 7:118, 1991b. Abstracted

Gibby GL, Jackson KI, Gravenstein JS, and Paulus DA. Development of Problem Categories for Computerized Preanesthetic Evaluation of Outpatients. *Journal of Clinical Monitoring.* 1992; **8**: 156-157. Abstracted.

Gill PE. *Practical Optimization.* Academic Press Inc, New York NY. 1981.

Glantz SA. *Primer of Biostatistics Third Edition.* McGraw-Hill, Inc. New York, NY, 1992 .

Gravenstein JS. Essential monitoring examined through different lenses. *Journal of Clinical Monitoring.* 1986; **2**: 22-29.

Gravenstein JS. The uses of the anesthesia record. *Journal of Clinical Monitoring.* 1989; **5**: 256-265.

Gravenstein JS, de Vries A, and Beneken JEW. Sampling intervals for clinical monitoring of variables during anesthesia. *Journal of Clinical Monitoring.* 1989; **5**: 17-21.

Green NM. Anesthesia and the development of surgery (1846-1896). *Anesthesia and analgesia.* 1979 Jan-Feb; **58**(1): 5-12.

Greenhow SG, Linkens DA, and Asbury AJ. Development of an expert system for anaesthetic control. *Computer Methods and Programs in Biomedicine.* 1992; **37**: 215-229.

Hartigan J. *Clustering algorithms.* New York, 1975, Wiley.

Hewlett-Packard. *HP Component Monitoring System User's Reference Guide.* Edition six, 1992. Part number M1046-9001C. Hewlett-Packard Böblingen, Germany.

Holman JG and Cookson MJ. Expert systems for medical applications. *Journal of medical engineering & technology.* 1987; **11**(4): 151-159.

Hug CC, Jr. Monitoring. In: *Anesthesia.* Volume I, Miller-RD [ed.], Chapter 6. Curchill Livingstone Inc. 1981; ISBN 0-443-08082-8.

International Organization for Standardization. Information Centre. *Information transfer: handbook on international standards governing information transfer : texts of ISO standards.* Geneva, 1977, International Organization for Standardization.

Jackson CJ and Scott RJ. A New Comprehensive Anaesthesic Record. *Anaesthesia and Intensive Care.* 1989; **17**(4): 475-481.

Jackson KI. *Patient outcome prediction for anesthesia.* Thesis (M.Eng) University of Florida. 1992.

Jain AK and Dubes RC. *Algorithms for clustering data.* 1988. Prentice Hall, Inc., Englewood Cliffs, NJ. ISBN 0-13-022278-X.

Julian RM. *Understanding Anesthesia.* 1984. Addison-Wesley Publishing Co. Menlo Park CA. ISBN 0-201-11601-4

Kahn MG, Fagan LM, and Tu S. Extensions to the Time-Oriented Database Model to Support Temporal Reasoning in Medical Expert Systems. *Methods of Information in Medicine.* 1991; **30**: 4-14.

Kernighan BW and Ritchie DM. *The C programming language. Second edition.* Englewood Cliffs, N.J., 1988, Prentice Hall .

Kestin IG, Miller BR, and Lockhart CH. Auditory Alarms during Anesthesia Monitoring. *Anesthesiology.* 1988; **69**:106-109.

King PH and Smith BE. Computerized monitoring at Vanderbilt University status and future directions. *International Journal of Clinical Monitoring and Computing.* 1991; **8**: 117-120.

Kristeller JL and Rodin J. Identifying eating patterns in male and female undergraduates using cluster analysis. *Addictive Behaviors.* 1989; **14**: 631-642.

Lemeshow S, Hosmer DW, Klar J, and Lwanga SK. *Adequacy of Sample Size in Health Studies.* John Wiley & Sons Ltd., Chichester, England. 1990.

Linnarsson D and Hallén B. *The information Content of an Automated Anesthesia Record.* In: Gravenstein JS, Newbower RA, Ream AK, et al, editors: *The automated anesthesia record and alarm systems.* Boston, 1987, Butterworths.

Mäkivirta A, Koski E, Kari A, and Sukuvaara T. The median filter as a preprocessor for a patient monitor limit alarm system in intensive care. *Computer methods and programs in biomedicine.* 1991 Feb-Mar; **34**(2-3): 139-144

Meijler AP. *Automation in anesthesia a relief? a systematic approach to computers in patient monitoring.* New York, 1987, Spinger-Verlag.

Miller PL. Critiquing Anesthetic Management: The "ATTENDING" Computer System. *Anesthesiology.* 1983; **58**: 362-369.

Miller PL. Medical plan-analysis by computer. *Computer Programs in Biomedicine.* 1984; **18**: 15-20.

Moller JT, Pedersen T, Rasmussen LS, Jensen PF, Pedersen BD, Ravlo O, Rasmussen NH, Espersen K, Johannessen NW, Cooper JB, Gravenstein JS, Chraemmer-Jørgensen B, Wilberg-Jørgensen F, Djernes M, Heslet L, and Johansen SH. Randomized Evaluation of Pulse Oximetry in 20,802 Patients: I. Design, Demography, Pulse Oximetry Failure Rate, and Overall Complication Rate. *Anesthesiology.* 1993a; **78**: 436-444.

Moller JT, Johannessen NW, Espersen K, Ravlo O, Pedersen BD, Jensen PF, Rasmussen NH, Rasmussen LS, Pedersen T, Cooper JB, Gravenstein JS, Chraemmer-Jørgensen B, Djernes M, Wilberg-Jørgensen F, Heslet L, and Johansen SH. Randomized Evaluation of Pulse Oximetry in 20,802 Patients: II. Perioperative Events and Postoperative Complications. *Anesthesiology.* 1993b; **78**: 445-453.

Murtagh F and Heck A. *Multivariate data analysis.* Dordrecht, Holland, 1987, Kluwer.

Nolan-Avila LS, Paganilli BE, and Norden-Paul RE. The Medical Information Bus. An Automated Method for Capturing Patient Data at the Bedside. *Computers in Nursing.* 1988; **6**(3): 115-121.

O'Carroll TM. Survey of alarms in an intensive therapy unit. *Anaesthesia.* 1986 Jul; **41**(7): 742-744

Orr JA and Westenskow DR. Evaluation of a breathing circuit alarm system based on neural networks. *Anesthesiology.* 1990; **73**(3A): A445.

Peters RM. Interactive microcomputer for acquisition of patient information. *Journal of Clinical Monitoring.* 1989; **5**: 266-269.

Phillips GD, Gordon PJ, and Cousins MJ. Computer-based Monitoring and Data Analysis in Anaesthesia and Intensive Care. *Anaesthesia and intensive care.* 1982 Aug; **10**(3): 223-228

Poore GV. History of Anaesthetics. *The Lancet.* October 12, 1872. Reprinted in: *Medical Times.* Nov 16, 1872; **1**:108.

Raftery H. Anaesthesia coming of age. *Irish Journal of Medical Science.* 1968 Nov; **7**(11): 523-530.

Ream KR. Automating the recording and improving the presentation of the anesthesia record. *Journal of Clinical Monitoring.* 1989; **5**: 270-283.

Reichgelt H. *Knowledge representation: an AI perspective.* Norwood NJ, 1991, Ablex Publishing Co.

Rennels GD and Miller PL. Artificial Intelligence research in Anesthesia and Intensive care. *Journal of Clinical Monitoring.* 1988; **4**: 274-289.

Schaaf C, Block FE. Evaluation of Alarm Sounds in the operating Room. *Journal of Clinical Monitoring.* 1989; **5**: 300.

Schecke T, Langen M, Popp HJ, Rau G, Häsmacher H, and Kalff G. Knowledge-based decision support for patient monitoring in cardioanesthesia. *International Journal of Clinical Monitoring and Computing.* 1992; **9**: 1-11.

Schneider AJL, and Redford JE. Historical Pulse Tracings Made during Anesthesia. *Anesthesiology.* 1979; **51**: 242-244.

Schreiber PJ and Schreiber J. Structured Alarm Systems for the Operating Room. *Journal of Clinical Monitoring.* 1989; **5**:201-204

Shuster JJ. *CRC Handbook of Sample Size Guidelines for Clinical Trials.* Boca Raton, FL, 1990, CRC Press Inc.

Simpson PK. *Artificial neural systems: foundations, paradigms, applications, and implementations.* 1990. Pergamon Press, Inc. Elmsford NY.

Sorkin RD. Why are people turning off our alarms? *Journal of the Acoustical Society of America.* 1988; **84**(3): 1107-1108

Taber's Cyclopedic Medical Dictionary, edition 16. Thomas CL (edt). 1989. FA Davis Company, Philadelphia.

Thomsen CE, Rosenfalck A, and Nørregaard Christensen K. Assessment of anaesthetic depth by clustering analysis and autoregressive modelling of electroencephalograms. *Computer Methods and Programs in Biomedicine.* 1991; **34**: 125-138.

Tinker JH, Dull DL, Caplan RA, Ward RJ, and Cheney FW. Role of Monitoring Devices in Prevention of Anesthetic Mishaps: A Closed Claims Analysis. *Anesthesiology.* 1989; **71**: 541-546.

Tou JT and Gonzalez RC. *Pattern Recognition Principles.* Addison-Wesley, Reading, MA. 1974.

United States. Congress. House. Committee on Government Operations. Special Task Force. Amtrak/Conrail crash : hearing before a Special Task Force of the Committee on Government Operations, House of Representatives, One hundredth Congress, first session, January 22, 1987. Washington : U.S. G.P.O. : For sale by the Supt. of Docs., Congressional Sales Office, 1987. pp. 81-82.

van der Aa JJ. *Intelligent alarms in anesthesia: a real time expert system application.* Ph.D. Thesis, Eindhoven University of Technology, The Netherlands. ISBN 90-9003303-3

van Meurs WL, Beneken JEW, Good ML, Lampotang S, Carovano RG, and Gravenstein JS. Physiologic model for the Gainesville Anesthesia Simulator. In: *Proceedings of the second European conference on engineering and medicine.* Elsevier, Amsterdam, The Netherlands. 1993.

van Oostrom JH. *Intelligent Alarms in Anesthesia: An implementation.* Eindhoven: Eindhoven University of Technology, Faculty of Electrical Engineering. EUT report 89-E-229. 1989. ISBN 90-6144-229-X

van Oostrom JH, van der Aa JJ, Gibby GL. An intraoperative data collection tool for Windows 3.0™. *Anesthesiology.* 1992; **77**(3A), A562. Abstracted

Velanovich V. The value of routine preoperative laboratory testing in predicting postoperative complications: A multivariate analysis. *Surgery.* Mar. 1991; **109**: 236-243.

Veselis RA, Reinsel R, Sommer S, and Carlon G. Use of neural network analysis to classify electroencephalographic patterns against the depth of Midazolam sedation in intensive care unit patients. *Journal of Clinical Monitoring.* 1991; **7**: 259-267.

Vitez TS. A Model for Quality Assurance in Anesthesiology. *Journal of Clinical Anesthesia.* 1990; **2**: 280-287.

Westenskow DR, Orr JA, Simon FH, Bender HJ, and Frankenberger H. Intelligent Alarms Reduce Anesthesiologist's Response Time to Critical Faults. *Anesthesiology.* 1992; **77**: 1074-1079.

Winston, PW. *Artificial Intelligence, second edition.* Reading MA, 1984, Addison-Wesley.

Witcher C. *Advantages of automated record keeping.* In: Gravenstein JS and Holzer JF, editors: *Safety and cost containement in anesthesia.* Boston, 1988, Butterworths.

Witcher C, Ream AK, Parsons D, Rubsamen D, Scott J, Chapeau M, Sterman W, and Siegel L. Anesthetic mishaps and the cost of monitoring: a proposed standard for monitoring equipment. *Journal of Clinical Monitoring.* 1988; **4**: 5-15.

Summary

Objective. To set intraoperative alarm limits automatically, based on information available about the patient and the procedure.

Introduction. Patient monitoring during anesthesia usually includes measurements of heart rate, systolic and diastolic blood pressures, end tidal CO_2 concentration, and arterial blood oxygen saturation. Alarm limits (upper and lower) are set to notify anesthesiologists of variables that transgress these limits. Alarm limits are individualized per patient based on the information obtained from the patient, primarily the preoperative evaluation (preop), and the anesthesiologist's knowledge and experience.

Because many monitors are used intraoperatively, it is a major task to program all the monitors with the desired limits. Many times no limits are set or they are left at the factory defaults or the intraoperative values of the previous case.

Most hospitals are moving to integrated networked systems that connect clinics, labs, and operating rooms, which makes it possible to use electronic patient information in a monitoring system. Intelligent monitoring systems and other systems such as Quality Assurance (QA) systems and automatic record keepers also need to know what acceptable limits of the patient variables are.

A system that takes information from the preoperative evaluation combined with anesthesiologist's knowledge, and suggests intra-operative alarm limits to the anesthesiologist was designed and implemented.

Methods. A systematic appoach was taken to implement a system based on the OSI reference model to record data from different intraoperative monitors. We documented how anesthesiologists use alarm limits, and how patients transgress these limits. Information on how to set alarm limits comes from the combination of two sources: the preoperative evaluation and the knowledge and experience of the anesthesiologist. The important data items from the preop that determine alarm limits are identified by experts. Because patients with the same or similar data require the same limits, patient groups are formed, based on the preoperative data, to facilitate the selection of appropriate limits. These groups are formed by using K-means clustering after parameter normalization, and are evaluated by experienced anesthesiologists. For each patient group a set of alarm limits is derived by asking an experienced anesthesiologist to indicate which limits would be appropriate in a case presented to him/her (data come from actual cases in the past). The limits assigned to each patient

group are the averages of the limits indicated by the experts for the patients belonging to that group. The patient groups are evaluated by determining if the limit sets for the groups were different. The limits assigned by the system are compared to alarm limits indicated by residents and faculty anesthesiologists intraoperatively.

Results. The information of the preoperative evaluation deemed important for selecting intraoperative alarm limits is: age, weight, systolic preoperative blood pressure, preoperative heart rate, ASA physical status, relative value of the procedure, problem index for blood pressure (comprised of past problems with stroke, shock, or aneurysm), and problem index for heart rate (comprised of past problems with coronary artery disease, hyperthyroidism, or heart block).

K-means clustering of these parameters of 5916 patients resulted initially in eight patient groups. For 106 patients (divided over the eight groups), expert anesthesiologists selected upper and lower alarm limits for systolic and diastolic blood pressure, heart rate, SpO_2 and end tidal CO_2 for the maintenance phase of anesthesia.

After evaluation of the differences in alarm limit settings for the different groups, six patient groups of limits resulted (two sets of two groups were merged because they had a similar set of limits).

The limits assigned by the system, based on expert opinions, were compared to the limits selected intraoperatively by faculty and resident anesthesiologists. 76% of the limits set by our system were within 15% of the limits selected intraoperatively.

In a seperate clinical study we defined the Clinical Operating Range (COR) of a patient's variables as the range that is clinically acceptable to the anesthesiologist. We showed that the 88 % of the COR limits were within the limits suggested by our system.

Discussion. The limits selected by our system are close to the ones selected intraoperatively by anesthesiologists, but no statement can be made on how good these limits are in terms of maximizing favorable anesthetic outcome. The limits database used in this study can be used as an indication of the standard of anesthetic monitoring practice.

Samenvatting

Doelstelling. Het automatisch instellen van alarmgrenzen voor patiëntbewaking tijdens anesthesie, gebaseerd op beschikbare informatie over de patiënt en de operatie.
Probleembeschrijving. Patiëntbewaking tijdens anesthesie bestaat normaliter uit meting van hartslag, systolische- en diastolische bloeddruk, eind expiratoire CO_2 concentratie, en zuurstofsaturatie van het arteriële bloed. Alarmgrenzen (boven en onder) worden ingesteld om de anesthesist te waarschuwen voor variabelen die ontoelaatbare waarden aannemen. Deze alarmgrenzen worden per patiënt bepaald, gebaseerd op over die patiënt beschikbare informatie, voornamelijk de preoperatieve evaluatie en de kennis en ervaring van de anesthesist.
Het is veel werk om alle alarmgrenzen in te stellen, gezien het aantal gebruikte monitors in de operatiekamer. Vaak worden dan ook geen alarmgrenzen ingesteld, of worden de standaard fabrieksinstellingen of de grenzen van de vorige operatie gebruikt.
De meeste ziekenhuizen hebben tegenwoordig geïntegreerde netwerksystemen die laboratoria, klinieken en operatiekamers met elkaar verbinden. Het wordt dan mogelijk om elektronische patiëntgegevens te gebruiken in patiëntbewakingsapperatuur. Intelligente patiëntbewakingsapperatuur en andere systemen zoals Kwaliteits Kontrole systemen en automatische "recordkeepers" moeten ook weten wat acceptabele alarmgrenzen voor de patiëntvariabelen zijn.
Wij hebben een systeem ontworpen dat informatie van de preoperatieve evaluatie verwerkt en, gecombineerd met kennis over anesthesie van de anesthesist, een aanbeveling voor alarmgrenzen geeft.
Methoden. Met een systematische aanpak hebben wij een systeem ontworpen, gebaseerd op het OSI referentiemodel, dat data kan opnemen van verschillende intraoperatieve patiëntbewakingsapperatuur. Wij hebben gedocumenteerd hoe de anesthesist alarmgrenzen gebruikt, en hoe patiëntvariabelen deze grenzen overschrijden. Informatie over hoe alarmgrenzen worden ingesteld komt uit twee bronnen: de preoperatieve evaluatie en de kennis en ervaring van de anesthesist. De belangrijkste data van de preoperatieve evaluatie zijn aangegeven door experts. Omdat patiënten met gelijke of gelijkwaardige data dezelfde grenzen behoeven, zijn patiëntengroepen gevormd, op basis van de preoperatieve data, om de selectie van geschikte grenzen mogelijk te maken.
De groepen zijn gevormd door de "K-means clustering" techniek toe te passen op genormaliseerde preoperatieve patiëntgegevens, en zijn vervolgens geëvalueerd door ervaren anesthesisten. Een reeks alarmgrenzen is afgeleid voor elke patiëntengroep door ervaren anesthesisten te vragen welke grenzen geschikt zijn voor een bepaalde

patiënt. De alarmgrenzen voor elke groep worden gevormd door het gemiddelde te nemen van de door de experts aangegeven grenzen voor die groep. De patiëntengroepen zijn geëvalueerd door te bepalen of de alarmgrenzen tussen de groepen verschillend waren. De grenzen aanbevolen door ons systeem zijn vergeleken met de grenzen die anesthesisten aangaven tijdens de operatie.

Resultaten. De informatie van de preoperatieve evaluatie die belangrijk bleek voor het selecteren van alarmgrenzen was: leeftijd, gewicht, preoperatieve systolische bloeddruk, preoperatieve hartslag, ASA status, relatieve waarde van de operatie, probleemindex voor bloeddruk (gebaseerd op de aanwezigheid van beroerte, shock of aneurysma) en een probleemindex voor hartslag (gebaseerd op de aanwezigheid van vernauwing van de kransslagaderen, schildklieraandoening of hartblok).

K-means groepering van deze parameters voor 5916 patiënten heeft in eerste instantie acht groepen opgeleverd. Ervaren anesthesisten hebben alarmgrenzen (boven en onder) aangegeven voor systolische en diastolische bloeddruk, hartslag, zuurstofsaturatie van het arteriële bloed, en eindexpiratoire CO_2 concentratie voor 106 patiënten, verdeeld over die acht groepen, tijdens de maintenance fase van de anesthesie.

Na evaluatie van de verschillen in de alarmgrenzen voor de verschillende groepen, bleven zes groepen van grenzen over (twee paar groepen zijn gecombineerd omdat zij dezelfde grenzen behoeven).

Wij hebben de grenzen, aangegeven door ons systeem vergeleken met grenzen geselecteerd door anesthesisten in de operatiekamer. 76% van de door ons systeem aangegeven grenzen waren binnen 15% van de grenzen in de operatiekamer.

In een aparte study hebben we de 'Clinical Operating Range' (COR) van de patiëntvariabelen gedefinieerd als het bereik dat klinisch geaccepteerd wordt door de anesthesist. Wij hebben aangetoond dat 88% van de COR grenzen binnen de grenzen lagen die ons systeem aangaf.

Discussie. De grenzen aangegeven door ons systeem zijn dichtbij de grenzen geselecteerd in de operatiekamer, maar wij kunnen geen uitspraak doen over hoe goed die grenzen zijn met betrekking tot het resultaat van de operatie. De database van alarmgrenzen, zoals gebruikt in ons onderzoek, kan gebruikt worden als een indicatie voor standaard patiëntbewaking.

Curriculum Vitae

Johannes Hugo Maria van Oostrom was born on November 6, 1963 in Utrecht, the Netherlands. After graduating from pre-university education (Atheneum β) at the Breul in Zeist in 1982, he attended the University of Technology in Eindhoven (TUE), the Netherlands. During the last year of his study he had the opportunity to do his thesis work in the Department of Anesthesiology at the University of Florida in Gainesville, U.S.A., under direct supervision of Dr. J.J. van der Aa, and supervised from the Medical Electrical Engineering group of TUE by Prof. J.E.W. Beneken. He obtained his masters degree in electrical engineering (ingenieur) in December 1988.

In January 1989 he started working towards a Ph.D. degree at the Department of Anesthesiology in Gainesville, U.S.A. During this period, he was involved in several multi-disciplinary research projects in anesthesia and broadened his insight and knowledge about the medical field and anesthesiology in specific. In his spare time he authored several public domain computer utilities that can be retrieved world-wide via the internet computer network.

He is married, and a member of the Institute of Electrical and Electronics Engineers (IEEE) and of its Engineering in Medicine and Biology Society.

STATEMENTS
pertaining to the dissertation of
J.H.M. van Oostrom:
'A System for Automatic Alarm Limit Setting in Anesthesia'.
Eindhoven, December 1 1993

1.

Although anesthesiologists often turn off the alarms on their monitors, they *do* use a mental picture of the patient and the appropriate limits this patient should stay within.
This dissertation

2.

Integration of pertinent information available about a patient, yields the bases to intraoperative problem solving.
This dissertation

3.

It is unfortunate that the Medical Information Bus is still not available. Therefore it is necessary, for the realization of integration of patient information around the patient, to develop a standard for existing equipment.
This dissertation

4.

Mortality and morbidity rates should be used to set priorities for medical and health research and care.

5.

The international Internet network that connects most universities and many companies is a valuable source of information for research, not limited to the technical field.

6.

The confidentiality aspect is usually underestimated or ignored when electronic patient information is made available on a network.

7.

In some countries it is mandated by law to have your car checked each year. This type of preventive medicine should be an integral part of health care and deserves a wider acceptance.

8.

The classification in racial groups by governments, for example when registering for classes at a university, does not help in the realization of Dr. Martin Luther King's dream from 1963.
["I have a dream" speech, August 28, 1963, ML King]

9.

The electronic distribution of some scientific journals in the near future will bring the risk of a different type of reading: let the computer decide which parts are important and skip the rest. This will not enhance the in-depth understanding of the literature.

10.

An anesthesiologist will never be replaced by equipment, although this fear seems to persist.