

CHEMOMETRIC ANALYSIS OF ACID-BASE MEASUREMENTS
A MULTIVARIATE APPROACH

Acknowledgement

The work presented in this thesis was partly supported by a grant for scientific research from the 'Onze Lieve Vrouwe Gasthuis' (Amsterdam, The Netherlands) and the 'St. Elisabeth Ziekenhuis' (Tilburg, The Netherlands).

Financial support by IMRO TRAMARKO bv for the publication of this thesis is gratefully acknowledged.

ISBN: 90-9013351-8

© M. Hekking, 1999

Chemometric Analysis of Acid-Base Measurements. A Multivariate Approach.

Thesis Erasmus University Rotterdam.

No part of this book may be reproduced without permission of the author.

CHEMOMETRIC ANALYSIS OF ACID-BASE MEASUREMENTS
A MULTIVARIATE APPROACH

CHEMOMETRISCHE ANALYSE VAN ZUUR-BASE METINGEN
EEN MULTIVARIATE BENADERING

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam op gezag van de
Rector Magnificus Prof.dr. P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 22 december 1999 om 15:45 uur

door

Marcel Hekking
geboren te 's-Gravenhage

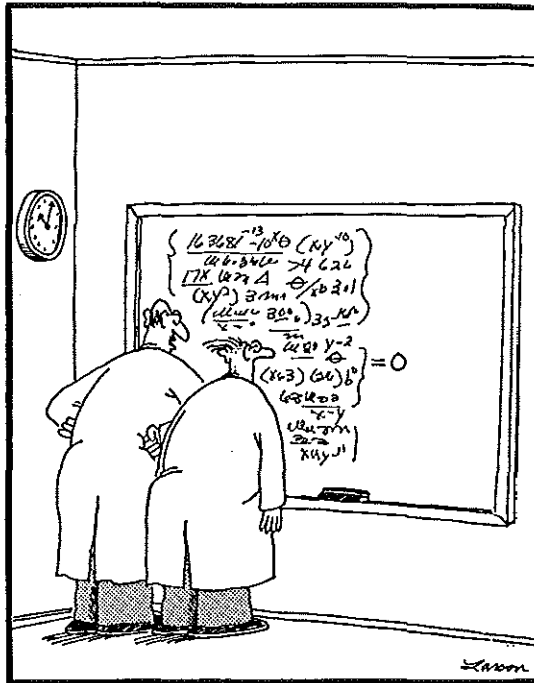
Promotiecommissie

Promotor Prof.dr. E.S. Gelsema

Overige leden: Prof.dr. H.A. Bruining
 Prof.dr.ir. H.L. Vader
 Prof.dr.ir. A. Hasman

Co-promotor Dr. J. Lindemans

THE FAR SIDE by Gary Larson



"No doubt about it, Ellington—we've mathematically expressed the purpose of the universe. Gad, how I love the thrill of scientific discovery!"

THE FAR SIDE © FARWORKS, INC. Used by permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.

Table of contents

1 General introduction	
1.1 The role of clinical chemistry in medicine	12
1.2 Arterial acid-base measurements in the ICU	12
1.3 Basic acid-base physiology	14
1.4 The clinical interpretation of acid-base parameters	16
1.4.1 <i>General nomenclature and terminology</i>	16
1.4.2 <i>The North American view; $a[\text{HCO}_3^-]$ and in vivo CO_2 buffer lines</i>	17
1.4.3 <i>The Scandinavian view; standard bicarbonate and base excess</i>	18
1.5 Objective and scope of this thesis	20
1.6 References	24
2 The application of principal component analysis (PCA) to reduce the dimensionality of trivariate arterial acid-base data distributions	
2.1 Introduction	28
2.2 Materials and methods	29
2.2.1 <i>Patient data</i>	29
2.2.2 <i>Standardisation</i>	30
2.2.3 <i>Principal component analysis</i>	30
2.3 Results	31
2.4 Discussion	35
2.5 Acknowledgements	37
2.6 References	38
3 The graphical representation and classification of arterial acid-base data in a principal component subspace	
3.1 Introduction	40
3.2 Methods	41
3.2.1 <i>Construction of the chart</i>	41
3.2.2 <i>Vectorial classification system for pH, PaCO_2 and base excess values</i>	43
3.3 Results	44
3.3.1 <i>The proposed chart for pH, PaCO_2 and $a[\text{HCO}_3^-]$ values</i>	44
3.3.2 <i>The proposed chart for pH, PaCO_2 and base excess values</i>	45
3.4 Discussion	47
3.5 References	49

4 The construction of patient-based bivariate reference models from trivariate arterial acid-base data distributions	
4.1 Introduction	52
4.2 Materials and methods	55
4.2.1 <i>Patient data</i>	55
4.2.2 <i>The multivariate reference model</i>	55
4.2.3 <i>Finding the Gaussian distributed core in a multivariate patient data set</i>	55
4.3 Results	58
4.4 Discussion	61
4.5 References	67
5 Computational methods and computer program descriptions	
5.1 Introduction	70
5.2 The implementation of numerical routines	70
5.3 The ABTRANS program	71
5.3.1 <i>Importing an acid-base data set (step 1)</i>	72
5.3.2 <i>Mapping variables (step 2)</i>	73
5.3.3 <i>Standardising the acid-base variables (step 3)</i>	73
5.3.4 <i>Performing PCA, constructing a bivariate reference region and defining the graphical outlines of the tri-axial chart (step 4)</i>	73
5.3.5 <i>Saving results into the ABCHART initialisation file (step 5)</i>	74
5.4 The ABCHART program	74
5.4.1 <i>Start-up window</i>	75
5.4.2 <i>Patient data editing window</i>	75
5.4.3 <i>Laboratory data editing window</i>	75
5.4.4 <i>Charts window</i>	76
5.4.5 <i>Automatic data-entry facility</i>	76
5.5 References	79
6 A re-appraisal of the tri-axial chart for monitoring arterial acid-base values — three case reports	
6.1 Abstract	82
6.2 Introduction	82
6.3 Materials and methods	83
6.3.1 <i>Patient cases</i>	83
6.3.2 <i>Construction of the chart</i>	83
6.3.3 <i>Monitoring parameter 1: abnormality of acid-base status</i>	84
6.3.4 <i>Monitoring parameter 2: rate of acid-base change</i>	84
6.4 Patient reports	85
6.4.1 <i>Patient 1 (Figure 6-4, Table 6-1)</i>	85
6.4.2 <i>Patient 2 (Figure 6-5, Table 6-2)</i>	88

6.4.3 Patient 3 (<i>Figure 6-6, Table 6-3</i>)	91
6.5 Discussion	93
6.6 Acknowledgement	94
6.7 References	95
7 Predicting hospital mortality of ICU patients from arterial acid-base measurements	
7.1 Abstract	98
7.2 Introduction	98
7.3 Materials and methods	99
7.4 Results	102
7.5 Discussion	105
7.6 Acknowledgement	109
7.7 References	110
8 Clinical evaluation of a tri-axial chart, multivariate reference model and vector classification method for arterial acid-base data	
8.1 Introduction	114
8.2 Materials and methods	115
8.2.1 <i>Set-up and patient data</i>	115
8.2.2 <i>Study design</i>	116
8.2.3 <i>Data analysis</i>	116
8.3 Results	117
8.3.1 <i>Effect of the introduction of ABCHART</i>	117
8.3.2 <i>Acid-base classification agreements</i>	118
8.3.3 <i>Comparison of reference models</i>	120
8.4 Discussion	120
8.5 Acknowledgement	122
8.6 References	123
9 Summary, general conclusions, and future research	
9.1 Summary	126
9.2 General conclusions and future research	129
10 Samenvatting, algemene conclusies en verder onderzoek	
10.1 Samenvatting	132
10.2 Algemene conclusies en verder onderzoek	135
Dankwoord	139
Curriculum vitae	141

1

General introduction

1.1 The role of clinical chemistry in medicine

Medicine is an art and a science in the service of fellow human beings [1]. On the basis of collected empirical data and information, clinicians select specific diagnoses, rule out other differential diagnoses and eventually make decisions about which and how specific therapeutic interventions are made for the benefit and health of their patients. For a proper interpretation, collected data and information must be compared with other, already existing, data and information to assess the exact value of the clinician's findings. Moreover, a clinician compares observed medical data of a patient with knowledge obtained during his or her training as a clinician and with the experience obtained by working with other patients.

A prerequisite in this paradigm, however, is that collected empirical data on which the diagnoses of a clinician are based must be as objective as possible. Clinical chemistry takes a pivotal role in this in the sense that the chemical characterisation of a patient's body fluid is one of the ways in medicine that can provide such objective data. Since the beginning of this century, clinical chemistry has evolved into a separate and independent discipline in the field of medicine [2-4]. Nowadays, most often a single central clinical chemistry laboratory takes care of the 'analytical needs' of one or more hospitals.

Tasks of the clinical chemist typically include the improvement of existing methods of chemical analysis, the development of new analytical methods and providing the clinician with as much information as possible on the basis of chemical analyses. Especially this last task forms the basis of what has become known as *chemometrics*, a branch of clinical chemistry that uses mathematical and statistical methods to extract a maximum of information from chemical analyses [5, 6].

This thesis presents a multivariate chemometric approach to the problems that are currently associated with the interpretation and evaluation of those laboratory measurements that are used to assess the arterial acid-base status of a patient in an intensive care unit (ICU).

1.2 Arterial acid-base measurements in the ICU

The ICU of today is a highly specialised ward in which expert medical, nursing and technical staff provides medical services to severely ill patients. It is characterised as a high-tech environment in which the real-time monitoring of vital functions plays a central role. The origin of the ICU can be traced back to the second half of the 19th century when special rooms, adjacent to the operating room, were used primarily for the purpose of postoperative care [7]. In the course of time, these recovery rooms evolved into specialised respiratory care units and shock and trauma units, eventually leading to the present day ICU. The modern ICU provides integrated cardiopulmon-

ary support for both medical and surgical patients suffering from severe respiratory and / or cardiac problems as a result of disease or trauma.

The most frequently ordered chemical test in the ICU is the arterial blood gas measurement [8]. Arterial blood gas measurements comprise those measurements of the patient's arterial blood that are used for the evaluation and interpretation of the patient's oxygen and acid-base status. Basic arterial blood gas measurements include: the partial pressure of oxygen (PaO_2), the oxygen saturation of haemoglobin, the pH of arterial blood, the partial pressure of carbon dioxide (PaCO_2) and the bicarbonate-ion concentration ($\text{a}[\text{HCO}_3^-]$). The first two measurements (PaO_2 and oxygen saturation) are used to evaluate the oxygen status, while the other three are used for the interpretation of the arterial acid-base status.

In a strict sense, the term blood gas measurements is incorrect, since only PaO_2 and PaCO_2 are true gas measurements and in modern chemical analysers $\text{a}[\text{HCO}_3^-]$ is not measured but calculated from measured pH and PaCO_2 . Moreover, two other derived acid-base parameters are generally considered part of the set of arterial blood gas measurements. These parameters are the standard bicarbonate-ion concentration (SB) and the base excess (BE). Their derivation and rationale are described in section 1.4.3 in more detail.

Since the second half of this century, the analysis of arterial blood for the purpose of acid-base characterisation has become a vital part of intensive care medicine. The importance of the acid-base characterisation of arterial blood is illustrated by the severe polio epidemic that struck Copenhagen (Denmark) in 1952 [9]. During this epidemic, hospitals in Copenhagen had to cope with a large number of patients needing intensive artificial respiration as a result of paralysis of the respiratory muscles. For a proper setting of the artificial respiration, the complete acid-base status of the patient had to be known. At that time, arterial blood of patients was seldom sampled for the purpose of performing blood gas measurements [10]. Arterial blood gas measurements were mainly performed in physiological laboratories and were not part of daily clinical practice. Techniques of measurement were cumbersome and needed large equipment. The clinical necessity of quickly knowing the patient's arterial acid-base status for the purpose of a proper adjustment of the artificial respiration inspired Poul Astrup to develop his equilibration method [9]. This method allowed a relatively quick determination of the three basic acid-base parameters by only measuring the pH of an arterial blood sample and the pH of the sample equilibrated at two known PaCO_2 gas tensions. The original PaCO_2 is calculated by interpolation [11, 12]. Since then, techniques of analysis developed and arterial acid-base measurements have become routine and indispensable in the daily clinical care of intensive care patients.

1.3 Basic acid-base physiology

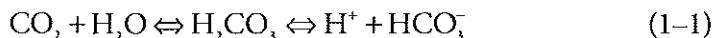
In chemical terms, acids are substances that are capable of donating hydrogen (H^+) ions while bases are substances capable of accepting H^+ ions. The amount of H^+ ions in the arterial blood determines its actual acidity. Acidity is measured as pH, which is, according to the definition of Sørensen, the negative logarithm of the H^+ concentration ($[H^+]$) [9].

The regulation of the amount of H^+ ions in the arterial blood and consequently its pH is one of the most powerful controlling mechanisms in the human body. Under normal physiologic conditions, the pH of arterial blood is kept within well-defined limits. This tight regulation of the H^+ concentration in arterial blood is essential since H^+ ions are highly reactive with negatively charged parts of molecules. Changes in H^+ concentration (intra-cellular as well as extra-cellular) therefore have a profound influence on the molecular configuration and consequently on protein function [13]. Hence, maintaining a constant pH ensures an optimal working condition for enzymes and other proteins. Moreover, large deviations in pH may have effects on the nervous system. If the body becomes too acidic, the nervous system can become so depressed that death can occur. On the other hand, if the body becomes too alkaline, the nervous system can become overexcited, resulting in death from tetanus of the respiratory muscle [14].

Two mechanisms exist to regulate pH of arterial blood: long term physiological buffering and short term chemical buffering. Physiological buffering is the redistribution, production, excretion and/or retention of (non-)volatile acids and bases by means of physiological processes. Chemical buffering is the result of the presence of weak acids and their conjugated bases in the arterial blood. Examples of chemical buffers in arterial blood are: inorganic phosphate, organic phosphate and haemoglobin.

One of the most important chemical buffer systems in the blood, however, is the bicarbonate ion (HCO_3^-)/carbon dioxide (CO_2) buffer system. It is mainly the presence of this buffer system that makes it possible for the human body to cope with the constant load of exogenous acids and bases and the vast amount of both volatile and non-volatile acids that are continuously generated as a result of normal metabolism.

The equation describing the HCO_3^-/CO_2 buffer system in blood is:



The left-hand side of this chemical reaction represents the formation of carbonic acid (H_2CO_3) from CO_2 and H_2O . Therefore, although CO_2 itself is not an acid, an elevation of the CO_2 in the blood increases the acidity of the blood through the formation of H_2CO_3 which immediately dissociates into protons (H^+) and bicarbonate ions (HCO_3^-).

Since the concentration of H_2CO_3 is so low in relation to the concentration of dissolved CO_2 and the concentration of HCO_3^- , the law of mass action for the $\text{HCO}_3^-/\text{CO}_2$ buffer system is:

$$K = \frac{[\text{H}^+] \times [\text{HCO}_3^-]}{[\text{CO}_2] \times [\text{H}_2\text{O}]}, \quad (1-2)$$

where K is a constant.

Because $[\text{H}_2\text{O}]$ is relatively constant in body fluids, it can be omitted from the equation and incorporated into the constant K , further indicated as K' [13]. Rewriting the resulting equation to solve $[\text{H}^+]$ yields the equation that Lawrence Joseph Henderson (1878-1942) first described in 1909 [10]:

$$[\text{H}^+] = K' \times \frac{[\text{CO}_2]}{[\text{HCO}_3^-]} \quad (1-3)$$

The concentration of dissolved CO_2 in blood ($[\text{CO}_2]$) is proportional to the partial pressure of CO_2 (PCO_2) in the gas with which the blood is in equilibrium. Therefore, $[\text{CO}_2]$ can be replaced by the partial pressure of CO_2 in the blood. Partial pressures are either measured in millimetres mercury (mmHg) or kilo-Pascal (kPa) where 1 mmHg = 0.133 kPa. The constant relating $[\text{CO}_2]$ in mmol/l to the PCO_2 is called the solubility constant. The solubility constant for $[\text{CO}_2]$ in plasma is 0.03 mmol per litre per mmHg or 0.225 mmol per litre per kPa.

Moreover, applying the pH concept of Sørensen, in 1917 Karl Albert Hasselbalch (1874-1962) introduced the Henderson-Hasselbalch equation:

$$\text{pH} = \text{pK}' + \log \frac{[\text{HCO}_3^-]}{\alpha \text{PCO}_2}, \quad (1-4)$$

where $\text{pK}' = 6.10$ and α is the solubility constant for $[\text{CO}_2]$ in plasma.

From equation 1-4 it is apparent that pH is the resultant of the ratio $a[\text{HCO}_3^-]/\text{PCO}_2$. Both PCO_2 and $a[\text{HCO}_3^-]$ can effectively be regulated by lungs and kidneys, respectively [15]. This feature in particular makes the $\text{HCO}_3^-/\text{CO}_2$ buffer system so effective in maintaining a constant arterial blood pH. Knowing pH, PCO_2 and $a[\text{HCO}_3^-]$ in the arterial blood of a patient is vital when interpreting the acid-base status of arterial blood. It gives information on both the respiratory and metabolic component of an acid-base disturbance and their joint effect on the acidity of the arterial blood.

Although Sørensen introduced the electrochemical measurement of H^+ ions as early as in 1909, it was not until 1932 that pH glass electrodes were produced commercially and used on a regular basis. Before that time, pH of blood was indirectly obtained from measuring total CO_2 and PCO_2 in the blood with the manometric Van Slyke

apparatus that Donald Dexter van Slyke (1883-1971) introduced in 1924 [9]. Around 1960 the CO_2 electrode was introduced into clinical chemistry. Today, chemical analysers measure pH and PCO_2 and calculate $[\text{HCO}_3^-]$ with the use of the Henderson-Hasselbalch equation (see Equation 1–4).

1.4 The clinical interpretation of acid-base parameters

An impairment in either the respiratory or metabolic function (or both) of the body may result in so-called acid-base disturbances [13]. For a proper treatment of these disturbances it is essential for an ICU clinician to be aware of the exact acid-base status of the arterial blood of an ICU patient. With the analysis of measured and calculated arterial acid-base parameters, the ICU clinician aims to find the underlying cause(s) of one or more acid-base disturbances in order to remove it with specific therapeutic interventions. Moreover, for patients receiving artificial respiration, the acid-base analysis of arterial blood is essential for setting the kind and degree of artificial respiration.

1.4.1 General nomenclature and terminology

Acid-base disorders can be divided into *primary*, *secondary* and *combined* acid-base disturbances. Primary acid-base disturbances are the result of impairment of either the respiratory function or the metabolic function of the body. Impairments of the respiratory function result in primary *respiratory* acid-base disturbances, whereas impairments in metabolic function result in *non-respiratory* or *metabolic* disturbances. Both respiratory and metabolic disturbances can be further divided into disturbances that tend to lower the pH, resulting in *acidemia*, and disturbances that tend to raise the pH, resulting in *alkalemia*. These acid-base disturbances are called *acidoses* and *alkaloses*, respectively. Hence, the terms acidosis and alkalosis refer to underlying pH-deranging physiologic processes, whereas the terms acidemia and alkalemia merely indicate the actual acidity of arterial blood. Multiple single primary acid-base disturbances can be present at the same time, resulting in *combined* acid-base disturbances. Moreover, as a response to primary acid-base disorders, the human body is capable of initiating compensating mechanisms. Primary respiratory disturbances trigger mechanisms in the kidneys that actively regulate the reabsorption of excreted HCO_3^- ions, thereby inducing metabolic compensating effects. Also, primary metabolic dysfunction eventually triggers the breathing centre, resulting in an adjustment of the respiration and consequently the PaCO_2 . These compensating processes result in *secondary* acid-base disturbances. The capability of the body to compensate for primary acid-base disturbances prevents large changes in the pH of arterial blood even though pathological processes may be present.

Respiratory compensations are very rapid and effective within minutes, while metabolic compensations can take up to three days to be fully effective. A metabolic compensation can, however, when in full working order, completely compensate a primary respiratory disturbance, while a respiratory compensation can only partially compensate primary metabolic acid-base disturbances.

It is apparent that for a proper treatment of an acid-base disturbance, the complete acid-base status of a patient should be known to a clinician. Although the body can compensate primary acid-base disturbances to a certain extent, therapeutic measurements must be taken as soon as possible to eliminate any primary acid-base disturbance. Moreover, severely ill patients on the ICU most often receive some form of artificial respiration. Being on mechanical ventilation means that the body cannot fully employ respiratory compensating mechanisms, making the ICU clinician even more responsible for keeping the pH of the arterial blood within acceptable boundaries.

For most ICU patients, an arterial blood gas analysis is performed on a routine basis, for instance every 3 or 6 hours. However, the interpretation of acid-base data is still regarded as difficult since several pieces of information must be evaluated at the same time in their clinical context. Multiple primary disturbances can be present at the same time, concealed by various degrees of compensation, making the diagnosis and monitoring of acid-base data a complex task.

This complexity is illustrated by the coexistence of two distinct methods for interpreting arterial acid-base parameters. One method uses *in vivo* information to interpret pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$, while the other method makes use of pH, PaCO_2 and a calculated *in vitro* parameter called base excess (BE). This latter method was developed around 1960 by Poul Astrup and Ole Siggaard-Andersen from Denmark and is therefore also known as the *Scandinavian view* [16].

Schwartz and Relman of the Tufts University School of Medicine in Boston (USA) criticised the *in vitro* approach and made a case for pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ [17]. This method is therefore also known as the *North American view*. The controversy between the two schools, which Bunker called 'The Great Trans-Atlantic Acid-Base Debate', still exists today, although many attempts were made to bridge the gap [16, 18-22].

1.4.2 The North American view; $\text{a}[\text{HCO}_3^-]$ and *in vivo* CO_2 buffer lines

In the North American view, a high value of PaCO_2 indicates a primary respiratory acidosis or a respiratory compensation for a metabolic alkalosis, while a low value of PaCO_2 indicates a primary respiratory alkalosis or a respiratory compensation for a primary metabolic acidosis. The metabolic component of an acid-base status is assessed with $\text{a}[\text{HCO}_3^-]$. A high value of $\text{a}[\text{HCO}_3^-]$ indicates a primary metabolic alkalosis or a metabolic compensation for a primary respiratory acidosis while a low

Table 1-1. *Compensations to primary acid-base disturbances in the North American view [13].*

disorder	primary change	compensatory response
metabolic acidosis	$\downarrow a[\text{HCO}_3^-]$	1.2 mmHg decrease in PaCO_2 for every 1 mmol/l fall in $a[\text{HCO}_3^-]$
metabolic alkalosis	$\uparrow a[\text{HCO}_3^-]$	0.7 mmHg elevation in PaCO_2 for every 1 mmol/l rise in $a[\text{HCO}_3^-]$
respiratory acidosis acute	$\uparrow \text{PaCO}_2$	1 mmol/l elevation in $a[\text{HCO}_3^-]$ for every 10 mmHg rise in PaCO_2
chronic		3.5 mmol/l elevation in $a[\text{HCO}_3^-]$ for every 10 mmHg rise in PaCO_2
respiratory alkalosis acute	$\downarrow \text{PaCO}_2$	2 mmol/l decrease in $a[\text{HCO}_3^-]$ for every 10 mmHg fall in PaCO_2
chronic		5 mmol/l decrease in $a[\text{HCO}_3^-]$ for every 10 mmHg fall in PaCO_2

$a[\text{HCO}_3^-]$ indicates a primary metabolic acidosis or a metabolic compensation for a primary respiratory acidosis. However, $a[\text{HCO}_3^-]$ cannot be used as a true metabolic parameter, since changes in PaCO_2 also effect $a[\text{HCO}_3^-]$.

The concept of the North American view is that *in vivo* data is used to calculate the expected rise or fall in $a[\text{HCO}_3^-]$ and/or PaCO_2 that occur in specific acid-base disorders. The empirically derived *in vivo* information has been compiled from a large number of clinical studies in which the normal compensatory reactions to each of the primary acid-base disorders has been investigated and quantified [23-30]. An observed value of $a[\text{HCO}_3^-]$ or PaCO_2 below or above the expected value of $a[\text{HCO}_3^-]$ or PaCO_2 is an indication for the presence and nature of a metabolic component or respiratory component of an acid-base disorder. Table 1-1 presents the empirically found expected compensatory rise and fall in $a[\text{HCO}_3^-]$ and PaCO_2 for the primary acid-base disturbances.

1.4.3 The Scandinavian view; standard bicarbonate and base excess

The North American view requires calculations to be performed at the bedside of a patient. Moreover, to predict the amount of rise or fall in primary acid-base values, the acid-base disturbance of a patient should be known *a priori*. To overcome the 'problems' of bedside calculations and the paradox of classifying an already known

acid-base disturbance, Astrup and Siggaard-Andersen developed the concept of the standard bicarbonate and the base excess as true metabolic acid-base parameters [11]. In 1960, Astrup described his equilibration method for the rapid measurement and calculation of the primary acid-base parameters pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ [12, 31]. In a microtonometer a blood sample is equilibrated with two known CO_2 gas mixtures, one with a high PaCO_2 and one with a low PaCO_2 . Plotting PaCO_2 and measured pH at both PaCO_2 values in a log PaCO_2 -pH diagram, and connecting the two points with a line yields the *in vitro* CO_2 equilibration curve. By measuring pH of the original blood sample and putting it in the plot, the actual PaCO_2 of the blood sample can be read from the CO_2 equilibration curve. With the Henderson-Hasselbalch equation $\text{a}[\text{HCO}_3^-]$ can be calculated.

With the log PaCO_2 -pH chart and the *in vitro* CO_2 equilibration curve of a patient, $\text{a}[\text{HCO}_3^-]$ can be calculated at any desired PaCO_2 value. Astrup proposed to use the $\text{a}[\text{HCO}_3^-]$ of a blood sample at a PaCO_2 of 40 mmHg as a true metabolic parameter, since this would be the concentration that would have been found in the blood sample if the influence of the respiration was eliminated. He called it the standard bicarbonate concentration or SB.

At the same time, Siggaard-Andersen completed his titration experiments in which he determined the CO_2 equilibration curves of normal blood and blood with known amounts of non-volatile acids and bases at a fixed PaCO_2 of 40 mmHg. Based on these experiments he added to the log PaCO_2 -pH diagram of Astrup a curved line representing the amount of non-volatile acid or base needed to titrate the blood sample at a PaCO_2 of 40 mmHg to a pH of 7.40 at a temperature of 37 °C. Astrup and Siggaard-Andersen called this the base excess or BE. Positive base excess values indicate a relative deficit of non-volatile acids while negative base excess values indicate a relative surplus of non-volatile acids. A base excess of 0 means that there is no metabolic component in the acid-base disorder. In modern analysers, BE is calculated from pH, PaCO_2 , $\text{a}[\text{HCO}_3^-]$ and the haemoglobin concentration of the arterial blood sample at hand.

The most important argument against the use of standard bicarbonate and base excess is that they are determined *in vitro*. The *in vitro* CO_2 equilibration curve is the equilibration curve of whole blood in a tube or syringe. It has been shown that *in vivo* buffering of protons is different from the *in vitro* buffering of protons [17]. This is mainly because *in vivo* buffering takes place in the extracellular fluid in which the haemoglobin concentration (a powerful chemical buffer) is lower than in whole blood. Both Siggaard-Andersen himself and Severinghaus proposed to calculate BE not with the measured haemoglobin concentration of the sample, but with a haemoglobin concentration of 5 g/dl, which is the Hb concentration relative to the total volume of extracellular fluid of the body [32, 33]. This BE is also known as BE_{ecf} (Base Excess of ex-

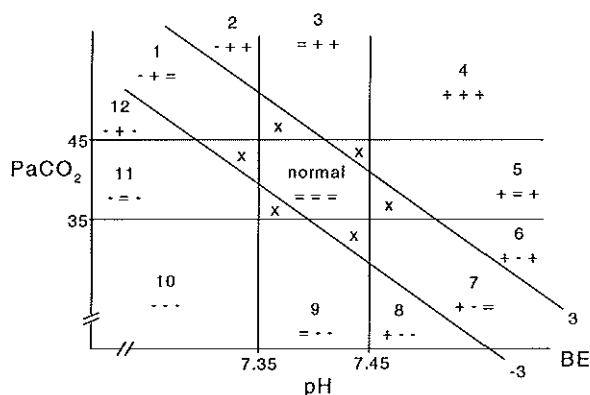


Figure 1-1. Areas of acid-base classification according to the method presented in Table 1-2. Combinations of low, high or normal observed values yield 12 specific acid-base disorder regions. In the normal region, all values are within their standard 95% normal reference intervals. The 'x' regions are formally not classifiable [34]. In these regions, one of the three observed acid-base values is outside its 95% univariate reference interval.

tracellular fluid), SBE (Standard Base Excess) and BE5 (Base Excess at a haemoglobin concentration of 5 g/dl).

With BE as a true metabolic parameter, classifying acid-base disturbances is now straightforward. Figure 1-1 and Table 1-2 show all possible acid-base classifications based on pH, PaCO_2 and BE. To determine whether an observed value for an acid-base parameter is too low, normal or too high, standard univariate 95% reference intervals are used. Table 1-3 presents the associated upper and lower cut-off values for the univariate 95% reference intervals of arterial pH, PaCO_2 , BE and $\text{a}[\text{HCO}_3^-]$.

1.5 Objective and scope of this thesis

Two main problems are currently associated with the interpretation and evaluation of arterial acid-base measurements in an intensive care setting.

The first problem occurs when classifying the acid-base variables pH, PaCO_2 and BE according to the method described in section 1.4.3. A strict adherence to the classification rules as described in Table 1-2 reveals that some combinations of observed values for the three acid-base variables can formally not be classified. This was found when an attempt was made to computerise the classification scheme of Astrup and Siggaard-Andersen in a rule-based expert system [34]. Typically, the *unclassifiable* situation occurs when only one of the three observed acid-base values is outside its 95% univariate reference interval, while the other two are within their 95% univariate

Table 1–2. Classification of acid-base disorders in the Scandinavian view. The signs ‘-’, ‘+’ and ‘=’ indicate an observed value being respectively below, above or within its 95% normal reference interval. See also Figure 1–1.

	pH	PaCO ₂	BE	classification
1	-	+	=	respiratory acidosis
2	-	+	+	partly compensated respiratory acidosis
3	=	+	+	compensated respiratory acidosis OR compensated metabolic alkalosis OR combined respiratory acidosis and metabolic alkalosis
4	+	+	+	partly compensated metabolic alkalosis
5	+	=	+	metabolic alkalosis
6	+	-	+	combined respiratory and metabolic alkalosis
7	+	-	=	respiratory alkalosis
8	+	-	-	partly compensated respiratory alkalosis
9	=	-	-	compensated respiratory alkalosis OR compensated metabolic acidosis OR combined respiratory alkalosis and metabolic acidosis
10	-	-	-	partly compensated metabolic acidosis
11	-	=	-	metabolic acidosis
12	-	+	-	combined respiratory and metabolic acidosis
	=	=	=	normal
x				unclassifiable

intervals. In Figure 1–1, this situation is represented by the triangular regions denoted by ‘x’.

The second problem originates from the use of the 95% univariate reference interval as the standard statistical model for evaluating the ‘normalcy’ of observed arterial acid-base values from intensive care patients.

A first critical note on the use of 95% reference intervals is that the determination of the respective reference intervals and the characteristics of the reference population are completely unknown. In general, reference intervals are derived from a representative

Table 1–3. Upper and lower limits of the standard 95% normal reference intervals for the acid-base variables in arterial blood.

Acid-base variable	lower limit	upper limit
pH	7.35	7.45
PaCO ₂	35 mmHg	45 mmHg
BE	-3 mmol/l	3 mmol/l
a[HCO ₃ ⁻]	21 mmol/l	27 mmol/l

sample of a (often) 'healthy' reference population [35]. The process of defining the reference criteria, the selection of reference individuals, analytical considerations and the use of statistical techniques for defining valid 95% univariate reference intervals are described in detail [36-40]. Nothing is known, however, about the determination of the 95% univariate reference intervals that are presented in Table 1-3. If we assume that the intervals are defined on a 'healthy' reference population, what is the value of these intervals in an intensive care setting where it is to be expected that most of the observed acid-base values will be outside these 'health'-based intervals?

A second critical note concerns the number of reference intervals used. Traditionally, the interpretation of the acid-base status involves the use of three separate 95% reference intervals for evaluating the acid-base variables: pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ in the North American view, or pH, PaCO_2 and BE in the Scandinavian view. From the Henderson-Hasselbalch equation (Equation 1-4), however, it is apparent that the relationship between pH, $\log \text{PCO}_2$ and $\log [\text{HCO}_3^-]$ is a linear one. This can best be appreciated when Equation 1-4 is rewritten as:

$$\text{pH} - \log \text{a}[\text{HCO}_3^-] + \log \text{PCO}_2 = \text{pK}' - \log \alpha, \quad (1-5)$$

with pK' and $\log \alpha$ both being constant.

Moreover, in Chapter 2 it will be demonstrated that the relationship between pH, PaCO_2 and BE is also (almost) linear. Consequently, as Madias [41] already pointed out, it is illogical and fundamentally wrong that three separate 95% univariate reference intervals are used, while only two of the three variables can change independently.

A third critical note on the use of univariate 95% reference intervals is that the 95% univariate interval is not the proper statistical model for evaluating arterial acid-base values. Theoretically, the use of more than one 95% univariate reference interval in case of a simultaneous evaluation of multiple variables – which is the case when interpreting arterial acid-base values – is prone to error and leads *a priori* to more false positive and false negative observations [42-44]. This will be illustrated in detail in Chapter 4.

This thesis describes a new multivariate statistical reference model for evaluating and classifying arterial acid-base variables in an intensive care environment that addresses all of the above mentioned problems. The essence of the model is that a single 95% multivariate statistical reference region is defined on a large reference population consisting of acid-base data coming from intensive care patients themselves. Furthermore, the multivariate reference model is not defined on the original acid-base measurements but rather on the values obtained after applying a mathematical data reduction transformation procedure. Finally, based on the outcome of this transformation, a

new way of classifying pH, PaCO₂ and BE values will be proposed that will have no *unclassifiable* categories, unlike the method described in 1.4.3.

The outline of this thesis is as follows. In Chapter 2, the mathematical data reduction technique will be introduced, together with the results of various transformed large acid-base data sets coming from several ICUs. In Chapter 3, a two-dimensional graphical representation of the three acid-base variables will be presented, based on the mathematical transformation as described in Chapter 2. Also, the new classification model for pH, PaCO₂ and BE combinations will be described. Then, in Chapter 4, the technique for defining a 95% multivariate patient-based reference region for the acid-base variables will be described. Chapter 5 presents the computational methods involved in the data reduction transformation procedure and the construction of the multivariate reference model. It also presents the prototype computer programs that were built for defining multivariate acid-base reference regions and describes their use in daily clinical practice. Chapter 6 exemplifies the use and practicability of the proposed graphical representation of acid-base data using measurements from three intensive care patients. In Chapters 7 and 8, the results of the clinical evaluation of the multivariate acid-base reference regions and classification model can be found. The thesis is concluded with a general discussion.

1.6 References

1. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. Tietz textbook of clinical chemistry. 2nd Edition, W.B. Saunders Company, 1994; 454-484.
2. Guder WG, Büttner J. Clinical chemistry in laboratory medicine in Europe-past, present and future challenges. *Eur J Clin Chem Clin Biochem* 1997; 35:487-494.
3. Sunderman FW, Sr. The foundation of clinical chemistry in the United States. *Clin Chem* 1994; 40:835-842.
4. Büttner J. Clinical chemistry as scientific discipline: historical perspectives. *Clin Chim Acta* 1994; 232:1-9.
5. Wold S. Chemometrics, why, what and where to next? *J Pharm Biomed Anal* 1991; 9:589-596.
6. Karjalainen EJ. The role of chemometrics in medical decision making. *Scand J Clin Lab Invest Suppl* 1990; 202:109-111.
7. Weil MH, Von Planta M, Rackow EC. Critical care medicine: introduction and historical perspective. In: Shoemaker W, ed. Textbook of Critical Care. 2nd Edition, Philadelphia: W.B. Saunders Company, 1989; 1-5.
8. Muakkassa FF, Rutledge R, Fakhry SM, et al. ABG's and arterial lines: the relationship to unnecessarily drawn arterial gas samples. *J Trauma* 1990; 30:1087-1095.
9. Severinghaus JW, Astrup PB. History of blood gas analysis. Boston: Little, Brown and Company, 1987, Lange BP, ed., *International Anesthesiology Clinics*; vol 25.
10. Astrup P, Severinghaus JW. The history of blood gases, acids and bases. The history of blood gases, acids and bases. 1st Edition, Copenhagen: Munksgaard International Publishers, 1986; 264-295.
11. Astrup P, Jørgensen K, Siggaard-Andersen O, et al. The acid-base metabolism, a new approach. *Lancet* 1960; 1035-1039.
12. Astrup P. A new approach to acid-base metabolism. *Clin Chem* 1961; 7:1-15.
13. Rose BD. Clinical Physiology of Acid-Base and Electrolyte Disorders. 4th Edition, New York: McGraw-Hill, Inc., 1994; 853.
14. Hainsworth R. Acid-base balance. Physiological Society Study Guides. Manchester: Manchester University Press, 1986; 155; vol 1.
15. Lane EE, Walker JF. Clinical arterial blood gas analysis. St. Louis: The C.V. Mosby Company, 1987; 247.
16. Bunker JP. The great trans-atlantic acid-base debate. *J Anesthesiol* 1965; 26:591-594.
17. Schwartz WB, Relman AS. A critique of the parameters used in the evaluation of acid-base disorders. "Whole-blood buffer base" and "standard bicarbonate" compared with blood pH and plasma bicarbonate concentration. *New Engl J Med* 1963; 268:1382-1388.
18. Rispens P, Zijlstra WG, Van Kampen EJ. Significance of bicarbonate for the evaluation of non-respiratory disturbances of acid-base balance. *Clin Chim Acta* 1974; 54:335-347.
19. Siggaard-Andersen O, Fogh-Andersen N. Base excess or buffer base (strong ion difference) as measure of a non-respiratory acid-base disturbance. *Acta Anaesthesiol Scand Suppl* 1995; 107:123-128.
20. Severinghaus JW. Acid-base balance nomogram. A Boston-Copenhagen détente. *Anesthesiology* 1976; 45:3-5.
21. Severinghaus JW. Acid-base balance controversy. Editorial introduction. *J Clin Monit* 1991; 7:274-275.
22. Severinghaus JW. Siggaard-Andersen and the "Great Trans-Atlantic Acid-Base Debate". *Scand J Clin Lab Invest* 1993; 53:99-104.

23. Arbus GS, Herbert LA, Levesque PR, et al. Characterization and clinical application of the "significance band" for acute respiratory alkalosis. *New Engl J Med* 1969; 280:117-123.
24. Bushinsky DA, Coe FL, Katzenberg C, et al. Arterial PCO₂ in chronic metabolic acidosis. *Kidney Int* 1982; 22:311-314.
25. Javaheri S, Shore NS, Rose B, et al. Compensatory hypoventilation in metabolic alkalosis. *Chest* 1982; 81:296-301.
26. Javaheri S, Kazemi H. Metabolic alkalosis and hypoventilation in humans. *Am Rev Respir Dis* 1987; 136:1011-1016.
27. Pierce NF, Fedson DS, Brigham KL, et al. The ventilatory response to acute base deficit in humans. Time course during development and correction of metabolic acidosis. *Ann Intern Med* 1970; 72:633-640.
28. Polak A, Haynie GD, Hays GM, et al. Effects of chronic hypercapnia on electrolyte and acid-base equilibrium. I. Adaptation. *J Clin Invest* 1961; 40:1223.
29. Van Yperselle de S, Brasseur L, De Coninck JD. The "carbon dioxide response curve" for chronic hypercapnia in man. *New Engl J Med* 1966; 275:117-122.
30. Gennari FJ, Goldstein MB, Schwartz WB. The nature of the renal adaptation to chronic hypocapnia. *J Clin Invest* 1972; 51:1722-1730.
31. Siggaard-Andersen O, Engel K, Jørgensen K, et al. A micro method for determination of pH, carbon dioxide tension, base excess and standard bicarbonate in capillary blood. *Scand J Clin Lab Invest* 1960; 12:172-176.
32. Siggaard-Andersen O. An acid-base chart for arterial blood with normal and pathophysiological reference areas. *Scand J Clin Lab Invest* 1971; 27:239-245.
33. Severinghaus JW. Acid-base balance controversy: case for standard-base excess as the measure of nonrespiratory acid-base imbalance. *J Clin Monit* 1991; 7:276-277.
34. Wulkan RW. Expert systems and multivariate analysis in clinical chemistry. Rotterdam: Erasmus University Rotterdam, 1992; 111 pp.
35. Solberg HE, Gräsbeck R. Reference values. *Adv Clin Chem* 1994; 27:1 -79.
36. Dybkær R, Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1987) on the theory of reference values. Part 6. Presentation of observed values related to reference values. *J Clin Chem Clin Biochem* 1987; 25:657-662.
37. PetitClerc C, Solberg HE. International Federation of Clinical Chemistry (IFCC). Approved Recommendation (1987) on the theory of reference values. Part 2. Selection of individuals for the production of reference values. *J Clin Chem Clin Biochem* 1987; 25:639-644.
38. Solberg HE, PetitClerc C. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *J Clin Chem Clin Biochem* 1988; 26:593-598.
39. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem* 1987; 25:645-656.

40. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem* 1987; 25:337-342.
41. Madias NE, Adroqué HJ, Horowitz GL, et al. A redefinition of normal acid-base equilibrium in man: Carbon dioxide tension as a key determinant of normal plasma bicarbonate concentration. *Kidney Int* 1979; 16:612-618.
42. Stambhuis IH, Bezemer PD, Kuik D. Evaluation of univariate ranges with a multivariate standard. *J Clin Epidemiol* 1988; 41:359-366.
43. Solberg HE. Multivariate reference regions. *Scand J Clin Lab Invest Suppl* 1995; 222:3-5.
44. Schoen I, Brooks SH. Judgment based on 95% confidence limits. *Statistical Considerations* 1969; 53:190-193.

The application of principal component analysis (PCA) to reduce the dimensionality of trivariate arterial acid-base data distributions

Marcel Hekking, Jan Lindemans, Edzard S. Gelsema

Parts are published in: International Journal of Bio-Medical Computing 1994; 209 - 221

2.1 Introduction

In clinical practice, the interpretation of the arterial acid-base status is performed by a simultaneous evaluation of three acid-base laboratory parameters; pH of arterial blood, the partial pressure of carbon dioxide (CO_2) in arterial blood (PaCO_2) for the evaluation of the respiratory component, and either the arterial bicarbonate-ion (HCO_3^-) concentration ($a[\text{HCO}_3^-]$) or the base excess (BE) for the evaluation of the metabolic component [1, 2]. Arterial acid-base values, however, are linearly related [1]. This means that if two of the three parameters are known, the third can be calculated. Thus, clinicians use three acid-base parameters to assess the acid-base status of a patient as if they were independent of each other, although only two of the three variables can change independently.

Although a strict linear relationship is not self-evident for pH, PaCO_2 and BE, an almost linear relationship is also present between these three variables. This was discovered during an earlier study [3] in which the distributions of large collections of pH, PaCO_2 and BE values were explored, using graphical software with capabilities of on-line three-dimensional rotation. During these explorations it was realised that, when plotting the combinations of the three variates as they occur in practice in three-dimensional space, the points are located on a surface with only a slight curvature.

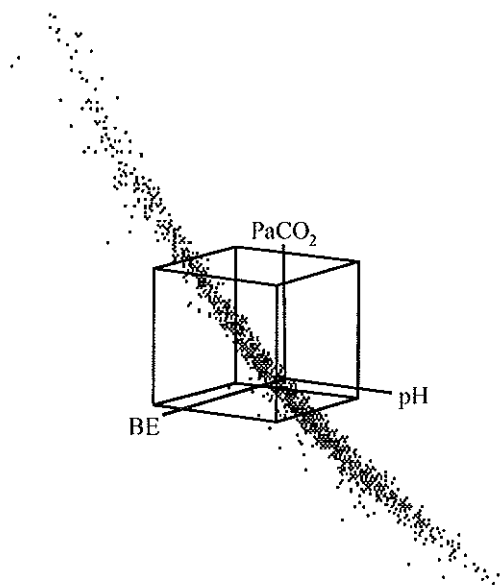


Figure 2-1. Three-dimensional view of a rotated pH, PaCO_2 and base excess (BE) data set. The cube represents the 95% reference volume as defined by the 95% univariate reference intervals of Table 2-1.

Figure 2-1 displays such a pH, PaCO_2 and BE distribution in three dimensions. The distribution is rotated in such a way that the curved plane of measurements can be easily viewed. The cube in the middle is the volume in three dimensions that represents the standard reference volume, as built from the three standard univariate 95% reference intervals for pH, PaCO_2 and BE of Table 2-1.

Having observed that the relationship between the arterial acid-base variables (both for combinations of pH, PaCO_2 and BE and pH, $\log \text{PaCO}_2$ and $\log a[\text{HCO}_3^-]$) is an almost linear one, the goal is to arrive at a mathematical description of this relationship and to investigate its departure from linearity. The mathematical technique to be used

Table 2-1. Means (m) and standard deviations (s) as derived from the standard 95% reference intervals for the arterial acid-base variables. The standard deviations are calculated by assuming the 95% reference intervals to be 4 standard deviations wide.

	95 % reference interval	m	s
pH	7.35 – 7.45	7.40	0.025
PaCO ₂ (mmHg)	35 – 45	40	2.5
PaCO ₂ (kPa)	4.7 – 6.0	5.33	0.325
BE (mmol/l)	-3 – 3	0	1.5
a[HCO ₃ ⁻] (mmol/l)	21 – 27	24	1.5
log PaCO ₂ (mmHg)	log 35 – log 40	log 40	(log 35 – log 40) / 4
log PaCO ₂ (kPa)	log 4.7 – log 6.0	log 5.33	(log 4.7 – log 6.0) / 4
log a[HCO ₃ ⁻]	log 21 – log 27	log 24	(log 21 – log 27) / 4

for such an investigation is a principal component analysis (PCA). PCA is a multivariate statistical technique for the compression of large data matrices [4, 5]. In this chapter, the results of a PCA of several distributions of acid-base data, coming from various ICUs, are described.

2.2 Materials and methods

2.2.1 Patient data

Six acid-base data sets from four different intensive care units were submitted to PCA. Data set *AZRbe* contains 1500 unselected combinations of pH, PaCO₂ and BE values from patients of the respiratory ICU of the Dijkzigt academic hospital, Rotterdam, The Netherlands. The term *unselected* means that no specific selection criteria were applied. In fact, all data sets are constructed by sampling the acid-base data as consecutively measured in the respective clinical laboratories. Data set *OLVGbe* contains 1500 unselected combinations of pH, PaCO₂ and BE values from patients of the general ICU of the OLVG hospital, Amsterdam, The Netherlands. Data set *OLVGab* comprises the 1500 combinations of pH, log PaCO₂ and log a[HCO₃⁻] values from the same patients as the *OLVGbe* data set. Data set *SKZbe* contains 1500 combinations of pH, PaCO₂ and BE values from patients of the neonatal ICU of the Sophia Children's hospital, Rotterdam, The Netherlands. The data set is composed of equal numbers of data in three age groups: new-borns younger than five days, infants between five days and one month of age, and infants aged between one month and one year. Data set *ELibe* contains 1500 unselected combinations of pH, PaCO₂ and BE values from the general ICU of the St. Elisabeth hospital, Tilburg, The Netherlands.

Data set *ELLab* comprises the 1500 combinations of pH, log PaCO₂ and log a[HCO₃⁻] from the same patients as the *ELLbe* data set.

2.2.2 Standardisation

Prior to the principal component analysis of an acid-base data set, each variable in the data set was standardised with fixed means and standard deviations according to:

$$z_i = \frac{x_i - m}{s}, \quad i = 1, \dots, N \quad (2-1)$$

where m and s are, respectively, the mean and standard deviation for the respective acid-base variables as presented in Table 2-1, while N is the total number of cases in the data set. The z_i -values are therefore the deviations from the mean m , measured in units of the corresponding standard deviation s .

2.2.3 Principal component analysis

The standardised data sets were then subjected to PCA. PCA is a mathematical transformation that enables the reduction of the number of variables in a multivariate data set whilst preserving as much of the original information as possible [4, 5]. Assuming a multivariate data set with p variables (x_1, x_2, \dots, x_p), PCA finds a new set of derived variables (z_1, z_2, \dots, z_p) that are linear functions of x_1, x_2, \dots, x_p with the following properties:

- z_1 has maximum possible variance among all possible linear functions of x_1, x_2, \dots, x_p .
- z_k has maximum possible variance among all possible linear functions of x_1, x_2, \dots, x_p , subject to z_k being uncorrelated with z_1, z_2, \dots, z_{k-1} , for $2 \leq k \leq p$ [4].

The derived variables z_1, z_2, \dots, z_p are called the principal components or PCs.

In linear algebraic terms, PCs are determined with an eigenvalue transformation of the variance-covariance matrix as derived from the multivariate data set. For a set of N vectors \mathbf{x}_i ($i = 1, \dots, N$) in a p -dimensional data set, the variance-covariance matrix V is defined as:

$$V = \frac{\sum_{i=1}^N (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^T}{N(N-1)} \quad (2-2)$$

where \mathbf{m} is the vector of the mean of the set \mathbf{x}_i ($i = 1, \dots, N$) and the superscript T indicates transposition of a vector, in the convention that an untransposed vector is a column vector. The eigenvalue transformation yields a transformation matrix U , which transforms the original vectors \mathbf{x} into vectors \mathbf{y} , according to $\mathbf{y} = U \mathbf{x}$, such that the variance-covariance matrix $W = UVU^T$ of the transformed vectors \mathbf{y} is a diagonal matrix. If U is constrained to be a unitary matrix, the component variances of the trans-

formed vectors y appear as eigenvalues (λ) in the analysis and are found as the diagonal elements of W .

Since the eigenvalue transformation diagonalises the variance-covariance matrix, the total variance in the set of original vectors x is decomposed into p orthogonal directions. Thus, for a set of p -dimensional vectors x for which it is observed that most of the variance is confined to a subspace of dimension $l < p$, it is expected that the components 1 through l of the transformed vectors y contain most of the useful information. The components $l + 1$ through p have only a small variance, and thus convey (almost) no information. In the present situation, $p = 3$ and due to the (almost) linear relationships between the variables in the standardised data sets, it is expected that $l = 2$.

2.3 Results

Table 2-2 presents the results of the principal component analysis of each data set. The eigenvalues λ are shown for each of the three principal components (hereafter referred to as PC1, PC2 and PC3). The eigenvalues λ explain the contribution of each of the principal component to the total variance in the data set prior to PCA. For instance, in the *AZRbe* data set, PC1, PC2 and PC3 explain 62.37%, 36.91% and 0.71% of the total variance in the initial data set, respectively. From Table 2-2 it can be concluded that for each data set, the percentage of variance explained by the third principal component (PC3) is only small compared to the variance explained by the first two principal components (PC1 and PC2) together. The explained variance by PC1 and PC2 for each data set is more than 99%. The data sets *OLVGab* and *ELIab* show the smallest explained variance by PC3; 0.03% and 0.09%, respectively. This is not surprising since these data sets consist of pH, $\log PaCO_2$ and $\log a[HCO_3^-]$ values and these variables are linearly related according to the Henderson-Hasselbalch equation (see Chapter 1) [1].

Table 2-2. Eigenvalues λ and contributions to the total variance in the initial data set (in brackets) for each principal component.

	PC1	PC2	PC3
AZRbe	20.54 (62.37%)	12.16 (36.91%)	0.235 (0.71%)
OLVGbe	26.04 (69.70%)	11.22 (30.03%)	0.101 (0.27%)
OLVGab	25.52 (74.30%)	8.82 (25.67%)	0.009 (0.03%)
SKZbe	21.09 (78.01%)	5.87 (21.72%)	0.072 (0.27%)
ELIbe	23.62 (64.55%)	12.83 (35.07%)	0.137 (0.37%)
ELIab	20.36 (58.91%)	14.17 (41.00%)	0.032 (0.09%)

For each data set, a matrix U can be built from the three separate normalised eigenvectors ϵ , which are used to calculate the associated principal component values from a combination of standardised original acid-base values. Table 2–3 shows the normalised eigenvectors ϵ of each principal component for all data sets. With the matrices U , new trivariate distributions of principal component values were calculated from the original standardised acid-base data sets.

Table 2–3. Normalised eigenvectors ϵ of each principal component as obtained after PCA. The eigenvectors for PC1, PC2 and PC3 are columns 1, 2 and 3, respectively of the eigenmatrix matrix U . The eigenmatrix U will be the input of calculations to be presented in the next chapters.

	PC1	PC2	PC3
AZRbe	(0.297, -0.885, -0.358)	(0.709, -0.046, 0.703)	(0.639, 0.463, -0.614)
OLVGbe	(-0.039, -0.777, -0.628)	(0.742, -0.444, 0.503)	(0.669, 0.446, -0.594)
OLVGab	(0.044, 0.686, 0.727)	(-0.839, 0.420, -0.346)	(0.542, 0.594, -0.594)
SKZbe	(0.635, -0.757, 0.153)	(0.425, 0.508, 0.749)	(0.645, 0.411, -0.649)
ELIbe	(0.732, -0.184, 0.656)	(-0.184, 0.874, 0.450)	(0.656, 0.450, -0.605)
ELIab	(0.734, -0.012, 0.679)	(-0.391, 0.811, 0.436)	(0.556, 0.585, -0.590)

Table 2–4 presents the characteristics of the resulting principal component value distributions. For each data set, the standard deviation of the PC3 distribution is small compared to the standard deviations of the PC1 and PC2 distributions. Data set *OLVGab* and *ELIab* have the smallest standard deviations for the third principal component value distribution: 0.097 and 0.178, respectively. This is in accordance with the results presented in Table 2–2.

Table 2–4. Characteristics (m is mean and s is the standard deviation) of the principal component value distributions.

	PC1		PC2		PC3	
	m	s	m	s	m	s
AZRbe	1.282	4.532	0.695	3.487	0.177	0.485
OLVGbe	-0.773	5.103	-0.315	3.349	-0.215	0.317
OLVGab	0.357	5.052	0.250	2.969	-0.282	0.097
SKZbe	-3.358	4.593	-2.676	2.423	-0.109	0.268
ELIbe	-1.025	4.860	0.062	3.582	-0.005	0.370
ELIab	-1.008	4.513	-0.038	3.765	-0.071	0.178

Since for each data set the amount of explained variance is more than 99% when only PC1 and PC2 are considered, there is no significant loss of information when the acid-base values are projected onto the plane spanned by PC1 and PC2. Hence, any

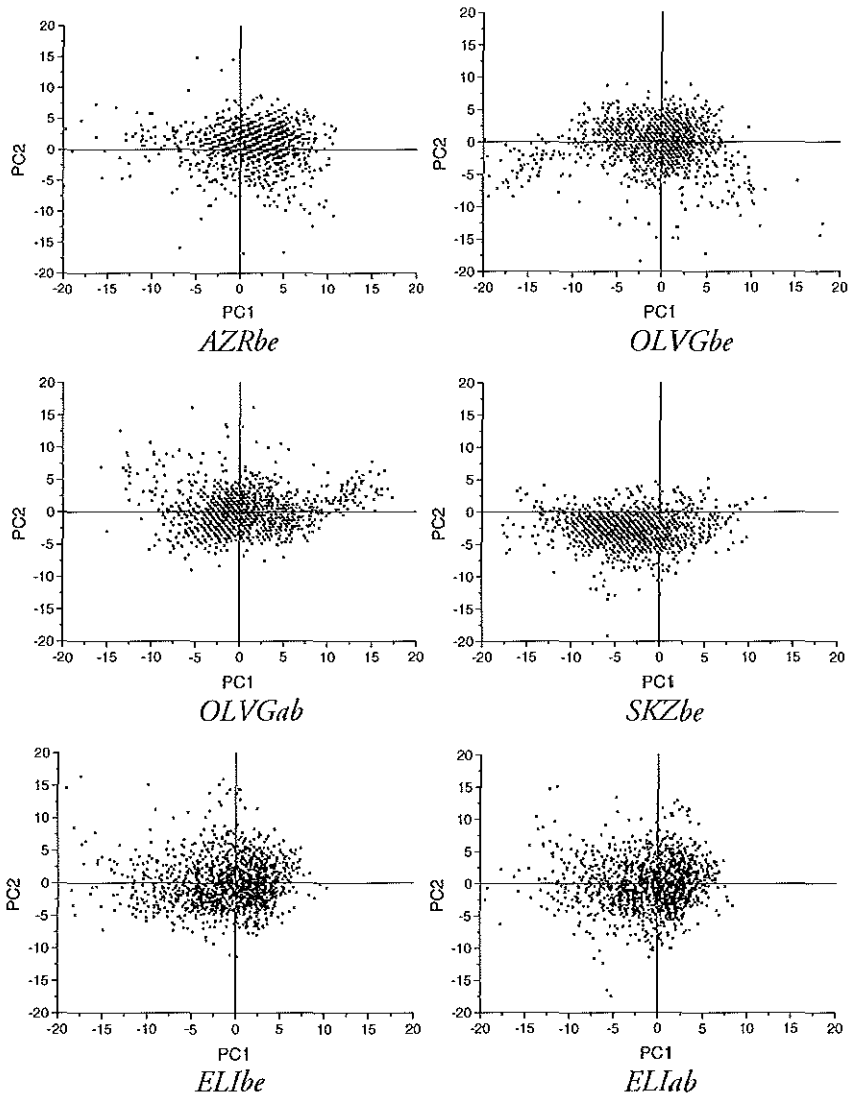


Figure 2-2. Projections of the 1500 acid-base observations on the plane spanned by the first two principal components (PC1 and PC2).

quantitative analysis based on PC1 and PC2 addresses the complete acid-base status. In Figure 2-2, scatterplots of PC2 versus PC1 are shown for all data sets.

Since the plane of measurements in case of a pH, PaCO_2 , and BE data set is slightly curved (see Figure 2-1), it is interesting to investigate the effect of the curvature on the distribution of PC3 values. Therefore, the PC3 distribution characteristics of two data sets were investigated. This was done by constructing box-whisker plots of groups of PC3 values that are increasingly further away from the bivariate PC1-PC2 mean. As

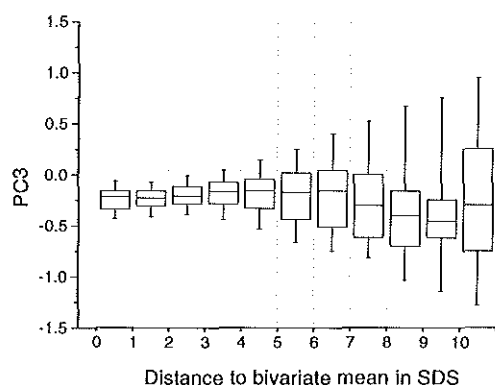


Figure 2-3. Box-whisker plots as a function of the distance from the mean in the PC1-PC2 plane for data set *OLVGbe*. SDS stands for 'standard deviation score'.

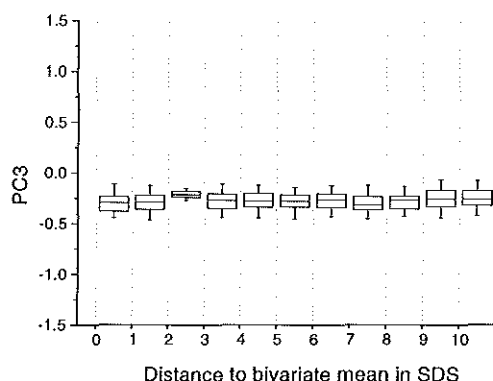


Figure 2-4. Box-whisker plots as a function of the distance from the mean in the PC1-PC2 plane for data set *OLVGab*.

a cut-off point, a distance of 1 standard deviation score was chosen with a maximum of 10, yielding 11 groups of data. A box-whisker plot provides a graphical representation of the distribution of values in a given data set. The outer top and bottom horizontal lines of the box-whisker plots indicate the 95th and 5th percentiles of a distribution, respectively. The top and bottom horizontal lines enclosing the box denote the 75th and the 25th percentile, respectively. The horizontal line inside the box denotes the median.

Figure 2-3 shows the box-whisker plots for the *OLVGbe* data set. In Figure 2-4 the box-whisker plots are shown for the *OLVGab* data set. Comparing both figures, it is apparent that with increasing distance from the mean in the plane spanned by the first two principal components PC1 and PC2, the variance in the PC3 distribution increases for data set *OLVGbe*, while the variance in the PC3 distribution of data set *OLVGab* remains the same for all distance strata. These figures illustrate the slight curvature of a PCA transformed pH, PaCO₂ and BE data set which is absent in a PCA transformed pH, log PaCO₂ and log a[HCO₃⁻] data set.

Figure 2-5 presents a histogram of the 1500 calculated PC3 values of the transformed *OLVGab* distribution. The straight line in the normal probability plot in the upper part of Figure 2-5 indicates that the 1500 PC3 values are normally distributed. This was confirmed with a Kolmogorov-Smirnov distribution fit test (D_{\max} of 0.03 with a p -value of 0.118). Since these PC3 values are normally distributed, a parametric 95% reference interval may be derived from this distribution as $m \pm 2s$, resulting in a reference range of -0.472 to -0.092. The calculated PC3 value of a pH, log PaCO₂, log a[HCO₃⁻] combination from an ICU patient of the OLVG hospital, transformed with the corresponding eigenvectors of Table 2-3, will have a probability of 95% of being

located within this interval. A similar analysis, however, on the 1500 PC3 values of the transformed data set *ELIab* showed a bimodal distribution of PC3 values (Figure 2–6). Consequently, the distribution was found to be significantly deviating from a normal distribution (D_{\max} of 0.263 with a p -value < 0.01).

2.4 Discussion

In 1979, Madias *et al.* [6] already noted that, when evaluating an acid-base status, it is illogical and fundamentally wrong to use pH, PaCO_2 as well as $\text{a}[\text{HCO}_3^-]$, while only two of these three variables are free to change independently. He proposed to evaluate acid-base disorders with only two of the three basic acid-base variables. This means, however, that clinicians are deprived of one of the three variables on which the interpretation of the acid-base status is traditionally based.

In this chapter, a solution is proposed that allows a quantitative analysis of all three basic acid-base variables while being faithful to the interdependence between them. A multivariate statistical technique called principal component analysis (PCA) was used to reduce the dimensionality of large trivariate distributions of acid-base variables. Results show that the acid-base status can be faithfully represented in a principal component subspace defined by the principal components PC1 and PC2, without significant loss of information. The distortion, measured as a percentage of variance not represented, is shown to be less than 0.7% for all the data sets investigated. The (small) percentage of explained

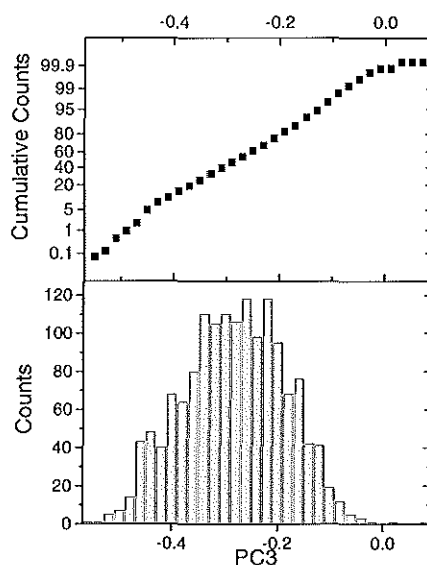


Figure 2–5. Histogram and normal probability plot of the 1500 third principal component values (PC3) of the OLVGab data set.

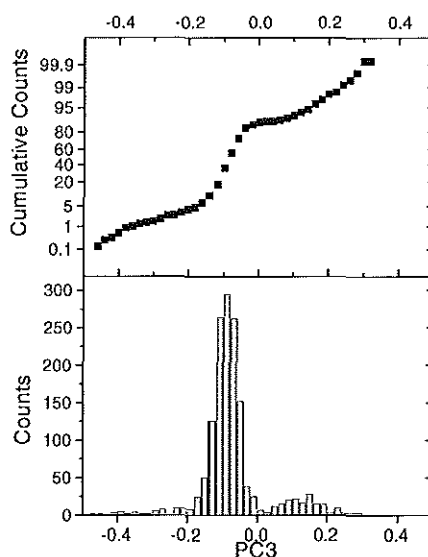


Figure 2–6. Histogram and normal probability plot of the 1500 third principal component values (PC3) of the ELIab data set.

variance by PC3 in data sets of pH, log PaCO₂ and log a[HCO₃⁻] (data sets *OLVGab* and *ELIab*) may be attributed to rounding effects and analytical imprecision. For the other data sets, consisting of pH, PaCO₂ and BE values, the curvature of the plane of measurements is an extra source of variance resulting in larger percentages of variance explained by PC3. However, this source of variance is only minor and for each data set it is therefore justified that quantitative analyses of acid-base disorders be based on PC1 and PC2 values after a PCA transformation, rather than on the original acid-base values. Furthermore, projection of the original points onto the PC1-PC2 subspace is (almost) distortionless. In Chapter 3, this characteristic is used to define a sound way to graphically represent all three acid-base variables in a single two-dimensional representation.

The minor variance in PC3 may also serve as a plausibility check for acid-base laboratory values; each transformed combination of pH, PaCO₂ and a[HCO₃⁻]/BE must lead to a small PC3 value. For a pH, log PaCO₂ and log a[HCO₃⁻] data set, PC3 must be within the 95% reference interval for PC3 as obtained from the PC3 values after PCA of an acid-base data set. For instance, the 95% reference interval for PC3 of the *OLVGab* data set was found to be -0.472 to 0.092. If a transformed combination of acid-base measurements is not within the interval, then it may be concluded that this specific combination of pH, PaCO₂ and a[HCO₃⁻] is not valid. Note that the interval is not equally centred around zero. From the definition of PCA one would expect that, when calculated means and standard deviations are used, the mean value for all principal component values would be zero. However, for the standardisation procedure the fixed means and standard deviations of Table 2-1 were used, leading to the observation that the mean values of the principal components are different from 0 for the various data sets, since they have different means and variances for the original acid-base values.

Checking whether the PC3 value of a transformed acid-base observation is within the 95% reference interval is only possible for data sets of pH, log PaCO₂ and log a[HCO₃⁻], since the variance in PC3 is independent of the distance of an observation to the PC1-PC2 bivariate mean (see Figure 2-6). Observations in a data set of pH, PaCO₂ and BE are located on a slightly curved plane of measurements, resulting in the effect that with increasing distances from the PC1-PC2 bivariate mean, the variance in PC3 increases (see Figure 2-3). To check the plausibility of a transformed pH, PaCO₂ and BE combination one could either use the variance in PC3 as found for data with distances larger than or equal to 10 standard deviations scores (≥ 10), or use the variance in PC3 in the associated distance group.

One could argue that the relationship between the acid-base variables could be described by studying the formula used in acid-base analysers to calculate a[HCO₃⁻] or BE from measured pH, PaCO₂ and haemoglobin. The advantage of the approach pre-

sented in this chapter, however, is that no prior knowledge is needed about the formula with which the $a[\text{HCO}_3^-]$ or the BE are calculated. The method, therefore, adapts itself to the instruments used.

2.5 Acknowledgements

I am indebted to dr. R.N.M. Weijers and dr. D. Zandstra for making the data sets *OLVGbe* and *OLVGab* available. Dr. B. van der Berg and dr. R.W. Wulkan contributed similar material: the data sets *AZRbe* and *SKZbe*, respectively. Finally, dr. J.E. van Puyenbroek and dr. B. Speelberg provided the data sets *ELIbe* and *ELIab*.

2.6 References

1. Rose BD. Clinical Physiology of Acid-Base and Electrolyte Disorders. 4th Edition, New York: McGraw-Hill, Inc., 1994; 853.
2. Astrup P, Jørgensen K, Siggaard-Andersen O, et al. The acid-base metabolism, a new approach. *Lancet* 1960;1035-1039.
3. Gelsema ES, Leijnse B, Wulkan RW. A multi-dimensional analysis of three chemical quantities in the blood. *Med Inform* 1991; 16:43-54.
4. Jolliffe IT, Morgan BJT. Principal component analysis and exploratory factor analysis. *Stat Meth Med Res* 1992; 1:69-95.
5. Jolliffe IT. Principal Component Analysis. New York: Springer-Verlag, 1986, *Springer Series in Statistics*; vol 12.
6. Madias NE, Adroqué HJ, Horowitz GL, et al. A redefinition of normal acid-base equilibrium in man: Carbon dioxide tension as a key determinant of normal plasma bicarbonate concentration. *Kidney Int* 1979; 16:612-618.

The graphical representation and classification of arterial acid-
base data in a principal component subspace

Marcel Hekking, Edzard S. Gelsema, Jan Lindemans

Parts are published in: International Journal of Bio-Medical Computing 1994; 209 –
221

3.1 Introduction

Numerous methods, nomograms, charts and graphics for the representation and evaluation of the primary acid-base variables have been developed since Henderson first demonstrated the importance of the $a[\text{HCO}_3^-]/\text{CO}_2$ buffer system for the acid-base equilibrium in human blood [1-7]. In 1921, Van Slyke published one of the first acid-base charts which was a plot of the bicarbonate-ion (HCO_3^-) concentration versus the pH of blood [8]. Later he altered this chart by adding iso-carbon-dioxide (CO_2) lines [9]. In 1931, Peters and Van Slyke published a plot in which the logarithm of the total plasma CO_2 content was plotted against the logarithm of the CO_2 tension. Many other charts developed since then are merely variations on the original acid-base charts introduced by Van Slyke.

In the 1950's, Davenport popularised the plot of the arterial bicarbonate-ion concentration ($a[\text{HCO}_3^-]$) versus pH with curved isopleths of partial pressures of carbon-dioxide in arterial blood (PaCO_2) for the representation and evaluation of acid-base disorders [10]. The 'followers' of the *in vivo* approach (see Chapter 1) mainly use this type of chart. When in 1960 Astrup and Siggaard-Andersen developed the concept of the base excess (BE) and the accompanying classification scheme, they proposed a chart for the representation of pH, PaCO_2 and BE values. The Siggaard-Andersen chart is a plot of logarithmic PaCO_2 versus pH in which equal BE values are represented by straight isopleths running from north-west to south-east [11, 12].

However, a major disadvantage of most acid-base representations, whether based on BE or $a[\text{HCO}_3^-]$, is that they try to represent three variables with only two coordinates. Representing three variables in only two dimensions leads to the phenomenon that equal changes in the acid-base status as measured with the three acid-base variables may not be displayed as equal distances in such a two-dimensional representation. These charts are therefore less suitable for the representation of consecutive acid-base observations from a single patient. As Siggaard-Andersen observed of his own chart: 'A major disadvantage of the chart is that the rate of changes is not easily visualised' [12].

To be useful in critical care situations, where serial analyses and interpretations of arterial blood gas values are crucial, an acid-base chart should represent all three acid-base variables in such a way that changes in any direction can be equally appreciated. Ideally, a graphical representation of three acid-base variables allowing the clinician to monitor the acid-base status over time would require the use of three-dimensional graphical software with options of on-line rotation. It is unlikely that a clinician will use such software in a clinical setting.

In this chapter, a standardised way is proposed, based on the data reduction method introduced in the preceding chapter, to represent pH, PaCO_2 and $a[\text{HCO}_3^-]/\text{BE}$ val-

ues in a two-dimensional chart exploiting their intrinsic two-dimensionality. The method allows a faithful graphical representation of all three basic acid-base variables in two dimensions. Moreover, for pH, PaCO_2 and BE values, a new way of classifying acid-base disorders is presented as a solution to the problem of 'unclassifiable' classes that occurs when classifying acid-base observations according to the Astrup and Siggaard-Andersen method as described in Chapter 1.

3.2 Methods

3.2.1 Construction of the chart

In the preceding chapter it was demonstrated that a standard mathematical transformation procedure called principal component analysis (PCA) of a trivariate acid-base data set results in a distribution of principal component values PC1 and PC2 that has almost the same information content as the original acid-base data set. Hence, the plane spanned by the PC1 and PC2 axes after PCA ensures a graphical representation of the complete acid-base status in two dimensions without significant loss of information. However, according to the definition of PCA, principal components are always uncorrelated [13]. This means that the projection of the original acid-base axes in the PC1-PC2 subspace is entirely dependent on the observed covariances in the original acid-base data set. As a result, PC1-PC2 subspaces of various PCA transformed acid-base data sets are not comparable. Figure 3-1 illustrates this by showing two PCA transformed trivariate acid-base distributions from the preceding chapters; *AZRbe* and *OLVGbe*. The original acid-base axes are projected onto the plane spanned by PC1 and PC2. It can be clearly seen that the projection of the original acid-base axes is different for the two data sets.

To obtain a standard appearance of the PC1-PC2 subspace, an extra rotation in this plane is added in such a way that the original pH-axis is always presented horizontally

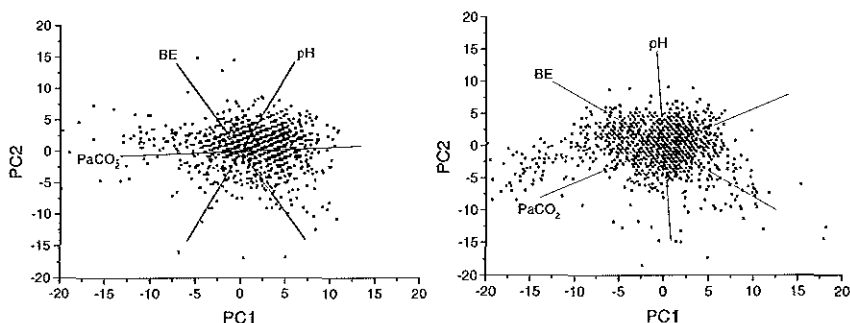


Figure 3-1. Original acid-base axes, projected on the PC1-PC2 subspaces for data sets *AZRbe* (left) and *OLVGbe* (right). The placement of the original acid-base axes is different as a result of different covariances in the original distributions.

with low pH values on the left and increasing values on the right. To do so, the eigenmatrix U of Chapter 2 is multiplied by a rotation matrix, yielding a final transformation matrix T ;

$$T = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} * U, \quad (3-1)$$

in which α is the angle between the projected original pH-axis and the PC1-axis line in the PC1-PC2 subspace (see Figure 3-2 for an example). The transformation matrix T calculates rotated PC1 and PC2 values (hereafter referred to as PC1' and PC2') from original standardised acid-base values. PC1' and PC2' values are in fact the co-ordinate values of the acid-base observation in the proposed acid-base chart.

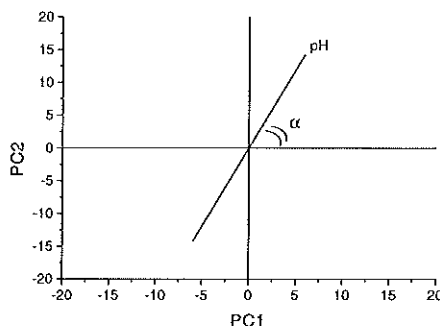


Figure 3-2. The angle α between the projected original pH-axis and the PC1-axis in the PC1-PC2 subspace of data set AZRbe.

Having found the final transformation matrix T , the outlines of the acid-base chart can now be drawn. First, original acid-base axes are drawn into the chart as follows. Since the acid-base data set is standardised (see Equation 2-1), each acid-base axis in the original coordinate system can be represented by a unit vector. For example, the unit vector representing the positive PaCO_2 -axis (values ≥ 40 mmHg) is $[010]$ since this vector points into the direction of equal pH (first element of the unit vector is 0), positive PaCO_2 (second element is 1) and equal BE or $a[\text{HCO}_3^-]$ (third element is 0). Multiplying this vector by matrix T yields PC1' and PC2' values that indicate the exact direction of the unit vector representing the positive PaCO_2 -axis in the proposed chart. This procedure is performed for the three original acid-base axes. Then, the standard 95% univariate reference intervals (see Table 1-3) are drawn into the chart. In three dimensions, the univariate intervals of the three acid-base variables form a

95% univariate reference cube. Figure 3-3 shows the intersection of a PC1-PC2 subspace with such a 95% univariate reference cube. Each side of the reference cube corresponds to a standard 95% lower upper cut-off value or upper cut-off value. Since the cube is represented in a standardised measurement space, and a 95% univariate interval is defined as $m \pm 2s$, each side has a value of -2 (lower cut-off value) or 2 (upper cut-off value). For each of the six intersection points (points 11, 8, 4, 10, 5, 1 in Figure 3-3), two of the three cut-off values are known. For example, point 11 is located on the side representing the 95% lower cut-off value for pH, on the side representing the 95% upper cut-off value for PaCO_2 but between the sides representing the upper- and lower cut-off values for the metabolic parameter. Let x , y and z be the 95% cut-off values for respectively pH, PaCO_2 and the metabolic parameter. Since principal components are orthogonal and the three variables are assumed to be linearly related, the third principal component value (PC3) must be 0 [13];

$$e_{31}x + e_{32}y + e_{33}z = 0, \quad (3-2)$$

where e_{31} , e_{32} and e_{33} are the three elements of the normalised eigenvector of the third principal component as listed in Table 2-3. With this equation, any third component of each of the six intersection points can be calculated. An intersection point must subsequently be multiplied by transformation matrix T to obtain its exact location in the proposed chart. The six transformed intersection points together yield the intersection with the 95% reference cube in the proposed chart.

3.2.2 Vectorial classification system for pH, PaCO_2 and base excess values

As described in Chapter 1, classifying acid-base disorders according to the Astrup and Siggaard-Andersen method is based on the simultaneous evaluation of observed pH, PaCO_2 and BE values. A specific combination of observed values for the three acid-base variables below, above or within their respective 95% reference intervals corresponds to a specific acid-base disorder classification (see Table 1-2). However, the ex-

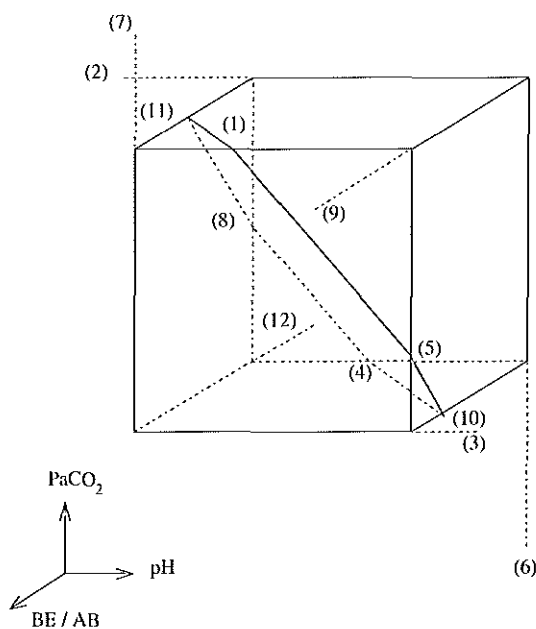


Figure 3-3. Intersection of a PC1-PC2 subspace with a 95% univariate reference cube. Arrows point into the direction of increasing values.

istence of non-classifiable combinations is inherent to a systematic use of these classification rules. As a solution to this problem, Gelsema *et al* [14] proposed to represent each of the 12 acid-base disturbances as used in the Astrup and Siggaard-Andersen classification scheme with a specific disorder vector. For instance, in the standardised measurement space, the unit vector $[-110]$ represents a respiratory acidosis (low pH, high PaCO_2 and a normal BE).

This vectorial classification scheme can be made visible in the proposed chart for pH, PaCO_2 and BE values. Multiplying all 12 unit acid-base disorder vectors by the transformation matrix T yield $\text{PC1}'$ and $\text{PC2}'$ values that indicate the exact directions of the 12 disorder vectors in the proposed chart. Determining all angles that a patient acid-base vector makes with each of the 12 disorder vectors in the proposed chart produces a classification. The acid-base disorder vector yielding the smaller angle is subsequently used for the actual classification of the observed patient acid-base values.

3.3 Results

3.3.1 The proposed chart for pH, PaCO_2 and $a[\text{HCO}_3^-]$ values

Figure 3-4 shows the proposed chart for the *ELIab* data set. The chart is based on the PCA transformation of 1500 pH, logarithmic PaCO_2 and logarithmic $a[\text{HCO}_3^-]$ values coming from patients of the ICU of the St. Elisabeth hospital. The hexagon in the middle is the two-dimensional representation of the standard 95% reference cube of Figure 2-1. Note that the original acid-base axes are no longer perpendicular as seen

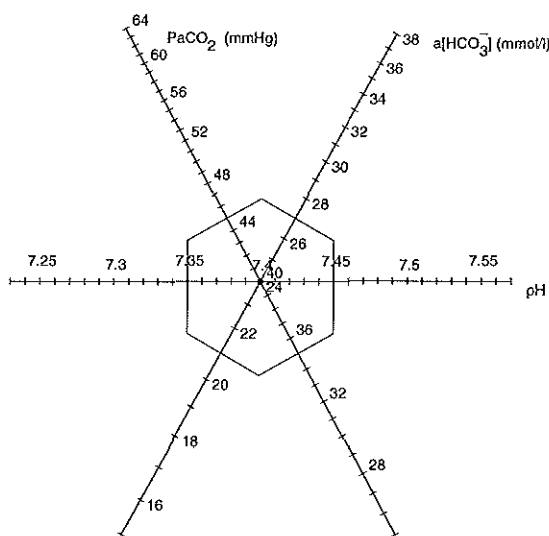


Figure 3-4. The proposed chart for pH, PaCO_2 and $a[\text{HCO}_3^-]$, as built from the data in the *ELIab* data set.

in other familiar acid-base charts. In Figure 3-4, the pH axis runs along the abscissa, while the PaCO_2 and $a[\text{HCO}_3^-]$ axes are placed at an angle close to 60° , resulting in a tri-axial placement of the original acid-base axes.

It may be argued that in this tri-axial configuration, information on the values of the original variables is lost. Indeed, the rotated principal components $\text{PC1}'$ and $\text{PC2}'$ are linear combinations of the original variables. Nevertheless, although the projected original acid-base axes are no longer perpendicular, each condition as represented by a point in the tri-axial

chart may still be interpreted in terms of the values of each of the original variables, by (mentally) projecting the point perpendicularly onto the respective axes.

3.3.2 The proposed chart for pH, PaCO_2 and base excess values

Figure 3–5 shows the proposed chart for the *ELIbe* data set. This chart is based on 1500 pH, PaCO_2 and BE values coming from patients of the ICU of the St. Elisabeth hospital. The regions corresponding to the 12 acid-base disorders of the Astrup and Siggaard-Andersen classification method are displayed. Based on this classification scheme, four types of regions can be seen in the chart:

- All variables within their respective standard reference ranges. This is the standard 95% reference volume. In the chart, this volume appears as the hexagonal shaped figure at the centre of the chart.
- One variable outside its 95% univariate reference interval. These regions appear as triangles sharing one side with the standard normal region. There are six such regions, since each of the three variables may be outside its reference interval either at the high or at the low end. These regions are ‘non-classifiable’ in the classification method of Astrup and Siggaard-Andersen.
- Two variables outside their 95% univariate reference intervals. These appear as re-

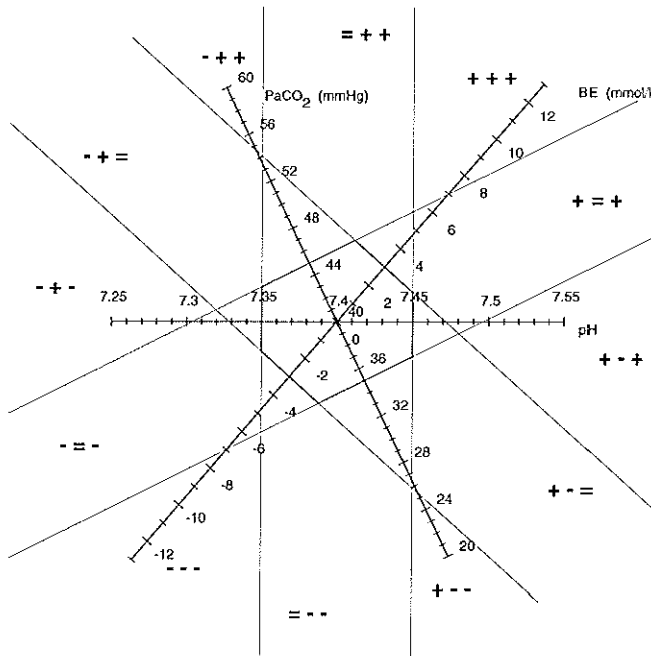


Figure 3–5. The proposed chart for pH, PaCO_2 and BE as built from the data in the *ELIbe* data set. The classical acid-base classification from Table 1–2 are plotted in (dotted lines).

gions, each sharing only one point with the border of the standard normal region. There are six such regions, each touching a different point of the normal hexagon. The region '= + +' is such a region; it contains cases for which pH is within reference, PaCO_2 and BE are outside, both at the high end. Depending on the history and status of the ICU patient, this region indicates either a fully compensated respiratory acidosis, a fully compensated metabolic alkalosis or a combined respiratory acidosis and metabolic alkalosis.

- All three variables are outside their 95% univariate reference interval. These appear as triangular regions with one side of the triangle in infinity and with no point in common with the border of the normal reference region. Again, there are six such regions. The region '+ + +' is one of these. It contains cases with all three variables outside their reference intervals at the high end: partly compensated metabolic alkalosis.

In Figure 3-6, each of the 12 acid-base disorder regions of Figure 3-5 is represented by a corresponding vector. The 12 vectors enable the classification of acid-base disorders according to the method of Astrup and Siggaard-Andersen without the disadvantage of having an *unclassifiable* class. There are two types of vectors:

- Vectors that are (more or less) perpendicular to an original acid-base axis. There are six such vectors and they correspond to the regions in Figure 3-5 that share only one point with the border of the standard normal region. Vector '- + =' is

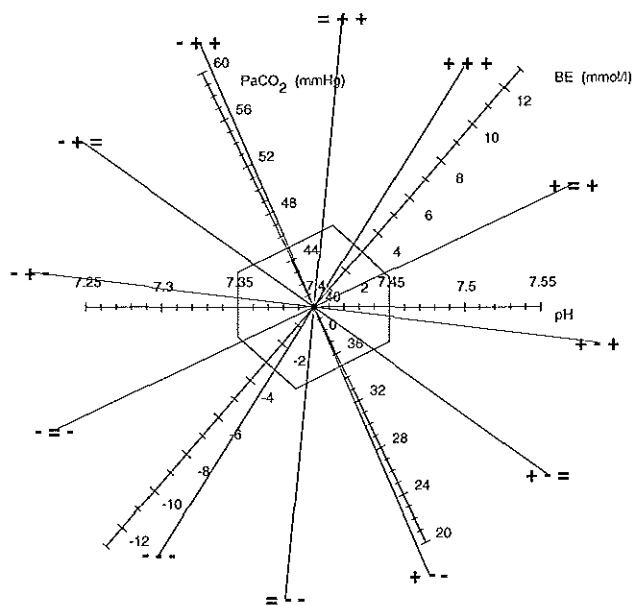


Figure 3-6. The ELIbe chart with the new vector classification scheme plotted in (dotted lines).

such a vector; it is almost perpendicular to the original BE axis and points towards low pH and high PaCO_2 values. Hence, this vector represents a pure respiratory acidosis.

- Vectors that are between the above mentioned six vectors. They represent the triangular regions of Figure 3–5 that have one side of the triangle in infinity and with no point in common with the border of the hexagon. There are six of these vectors. Vector ‘- + +’ is such a vector. This vector points into the direction of low pH, high PaCO_2 , and high BE: a partially compensated respiratory acidosis.

3.4 Discussion

Intensive care medicine of today is characterised by a voluminous production of measured and calculated variables. Continuous monitoring of vital signs, laboratory values and ventilator settings generates a vast and never ending stream of data to be evaluated and interpreted by ICU personnel. Since there is such an overload of information on the ICU there is a great need to develop decision aids to enable ICU personnel to become quickly orientated in the pathophysiological state of the patients under their care [15].

The traditional method of communicating medical quantitative information is the numerical display of data. However, a number of studies show that the graphical representation of medical data greatly improves human information processing. For instance, Cole found that the graphical representation of ventilator data led to a faster interpretation of these data as compared to the traditional numerical display, while Elting demonstrated that the assimilation of information was significantly faster and more accurate when time-dependent information variables are displayed graphically rather than in a tabular manner [16, 17]. In simulated sessions of anaesthesiology monitoring, Gurushanthaia demonstrated a significant improvement in accuracy and speed in detecting changes in the values of physiologic variables when a graphical display of the data was used rather than a textual one [18].

Numerical interfaces make the data, required to complete tasks of diagnosing and classifying, available but often do not provide the *information* necessary to support a physician in his decision making [19]. The advantage of graphical displays over numerical representations probably lies in the exploitation of the still unsurpassed pattern recognition capabilities of the human mind. Regular use of graphical displays by physicians may induce a learning effect so that specific pathophysiological states of a patient are recognised in a single glance. Because it is difficult for the clinician to memorise all the data, especially the display of previous data together with the current data is very useful in an intensive care setting [20]. Therefore, plotting consecutive data in charts and time trend plots may be of high value in such a setting.

In this chapter, a two-dimensional chart is described for the representation of trivariate acid-base data sets. The chart is useful for the graphical monitoring of the acid-base status of a patient, since acid-base changes between consecutive observations are faithfully displayed. The rationale of the chart is based on multivariate statistical principles.

Only after the development of this chart it was found that as early as in 1931, Hastings and Steinhaus proposed a chart that has a striking similarity to the one developed here [21]. This kind of acid-base charting was mainly used in a research environment since analysing the acid-base status of arterial blood was far from a routine process around 1930. Only a limited number of studies referred to the original publication of Hastings and Steinhaus [22-24]. The Hastings and Steinhaus acid-base chart gradually sank into oblivion, despite some attempts to revive it [25-27]. The similarities of the Hastings and Steinhaus chart with the chart described in this chapter and the advantages of a tri-axially configured acid-base chart will be further discussed in Chapter 6.

3.5 References

1. Siggaard-Andersen O, Engel K. A new acid-base nomogram. An improved method for the calculation of the relevant blood acid-base data. *Scand J Clin Lab Invest* 1960; 12:177-186.
2. Kaldor G, Rada R. Computerised evaluation of acid-base disorders based on a nine-cell decision matrix. *Med Biol Eng Comput* 1985; 23:269-273.
3. Kintner EP. The A/B ratio. A new approach to acid-base balance. *Am J Clin Pathol* 1967; 47:614-621.
4. Müller-Plathe O. A nomogram for the interpretation of acid-base data. *J Clin Chem Clin Biochem* 1987; 25:795-798.
5. Arbus GS. An in vivo acid-base nomogram for clinical use. *Can Med Assoc J* 1973; 109:291-293.
6. Baron DN. Radial presentation of results of blood acid-base analyses. *Ann Clin Biochem* 1985; 22:359-361.
7. Severinghaus JW, Astrup PB. History of blood gas analysis. Boston: Little, Brown and Company, 1987, Lange BP, ed., *International Anesthesiology Clinics*; vol 25.
8. Van Slyke DD. Some points of acid-base history in physiology and medicine. In: Whipple HE, ed. Current concepts of acid-base measurement. New York: The New York Academy of Sciences, 1964; 1-274; vol 133.
9. Hastings AB. Part I: acid-base measurements *In Vitro*. Introductory remarks. In: Whipple HE, ed. Current concepts of acid-base measurement. New York: The New York Academy of Sciences, 1964; 1-274; vol 133.
10. Davenport HW. The ABC of acid-base chemistry: the elements of physiological blood-gas chemistry for medical students and physicians. 6th Edition, Chicago: University of Chicago Press, 1974.
11. Astrup P, Jørgensen K, Siggaard-Andersen O, et al. The acid-base metabolism, a new approach. *Lancet* 1960;1035-1039.
12. Siggaard-Andersen O. An acid-base chart for arterial blood with normal and pathophysiological reference areas. *Scand J Clin Lab Invest* 1971; 27:239-245.
13. Jolliffe IT. Principal Component Analysis. New York: Springer-Verlag, 1986, *Springer Series in Statistics*; vol 12.
14. Gelsema ES, Leijnse B, Wulkan RW. A multi-dimensional analysis of three chemical quantities in the blood. *Med Inform* 1991; 16:43-54.
15. Mrochen H, Hieronymi U, Meyer M. Physiological profiles and therapeutic goals-graphical aids support quick orientation in intensive care. *Int J Clin Monit Comput* 1991; 8:207-212.
16. Cole WG, Stewart JG. Human performance evaluation of a metaphor graphic display for respiratory data. *Meth Inform Med* 1994; 33:390-396.
17. Elting LS, Bodey GP. Is a picture worth a thousand medical words? A randomized trial of reporting formats for medical research data. *Meth Inform Med* 1991; 30:145-150.
18. Gurushanthaiah K, Weinger MB, Englund CE. Visual display format affects the ability of anesthesiologists to detect acute physiologic changes. A laboratory study employing a clinical display simulator. *Anesthesiology* 1995; 83:1184-1193.
19. Bennett KB, Flach JM. Graphical displays: implications for divided attention, focused attention, and problem solving. *Hum Factors* 1992; 34:513-533.
20. Clemmer TP, Gardner RM. Medical informatics in the intensive care unit: state of the art 1991. *Int J Clin Monit Comput* 1991; 8:237-250.

CHAPTER 3

21. Hastings AB, Steinhaus AH. A new chart for the interpretation of acid-base changes and its application to exercise. *Am J Physiol* 1931; 96:538-540.
22. Cassels DE, Morse M. Arterial blood gases and acid-base balance in normal children. *J Clin Invest* 1953; 32:824-836.
23. Shock NW, Hastings AB. Studies of the acid-base balance of the blood. III. Variation in the acid-base balance of the blood in normal individuals. *J Biol Chem* 1935; 104:585-600.
24. Shock NW, Hastings AB. Studies of the acid-base balance of the blood. IV. Characterization and interpretation of displacement of the acid-base balance. *J Biol Chem* 1935; 112:239-263.
25. Van Kampen EJ. Throwing a curve at laboratory error. *Diagn Med* 1980; March/April:55-61.
26. Van Kampen EJ. A nomogram for the interpretation of acid-base data. *J Clin Chem Clin Biochem* 1988; 26:149-150.
27. Rispen P, Zijlstra WG, Van Kampen EJ. Significance of bicarbonate for the evaluation of non-respiratory disturbances of acid-base balance. *Clin Chim Acta* 1974; 54:335-347.

4

The construction of patient-based bivariate reference models from trivariate arterial acid-base data distributions

Marcel Hekking, Edzard S. Gelsema, Jan Lindemans

Parts are published in: Clinical Biochemistry 1995; 6:581–585

4.1 Introduction

A central question in clinical chemistry is 'Is the observed laboratory value for this person acceptable or not?'. The standard procedure for answering this question is to compare the observed laboratory value with a relevant univariate reference interval for that particular analyte [1-3]. The univariate reference interval has served for years as a standard for the judgement of 'normalcy' of laboratory values [4]. Many different approaches exist to determine a univariate reference interval [5-7]. In general, a reference interval is constructed by taking a random sample from a specific (often healthy) reference population. New observed values are then compared with the values of the sample. Obviously, if the observed value is clearly not within the range of the sample values then one can safely assume the observed value to be abnormal with respect to the chosen reference population. However, with observed values coming closer to the range of values in the sample, there has to be some decision point as whether to consider the observed value to be typical for the reference population or not. This decision point has been arbitrarily set at 5%; when a decision is made one accepts a 5% probability of erroneously classifying an observed value as not typical for the chosen reference population (in hypothesis testing known as a type I error) [8]. It has been shown that, in case the values of the sample follow a normal (or Gaussian) distribution, a parametric reference interval may be constructed by taking 2 standard deviations around the mean of the sample.

With respect to the statistical determination and the application of 95% univariate reference intervals for arterial pH, PaCO_2 , $\text{a}[\text{HCO}_3^-]$ and BE on an intensive care department, the following problems exist.

First, there is no clear consensus about how to define a useful reference population. If a population of healthy persons is taken as a reference population, reference intervals may be too narrow to be of practical use in clinical situations. Moreover, for quantities that are obtained by invasive procedures, the use of healthy volunteers may be ethically unsound [9].

A second problem stems from the intrinsic two-dimensionality of a trivariate arterial acid-base data set. The relationship between pH, logarithmic PaCO_2 and logarithmic $\text{a}[\text{HCO}_3^-]$, as described by the Henderson-Hasselbalch equation, is linear and even the relationship between pH, PaCO_2 and BE is a nearly linear one (see preceding chapters). Accordingly, it is unnecessary and illogical to use reference intervals for all three acid-base variables, since only two of the three can change independently [10].

Thirdly, there are problems concerning the selection of the proper statistical model. The use of the parametric univariate 95% reference interval as the statistical model for evaluating laboratory values is well established in clinical chemistry [2]. However, already in 1969, Schoen and Brooks [11] reported a statistical dilemma resulting from

the use of multiple 95% reference intervals. A person, evaluated with a single 95% reference interval has an *a priori* probability of 5% of being incorrectly diagnosed as abnormal. The same person evaluated with 10 such 95% reference intervals for 10 independent analytes will have an *a priori* probability of 40% ($1 - 0.95^{10} = 0.4$) for being incorrectly diagnosed as abnormal for one or more intervals. Hence, in situations of a simultaneous interpretation of multiple variables, the use of single 95% reference intervals results in an increase in the number of false positive observations by chance alone [12]. One way to deal with this dilemma is to automatically adjust the reference intervals in case of multiple variables in such a way that the probability of falling within all univariate reference intervals remains 95% (Bonferroni adjustment) [13, 14]. Another option is the use of the multivariate reference model [7, 14-19].

The essence of a multivariate reference model is that it treats two or more univariate distributions as one joint multivariate distribution rather than separate distributions. Figure 4-1 displays such a joint distribution for two hypothetical standard univariate

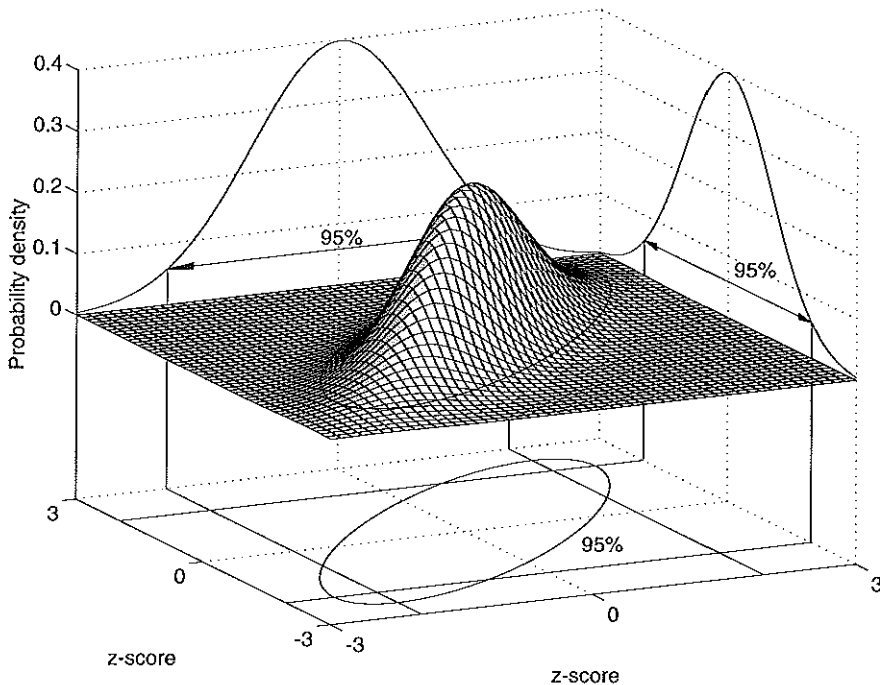


Figure 4-1. The joint distribution of two hypothetical standard normal univariate distributions. The correlation coefficient r between both variables has been set to 0.8. Projections of horizontal cross-sections at the 5% probability density level yield two types of reference regions: the square and the ellipse. Differences between these regions are explained further in Figure 4-2.

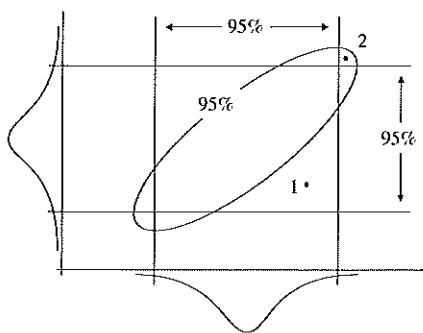


Figure 4-2. Discrepancy between a multiple univariate reference region and a single multivariate reference region. Plotted horizontally and vertically are the probability densities of analyte 1 and 2 of Figure 4-1, respectively. Using the region enclosed by the two univariate reference intervals as a reference region, observations 1 and 2 are a false negative and a false positive observation, respectively when the 95% reference ellipse is taken as the standard [7, 20].

distributions. Both univariate distributions are Gaussian, as illustrated by their bell-shaped curves. The joint distribution of both marginal distributions is the bell-shaped 'mountain' in the middle. Univariate reference intervals are obtained by making horizontal cross-section of the marginal distributions at the 5% probability density level. At the bottom the resultant squared reference region can be found. A cross-section of the multivariate distribution at the 5% probability level, however, results in an elliptical reference region. In analogy to the univariate case, an observation falling outside this ellipse has an *a priori* probability of 5% of being incorrectly diagnosed as abnormal.

The major advantage of using a multivariate reference region as the true refer-

ence region rather than the region enclosed by the two univariate reference intervals is that it takes into account possible correlations between the variables. This is illustrated in Figure 4-2. The rectangle represents the reference region when using both univariate reference intervals. If the ellipse is taken as the standard, false positive and false negative observations may occur. For instance, observation 1 falls within both univariate intervals and is therefore normal for both variables separately. The combination of the two, however, is far from normal due to the positive correlation between the variables. On the other hand, observation 2 is abnormal for both univariate variables but is indeed normal if the correlation between the variables is taken into account. These effects become stronger as the correlation between the variables increases but are present even if the variates involved are uncorrelated [20].

In this chapter, a method for deriving reference models for arterial acid-base data is presented that addresses all of the problems discussed above. A method is proposed that calculates multivariate reference models from laboratory values of intensive care patients themselves. Intensive care patients rather than healthy individuals are therefore the reference population. Moreover, in concordance with the observed linearity between the arterial acid-base parameters, multivariate models are built on the first two rotated principal component values as calculated with the techniques described in the preceding chapters.

4.2 Materials and methods

4.2.1 Patient data

For the analyses, the data sets presented in Chapter 2 were used; *AZRbe*, *OLVGbe*, *OLVGab*, *SKZbe*, *ELIbe* and *ELIab*. For each data set, the transformation matrix T was determined according to the descriptions in the preceding chapter. Each case in the original data set was then transformed with the transformation matrix T yielding new data sets of rotated principal component values $PC1'$, $PC2'$ and $PC3'$. The bivariate distributions of $PC1'$ and $PC2'$ are the input for the calculations presented below.

4.2.2 The multivariate reference model

The multivariate reference model is defined as follows. Assuming a theoretical multivariate Gaussian distribution with a known mean vector and variance-covariance matrix (μ, Σ) the squared Mahalanobis distances between a vector x and the mean μ :

$$d^2 = (x - \mu)^T \Sigma^{-1} (x - \mu) \quad (4-1)$$

are $\chi^2(k)$ -distributed where k is the dimensionality of the multivariate model [21]. The superscript T stands for the transposition of a column-vector to a row-vector, x is the observation vector and Σ^{-1} is the inverse of the variance-covariance matrix. The mean vector μ and the variance-covariance matrix Σ are called the model parameters. The 95% multivariate reference region includes all cases x with a d^2 smaller than or equal to the 0.95 fractile of a $\chi^2(k)$ -distribution [22]. In practice, the model parameters are unknown and replaced by the sample estimates m and S , respectively.

To correct for the uncertainty in the sample estimates of the mean vector m and the variance-covariance matrix S when evaluating single multivariate observations, Chew and Albert proposed to use a 0.95 cut-off fractile that is based on the F -distribution [17, 23];

$$C = k(N^2 - 1) F(0.95; k, N - k) / N(N - k) \quad (4-2)$$

where k is the number of variables in the analysis, N is the number of cases and $F(0.95; k, N - k)$ is the 0.95 fractile of the F -distribution for k and $N - k$ degrees of freedom. In geometrical terms, 0.95 fractiles delimit specific regions in k -dimensional space and are known as 95% equal probability ellipses (for $k = 2$) or 95% equal probability ellipsoids (for $k > 2$). The region delimited by C is also known as the 95% prediction region [17].

4.2.3 Finding the Gaussian distributed core in a multivariate patient data set

In this section, an iterative trimming procedure is described for the determination of a valid acid-base bivariate 95% prediction region from the $PC1'$ and $PC2'$ values of an

ICU patient data distribution. The method assumes that the patient data distribution is composed of two sorts of data; 1) a bivariate Gaussian distributed part of observations at the centre of the distribution (hereafter called the background model distribution) and 2) a contaminating part of outliers in the outer regions of the distribution. The method aims at finding the background model distribution by subsequently removing aberrant observations from the outer regions of the bivariate distribution until the remaining trimmed distribution is found to be bivariate Gaussian.

Under the assumption that the remaining observations of a trimmed bivariate distribution belong to a wider bivariate Gaussian distribution, the standard deviations of the marginal distributions of this wider bivariate Gaussian distribution can be approximated by the standard deviations of the trimmed distributions. However, clearly, these approximations are underestimates. Therefore, prior to the construction of the variance-covariance matrix from estimated marginal standard deviations, a correction must be performed. A table with standard deviation correction factors was constructed empirically beforehand as follows. A standard bivariate Gaussian distribution of 6000 cases was generated from two standard univariate Gaussian distributions (mean = 0 and standard deviation = 1) with the use of the SPSS statistical software package (SPSS for Windows release 6.0, Chicago). The resultant bivariate distribution was then trimmed at specific $\chi^2(2)$ fractiles. At each χ^2 cut-off fractile, the correction factor was determined by dividing the standard deviation of the original marginal distribution by the standard deviation of the trimmed distribution. In Figure 4-3, correction factors (CF) are plotted for 24 fixed χ^2 cut-off fractiles.

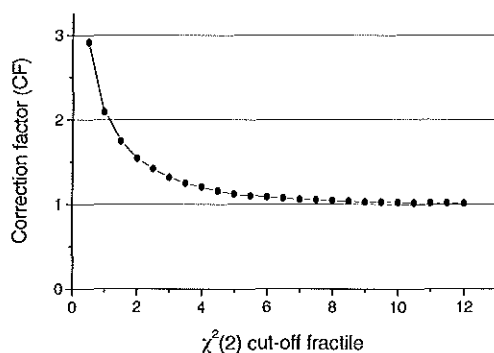


Figure 4-3. The standard deviation correction factor (CF) at 24 specific χ^2 cut-off fractiles for a bivariate Gaussian distribution.

Trimming the bivariate distribution now proceeds as follows (see also Figure 4-4).

- 1) In the first iteration on the untrimmed bivariate distribution, no cases are removed and the correction factor (CF) is therefore set to 1.

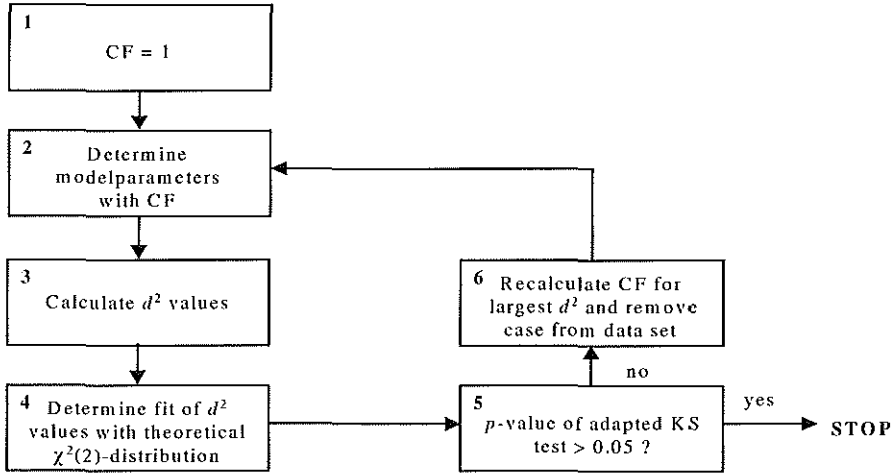


Figure 4–4. Flow-chart of trimming procedure to determine the background model parameters in a bivariate distribution of patient data. CF is the correction factor for the estimated standard deviations and KS stands for the size-adjusted Kolmogorov-Smirnov test. See text for further explanation.

- 2) The model parameters (mean vector and variance-covariance matrix) are constructed from the estimated means, corrected standard deviations and estimated correlation coefficient.
- 3) With the estimated model parameters, a d^2 value (Equation 4–1) is calculated for each case in the data set.
- 4) The goodness-of-fit of the observed cumulative probability distribution of d^2 values with the theoretical cumulative $\chi^2(2)$ probability distribution is now determined. In a bivariate Gaussian distribution, the d^2 values are distributed according to a χ^2 -distribution with 2 degrees of freedom [17, 18, 24, 25]. This observation is used to verify the bivariate Gaussian assumption. With a 1-dimensional goodness-of-fit test like the Kolmogorov-Smirnov (KS) test it can be tested whether the calculated d^2 values indeed follow a $\chi^2(2)$ -distribution [24]. The test statistic of the KS test is the largest difference in probability between an observed cumulative probability distribution and a specific theoretical cumulative probability distribution [25]. If the KS test statistic is small enough, the hypothesis of the bivariate distribution being Gaussian is assumed to be verified.

However, the KS test is designed to be used in non-trimmed distributions only and an adaptation of the test was necessary. Under the assumption that the remaining cases of a trimmed bivariate distribution are part of a wider bivariate Gaussian distribution, the number of cases in this wider bivariate Gaussian distribution can be

reconstructed for a given d^2 trimming value. Thus, for a ranked array of d^2 values (d_1, \dots, d_N) of a trimmed distribution, the adapted KS test statistic D_{\max} is defined as:

$$\begin{aligned} D^+ &= \max(i/(N/Pd_N) - Pd_i), & (i = 1, \dots, N), \\ D^- &= \max(Pd_i - (i-1)/(N/Pd_N)), & (i = 1, \dots, N), \\ D_{\max} &= \max(D^+, D^-), \end{aligned} \quad (4-3)$$

where Pd_i is the theoretical cumulative $\chi^2(2)$ probability for d_i , Pd_N is the cumulative theoretical $\chi^2(2)$ probability for d_N and consequently N/Pd_N is the estimated number of cases in the wider untrimmed bivariate Gaussian distribution. The p -value for a size-adjusted KS test statistic $*D_{\max}$ ($D_{\max} \times (\sqrt{N} - 0.01 + 0.85/\sqrt{N})$) in the range of 0.01-0.15 can be calculated as [26]:

$$p\text{-value} = 6.18 - 17.53 \times *D_{\max} + 16.75 \times *D_{\max}^2 - 5.39 \times *D_{\max}^3 \quad (4-4)$$

- 5) If the adapted KS test yields a p -value larger than $\alpha = 0.05$, there is not enough evidence to reject the H_0 -hypothesis of fit. It may then be concluded that the trimmed distribution is indeed part of a wider bivariate Gaussian distribution and the procedure is stopped. The estimated model parameters (mean vector \mathbf{m} and variance-covariance matrix S as built from the corrected standard deviations and the estimated correlation coefficient) are now the background model parameters and can be used to define bivariate reference regions.
- 6) If the adapted KS test yields a p -value smaller than or equal to $\alpha = 0.05$, the H_0 -hypothesis of fit is rejected and the H_A -hypothesis (lack of fit) is accepted. This means that the trimmed bivariate distribution is not part of a wider bivariate Gaussian distribution and the case with the largest d^2 value is removed from the data set. A correction factor (CF) corresponding to this d^2 value is calculated and a new iteration starts with the trimmed distribution and the new correction factor.

4.3 Results

For each data set, the trimming procedure described in the preceding section succeeded in establishing the background model parameters. Results can be found in Table 4-1. In Figure 4-5, the cumulative probability plot of the d^2 -values of the *ELIbe* data set can be found as an example of the trimming procedure. Note that the remaining cases of the trimmed data set (thick line) follow the theoretical cumulative $\chi^2(2)$ probability distribution (thin line). The corresponding adapted size-adjusted KS test-statistic is 0.855 with an associated p -value of 0.072. Hence, the conclusion can be drawn that the cases of this trimmed data set are part of a wider bivariate Gaussian distribution.

From the means (m), standard deviations (s) and correlation coefficients (r) presented in Table 4-1, the final background model parameters for each data set can be con-

Table 4-1. Distribution characteristics (m is mean, s is standard deviation and r is Pearson's product correlation coefficient between PC1' and PC2') after trimming.

data set	N total	trim med	$*D_{\max}$	m PC1'	m PC2'	s PC1'	s PC2'	r
AZRbe	1500	264	0.891 ($p = 0.050$)	1.93	-1.41	2.46	2.99	-0.20
OLVGbe	1500	151	0.889 ($p = 0.051$)	0.17	0.08	2.84	3.53	0.16
OLVGab	1500	125	0.875 ($p = 0.059$)	0.17	0.22	2.42	4.00	-0.01
SKZbe	1500	139	0.890 ($p = 0.051$)	-4.35	-0.34	3.51	2.94	-0.63
ELIbe	1500	382	0.855 ($p = 0.072$)	0.53	-0.20	3.40	3.39	0.27
ELIab	1500	307	0.882 ($p = 0.055$)	0.31	-0.15	3.16	3.63	0.18

structured. With the background model parameters, 95% and 30% equal probability ellipses were constructed for each data set and displayed in their respective tri-axial charts (Figure 4-6 to Figure 4-11). The differences between the corrected 0.95 and 0.30 fractiles (Equation 4-2) for the *ELIbe* data set (respectively 6.01 and 0.714) and the actual 0.95 and 0.30 $\chi^2(2)$ fractiles (respectively 5.99 and 0.713) are only small. Since the number of trimmed cases in this data set is the largest of all data sets (382 cases), differences between corrected and actual fractiles will be even smaller for other data sets. Therefore, equal probability ellipses for each data set are based on the true $\chi^2(2)$ fractiles rather than on their corrections.

Compared to the standard univariate reference region (the hexagon shaped figure) the

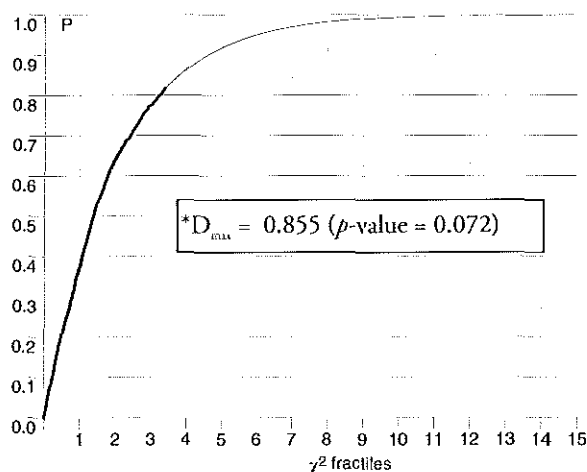


Figure 4-5. Fit of the observed cumulative probabilities of the remaining d^2 values of the trimmed *ELIbe* data set (thick line) with the theoretical cumulative $\chi^2(2)$ probability distribution (thin line). P is the cumulative probability. $*D_{\max}$ is the test statistic of the adapted and size-adjusted Kolmogorov-Smirnov test, indicating the degree of fit of the d^2 values with the theoretical distribution. In this case there is not enough evidence ($p > 0.05$) to reject the H_0 -hypothesis of fit.

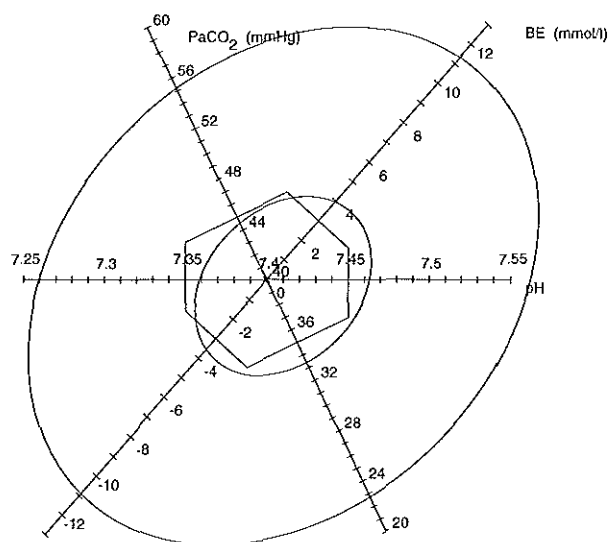


Figure 4-6. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the *ELIbe* data set

95% equal probability ellipses (outer ellipses) are relatively large. The location and orientation of the ellipses for the data sets *ELIbe* (Figure 4-6) and *OLVGbe* (Figure 4-7) are comparable. Their 30% equal probability ellipses both cover almost the entire standard univariate reference region. Both data sets come from a general ICU of a non-academic hospital. However, the ellipses of the *AZRbe* data set (Figure 4-8) are shifted towards a region with higher pH values and lower PaCO_2 values. The correlation coefficient r between $\text{PC1}'$ and $\text{PC2}'$ is -0.20 (see Table 4-1). Location and shape of the ellipses of the *AZRbe* data set indicate that a large portion of the acid-base data from this data set is indicative for a respiratory alkalosis. The *AZRbe* data set comes from the respiratory ICU of the academic hospital Dijkzigt. In Figure 4-9, the ellipses of the neonatal ICU of the Sophia Children's hospital can be found. With respect to the hexagon in this figure and other data sets, the ellipses are shifted towards lower pH values, higher PaCO_2 values and lower BE values. This area is associated with combined respiratory and metabolic acidoses. Moreover, most of the variation in the data set is caused by variation in PaCO_2 values rather than BE values. This can be easily appreciated by (mentally) projecting the ellipses onto the PaCO_2 and BE axis, respectively. The projected ellipses would cover a much larger area on the PaCO_2 axis than on the BE axis.

For the models based on the actual bicarbonate concentration (Figure 4-10 and Figure 4-11), it can be seen that the equal probability ellipses are also very wide but centred around the origin of the chart.

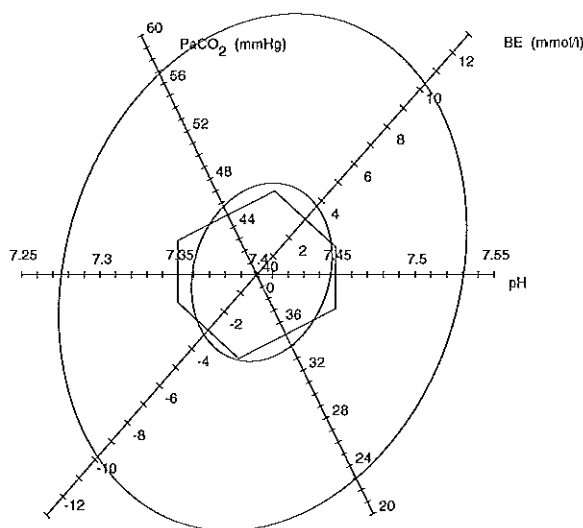


Figure 4-7. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the OLVGe data set

4.4 Discussion

In this chapter, an iterative method is described for the determination of bivariate reference models for acid-base variables, based on values coming from intensive care patients themselves. The proposed method closely resembles the iterative method described by Gelsema *et al.* for defining multivariate reference models from patient data [27]. In this method, the background model parameters are estimated by iteratively including an increasing number of observations from the centre of a multivariate distribution and verifying whether the included cases still belong to a multivariate Gaussian distribution. In general, however, there are fewer *aberrant* observations than *normal* observations in an unselected patient distribution and the method of Gelsema *et al.* may therefore be less efficient. Moreover, in the method of Gelsema the verification of the multivariate Gaussian assumption consists of a visual inspection of graphical output. In the method proposed in this chapter, the 1-dimensional Kolmogorov-Smirnov goodness-of-fit test is used to statistically verify the underlying multivariate Gaussian assumption. This makes the method suitable for an objective and fully automated process. Although the method was designed for the analysis of acid-base data, it can be used to define multivariate reference models for any combination of laboratory data or other measurements.

In clinical chemistry, the derivation of valid reference models from unselected patient data sets has always been attractive. No special sampling procedures are necessary, since routine daily measurements from the clinical chemical laboratory can be used.

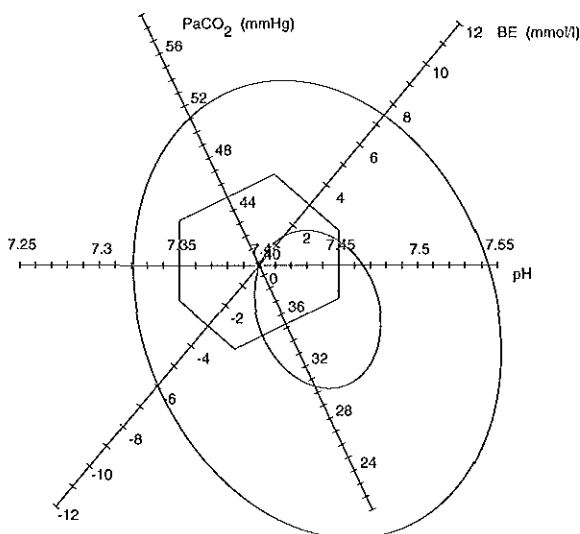


Figure 4-8. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the AZRbe data set.

No extra costs for performing laboratory tests are involved and there is a direct relation of the reference model with the target population [28].

In his thesis, Naus [12] compared 5 different methods to define valid reference intervals from unselected univariate patient data distributions: the method of Hoffman, Neuman, Pryce, Bechtel and Bhattacharya. It was shown that the Bhattacharya method was superior in terms of ease of use and reliability. The essence of the Bhattacharya method is that it determines the Gaussian component (if present) in an unselected univariate reference distribution of patient data. The method starts by dividing a frequency distribution into a number of equally spaced classes. If a Gaussian component is present, plotting the logarithm of the ratio of the frequencies in two subsequent classes against the midpoint of the first class results in a straight line ($y = ax + b$) somewhere in the graph. The estimated mean (m) and variance (s^2) of the Gaussian component are then calculated respectively as $-b/a + 0.5h$ and $-h/a - h^2/12$ (with h being the width of classes) [29]. For a detailed description of the Bhattacharya method the reader is referred to the thesis of Naus.

Baadenhuisen and Smit [28] used a modified Bhattacharya procedure to determine reference intervals for univariate unselected and skewed distributions. Oosterhuis [29] compared the Bhattacharya method for defining univariate reference intervals with the method proposed by the IFCC, which involves the evaluation of data from blood donor populations. Naus *et al.* [30] used the Bhattacharya method for the determination of reference intervals for a number of haematological parameters.

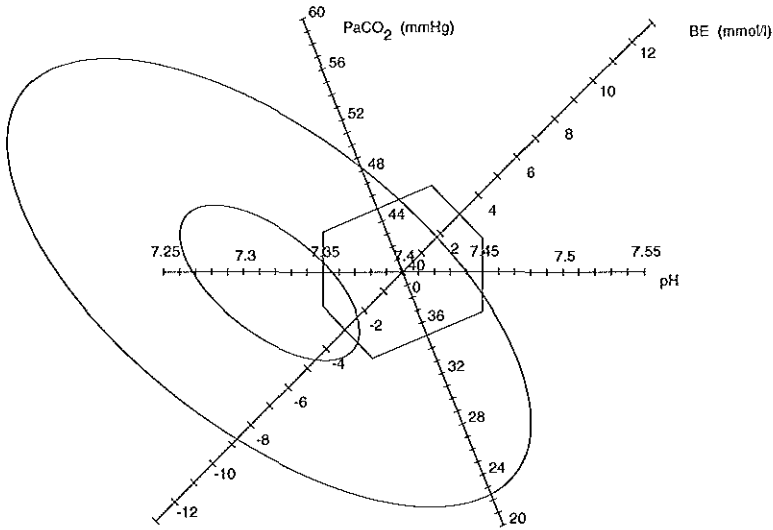


Figure 4-9. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the SKZbe data set.

Naus also used the Bhattacharya method for the determination of a normal bivariate reference region from patient data for the combination of total protein and albumin [31]. His approach starts with calculating the means and variances for the two marginal distributions using the Bhattacharya method. Then, the covariance between both variables (hereafter called x and y) are determined as $(s^2_{sum} - s^2_{dif}) / 4$. The term s^2_{sum} is the variance of the distribution of the sums of x and y , determined with the Bhattacharya method. The term s^2_{dif} is the variance of the distribution of the differences of x and y , determined with the Bhattacharya method. The resulting variance-covariance matrix (constructed with the aid of the Bhattacharya method) and the mean values are then used as the background model parameters to calculate bivariate reference regions.

Major shortcomings of Naus' method for defining multivariate normal reference regions from patient data are: 1) it cannot be used in situations where the Bhattacharya method fails to detect a Gaussian component in one or more marginal distributions; 2) the marginal distributions being Gaussian does not automatically imply that the joint distribution is also Gaussian; 3) computation time increases dramatically when the number of included variables increases; 4) the Bhattacharya method generally requires a substantial number of cases to be reliable [12].

For the iterative method described in this chapter, the prerequisite of the marginal distributions being Gaussian no longer holds, since the trimming of aberrant observations and the verification of the multivariate Gaussian assumption is performed directly on the joint distribution rather than on the marginal distributions. Further-

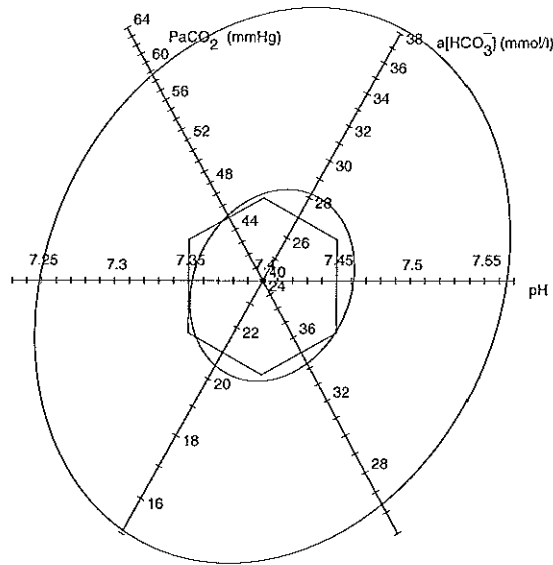


Figure 4-10. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the ELlab data set.

more, the proposed method can automatically and straightforwardly be applied to both the univariate and the multivariate case. A special computer program for defining and testing multivariate normal reference models derived from patient data sets as described in this chapter can be found in [32].

In the discussion about the usefulness of the proposed bivariate reference models for arterial acid-base data in an intensive care setting, two central questions arise. First, what is the clinical value of reference models that are based on patient data rather than on healthy reference populations? Second, what is the clinical value of using a multivariate reference model as compared to using the classical univariate reference intervals?

Ad 1. What is the clinical value of patient-based reference models? The selection of the reference population is considered the most crucial part in the process of building reference models. In general, reference populations consist of ambulant, subjectively healthy students, laboratory staff, blood donors, etcetera. But are these reference models valid? What is the value of *health*-associated reference models for specific groups of patients? Much confusion arises about the fact that health is relative [33]. A patient can be regarded ill in one respect and healthy in another. Laboratory values found in a person at a young age could indicate health but the same values determined on the same patient at an older age could be indicative of disease. Moreover, the diagnosis of health cannot be based on excluding pathology only [7]. If no signs of disease are pre-

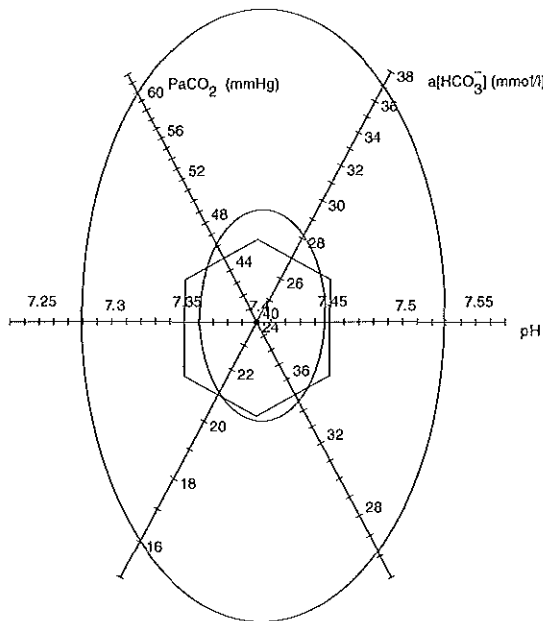


Figure 4-11. The 30% equal probability ellipse (inner ellipse) and the 95% equal probability ellipse (outer ellipse) for the OLVGab data set.

sent, uncertainty remains because these signs could possibly be found on closer examination. As Gräsback defines: *Health is characterised by a minimum of subjective feelings and objective signs of disease, assessed in relation to the social situation of the subject and the purpose of the medical activity, and it is in the absolute sense an unattainable ideal state* [33]. In the light of this definition of health it seems appropriate that reference models should not only be derived from *healthy* people but from a diversity of populations to suit a diversity of purposes.

In order to answer the question whether the patient-based bivariate reference region as defined in this chapter could be useful in an intensive care setting, one should first consider the mechanisms that may shape a reference population. One of those mechanisms is the nature of the ICU from which the patient reference population is taken. Arterial acid-base data from a neonatal ICU will be quite different from acid-base data coming from a respiratory ICU for adults. This is illustrated by the reference ellipses that are based on the neonatal ICU data set (Figure 4-9). The ellipses are shifted towards the area of combined respiratory and metabolic acidoses. This is not the case for the adult respiratory intensive care units. This shift towards the combined acidosis area for the neonatal ICU is not surprising, since the majority of neonatal patients are premature new-borns with an insufficient respiratory and metabolic system, leading to an inadequate clearance of both volatile and non-volatile acids.

Another mechanism that might influence the shape and position of a patient-based reference region is the difference in protocols used on an ICU and the preferences of the clinicians working at that ICU. For instance, from discussions with clinicians at the respiratory ICU of the academic hospital Dijkzigt (data set *AZRbe*) it became clear that there was a tendency to keep ventilated patients in a moderate state of hyperventilation [19]. This explains the shift of the ellipses of the *AZRbe* data set (Figure 4–8) towards the area of respiratory alkaloses.

In conclusion, a multivariate reference region based on an ICU population is not only patient-based but also clinic-dependent. It gives an indication as to which of the patients in one particular ICU are most in need of care.

Ad 2. What is the clinical value of a multivariate reference model compared to the use of classical univariate reference intervals? The Mahalanobis distance (Equation 4–1) enables the monitoring of the original three laboratory values with only one single parameter. Monitoring the acid-base status with this single multivariate index may well be advantageous for ICU personnel because of its ease of interpretation. In a future acid-base monitoring system, one single threshold needs to be set instead of three separate thresholds for all laboratory acid-base variables. Also, using the Mahalanobis distance as a monitoring parameter for arterial acid-base values could substantially reduce the number of false-positive alarms as described in the introduction. It should be noted here that the Mahalanobis distance does not add new concepts to the interpretation of acid-base data. Merely, another statistical tool is presented for the analysis of acid-base data that deals with the fundamental problems associated with the simultaneous analysis of more than one variable. The proposed technique does not deprive a clinician of his usual way of interpreting acid-base data of a patient, but patient outcome could be indirectly influenced because of the possible reduction in false positive alarms and the simplicity of analysing a single parameter, giving the clinician more time for other aspects of patient care.

4.5 References

1. Solberg HE, PetitClerc C. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. *J Clin Chem Clin Biochem* 1988; 26:593-598.
2. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. *J Clin Chem Clin Biochem* 1987; 25:337-342.
3. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values. Approved recommendation (1988) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *J Clin Chem Clin Biochem* 1987; 25:645-656.
4. Boyd JC. Perspectives on the use of chemometrics in laboratory medicine. *Clin Chem* 1986; 32:1726-1733.
5. Bezemer PD. Referentiewaarden - een verkenning van methoden voor het bepalen van 'normale waarden'. Amsterdam: Vrije Universiteit Amsterdam, 1981; 180 pp.
6. Solberg HE, Gräsback R. Reference values. *Adv Clin Chem* 1994; 27:1 -79.
7. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. Tietz textbook of clinical chemistry. 2nd Edition, W.B. Saunders Company, 1994; 454-484.
8. Dixon WJ, Massey FJ. Introduction to statistical analysis. Student Edition, Singapore: McGraw-Hill, Inc., 1983; 1-678.
9. Gelsema ES, Leijnse B, Wulkan RW. A multi-dimensional analysis of three chemical quantities in the blood. *Med Inform* 1991; 16:43-54.
10. Madias NE, Adroqué HJ, Horowitz GL, et al. A redefinition of normal acid-base equilibrium in man: Carbon dioxide tension as a key determinant of normal plasma bicarbonate concentration. *Kidney Int* 1979; 16:612-618.
11. Schoen I, Brooks SH. Judgment based on 95% confidence limits. *Statistical Considerations* 1969; 53:190-193.
12. Naus AJ. De berekening van referentiewaarden in de klinische chemie uit analyseresultaten van een patientenpopulatie. Maastricht: Rijksuniversiteit Maastricht, 1982; 154 pp.
13. Slotnick HB, Etzell P. Multivariate interpretation of laboratory tests used in monitoring patients. *Clin Chem* 1990; 36:748-751.
14. Winkel P. Multivariate analysis and expert systems. *Scand J Clin Lab Invest* 1994; 219:12-24.
15. Winkel P, Lyngbye J, Jörgensen K. The normal region - a multivariate problem. *Scand J Clin Lab Invest* 1972; 30:339-344.
16. Grams RR, Johnson EA, Benson ES. Laboratory data analysis system: section III-multivariate normality. *Am J Clin Pathol* 1972; 85:188-199.
17. Albert A, Harris EK. Multivariate interpretation of clinical laboratory data. New York: Marcel Dekker Inc, 1987, Owen DB, Cornell RG, eds., *STATISTICS: Textbooks and Monographs*; vol 75.
18. Harris EK. Statistical aspects of reference values in clinical pathology. *Prog Clin Pathol* 1981; 8:45-66.
19. Wulkan RW. Expert systems and multivariate analysis in clinical chemistry. Rotterdam: Erasmus University Rotterdam, 1992; 111 pp.

20. Stamhuis IH, Bezemer PD, Kuik D. Evaluation of univariate ranges with a multivariate standard. *J Clin Epidemiol* 1988; 41:359-366.
21. Mahalanobis PC. On the generalized distance in statistics. *Proc Natl Inst Sci India* 1936; 2:49-56.
22. Linner K. Influence of sampling variation and analytical errors on the performance of the multivariate reference region. *Meth Inform Med* 1988; 27:37-42.
23. Chew V. Confidence, prediction, and tolerance regions for the multivariate normal distribution. *J Am Stat Assoc* 1966; 61:605-617.
24. Boyd JC, Lacher DA. The multivariate reference range: an alternative interpretation of multi-test profiles. *Clin Chem* 1982; 28:259-265.
25. Mardia KV. Tests of univariate and multivariate normality. In: Krishnaiah PR, ed. *Handbook of Statistics*. Amsterdam: North-Holland Publishing Co, 1980; 279-320; vol 1: Analyses of Variance.
26. Solberg HE. Statistical treatment of reference values in laboratory medicine: Testing the goodness-of-fit of an observed distribution to the Gaussian distribution. *Scand J Clin Lab Invest Suppl* 1986; 46:125-132.
27. Gelsema ES, Leijnse B, Wulkan RW. Detection of aberrant observations in a background of an unknown multidimensional gaussian distribution. *Meth Inform Med* 1990; 29:236-242.
28. Baadenhuijsen H, Smit JC. Indirect estimation of clinical chemical reference intervals from total hospital patient data: application of a modified Bhattacharya procedure. *J Clin Chem Clin Biochem* 1985; 23:829-839.
29. Oosterhuis WP. Application of statistics in the clinical laboratory with emphasis on multivariate analysis. Leiden: Rijksuniversiteit Leiden, 1994; 104 pp.
30. Naus AJ, Borst A, Kuppens PS. The use of patient data for the calculation of reference values for some haematological parameters. *J Clin Chem Clin Biochem* 1980; 18:621-625.
31. Naus AJ, Borst A, Kuppens PS. Determination of n-dimensional reference ellipsoids using patient data. *J Clin Chem Clin Biochem* 1982; 20:75-80.
32. Hekking M, Lindemans J, Gelsema ES. A computer program for constructing multivariate reference models. *Comput Methods Programs Biomed* 1997; 53:191-200.
33. Gräsbeck R. Reference values, why and how. *Scand J Clin Lab Invest* 1990; 201:45-53.

5

Computational methods and computer program descriptions

Marcel Hekking

Parts are published in: Computer Methods and Programs in Biomedicine 1997; 191-200 and the Proceedings of the Nineteenth Annual Symposium on Computer Applications in Medical Care. 1995; 52-56.

5.1 Introduction

The calculation of patient-based multivariate acid-base reference regions with the methods described in the preceding chapters and the plotting of acid-base trajectories of ICU patients in the tri-axial chart have been implemented in two separate prototype computer programs. These programs are ABTRANS (Acid-Base TRANSformation system) and ABCHART (Acid-Base CHARTing system). Both programs implement the theory and associated equations presented in the preceding chapters. Based on the output of the ABTRANS program, ABCHART can plot acid-base observations of ICU patients in a tri-axial chart, calculate Mahalanobis-distances and make acid-base classifications according to the vector method described in Chapter 3. Typically, for a given ICU the ABTRANS program is used once, while the ABCHART program is subsequently used in a monitoring fashion, while making use of the information generated by the ABTRANS program.

In this chapter, the two prototype computer programs and the associated computational methods will be described in detail. Both computer programs were developed for the Microsoft® Windows™ platform with the use of Microsoft's VISUAL BASIC™ 3.0, Professional Edition (VB). The programs were developed on a 66 MHz 80486 personal computer (PC) with 12 megabytes of internal memory and require Microsoft® Windows™ 3.0 or higher [1]. VB is a rapid application development tool that enables a fast development of event driven programs with a graphical user interface. Special attention has been given to a friendly user-interface for the ABCHART program since this program is to be used by ICU personnel.

5.2 The implementation of numerical routines

The mathematical techniques used in both computer programs require a number of special numerical procedures. Rather than develop these numerical routines within VB from scratch, a numerical routine library (*Numerical Recipes in C. The Art of Scientific Computing*) was used [2]. This library provided the necessary numerical routines in C. With the use of the Borland® C" compiler version 4.5 [3], these routines were compiled into machine language and stored in a dynamic-link-library (DLL). Within VB, the functions and procedures stored in the DLL file were declared according to the standard declaration syntax of VB, making them available in the VB programming environment.

Linking these ready-made compiled numerical routines to the programs had some major advantages: the routines are reliable and highly efficient, no extensive programming or debugging was needed, and routines in compiled DLLs are faster than routines written in VB. Table 5-1 summarises the most relevant numerical routines that are used in the ABTRANS and ABCHART computer programs.

Table 5-1. Relevant numerical routines used in the programs ABTRANS and ABCHART.

name	type	description
invert	NR-procedure	Calculates the inverse of a matrix. Uses the NR-procedures <i>ludcmp</i> and <i>lubkdsb</i> .
moment	NR-procedure	Calculates mean, standard deviation, variance, kurtosis and skewness from an array of input values.
gammp	NR-function	Incomplete gamma function for the calculation of the theoretical cumulative probability for a specific χ^2 -value with a given number of degrees of freedom.
erff	NR-function	Error function for the calculation of a theoretical cumulative probability for a specific Gaussian z-score.
pearsn	NR-procedure	Returns Pearson's correlation coefficient r and corresponding p -value for two arrays of input values.
betai	NR-function	Incomplete beta function for the calculation of a theoretical cumulative probability for a specific F-value with a given number of degrees of freedom.
jacobi	NR-function	Returns the eigenvectors and eigenvalues of a symmetrical matrix.
eigsrt	NR-function	Sorts eigenvectors in the eigenvalue matrix after 'jacobi'.
sKS	VB-function	Calculates the adapted size-adjusted Kolmogorov-Smirnov test-statistic and associated p -value for a theoretical and empirical cumulative probability distribution.

All routines and functions of type NR are from the numerical routines library *Numerical Recipes in C. The Art of Scientific Computing* (Used with permission. Copyright © 1987-1992, Numerical Recipes Software) [2] and are located in the file NUMREC.DLL. The function of type VB is written in Visual Basic.

5.3 The ABTRANS program

ABTRANS is a software tool for:

1. performing a principal component analysis (PCA) of a large data set of acid-base data;
2. determining a valid bivariate reference model on the resulting rotated principal components PC1' and PC2';
3. defining the graphical outlines of the associated tri-axial chart.

Numerical routines used in ABTRANS include: calculating means, standard deviations, correlation coefficients, eigenvalue transformations, inverting variance-covariance matrices and goodness-of-fit testing.

ABTRANS is a single-windowed application that provides the user with a stepwise approach to these functions. At the left hand side of the window, the different steps in

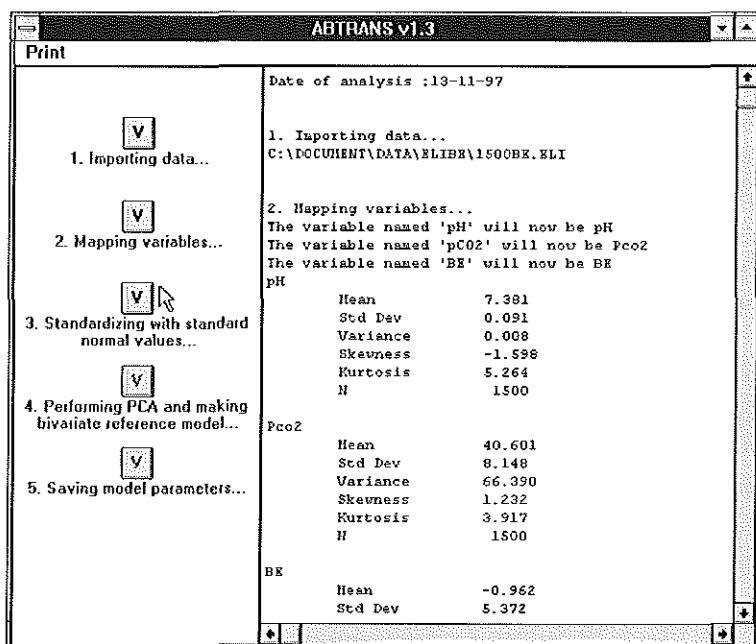


Figure 5-1. Main window of the ABTRANS program to guide a user through the process of defining a multivariate reference model for acid-base data. At the left-hand side of the window, the essential five steps to define a complete model are displayed. Pressing the associated button activates each step. Results of each step are displayed at the right-hand side of the window.

the analysis are displayed. Each step can be activated by pressing the associated button (see Figure 5-1). At the right hand side of the window, the results of each step are displayed. The results can also be sent to a printer. Completing a whole cycle of calculations involves the following five steps:

1. importing an ICU patient acid-base data set;
2. mapping imported variables to fixed internal variables;
3. standardising acid-base values;
4. performing a principal component analysis (PCA), constructing a bivariate reference region on the PC1'-PC2' distributions and defining the graphical outlines of the tri-axial chart;
5. transferring the results of the previous operations to a special Windows initialisation file (ini-file).

5.3.1 Importing an acid-base data set (step 1)

A session starts by importing a tab-separated, comma-separated or space-separated ASCII-file containing the acid-base variables on which the bivariate reference region

and the chart are to be designed. Typically, such a file consists of 500 or more acid-base measurements.

5.3.2 Mapping variables (step 2)

The imported data set may contain other variables than the necessary acid-base variables. At this step the user can indicate which of the imported variables are respectively pH, PaCO_2 , $\text{a}[\text{HCO}_3^-]$ and/or Base Excess (BE). Moreover, a choice can be made of which metabolic variable ($\text{a}[\text{HCO}_3^-]$ or BE) is to be used in the model calculations. Hence, reference models and tri-axial charts can be built on either the BE or the $\text{a}[\text{HCO}_3^-]$ variable. Furthermore, the user can indicate whether to use logarithms of PaCO_2 and $\text{a}[\text{HCO}_3^-]$ or the original measurements. When the acid-base variables are imported and the choices are made, mean and standard deviations are calculated with the routine *moments* (see Table 5-1) and displayed.

5.3.3 Standardising the acid-base variables (step 3)

The next step involves the standardisation of the original acid-base variables with a predefined set of means and standard deviations as described in Chapter 2. The program shows a window with the means and standard deviations that are used for the standardisation already filled in. The user can change these values.

5.3.4 Performing PCA, constructing a bivariate reference region and defining the graphical outlines of the tri-axial chart (step 4)

In linear algebraic terms a principal component analysis (PCA) is an eigenvalue transformation of the variance-covariance matrix of a standardised multivariate data set. For the actual eigenvalue transformation the routine *jacobi* of Table 5-1 is used. The output of this routine is a vector containing the eigenvalues and an eigenmatrix the columns of which contain the normalised eigenvectors (*i.e.* the principal components). The routine *eigsrt* of Table 5-1 is then used to sort the eigenvalues vector in descending order (resulting in the eigenvalue vector ϵ of Chapter 2), while rearranging the columns of the eigenmatrix correspondingly (resulting in the eigenmatrix U of Chapter 2).

The product of U and a special rotation matrix, as described in Chapter 3, yields the final transformation matrix T . Then, for all imported cases, the rotated principal components ($\text{PC1}'$, $\text{PC2}'$ and $\text{PC3}'$) are calculated using transformation matrix T .

The resulting data set of rotated principal components $\text{PC1}'$ and $\text{PC2}'$ are then input for the module that determines the bivariate reference ellipse as described in Chapter 4. For the calculation of the squared Mahalanobis distances, the routine *invert* of Table 5-1 was used for the determination of the inverse of the variance-covariance matrix. For the adapted Kolmogorov-Smirnov goodness-of-fit test, the routine *sKS* of

Table 5–1 was used. The routine *gammp* was used for the calculation of the cumulative probabilities at specific $\chi^2(2)$ -fractiles. For the correction of a 0.95 $\chi^2(2)$ -fractile (Equation 4–2), the routine *betai* was used.

Finally, the projections of the original acid-base axes, the univariate reference hexagon, the Astrup and Siggaard-Andersen regions and the disorder classification vectors in the tri-axial chart are determined.

5.3.5 Saving results into the ABCHART initialisation file (step 5)

In this last step, all calculated results are saved into the ABCHART initialisation file. This initialisation file can be read by the ABCHART program at start-up. The ABCHART program then uses the information stored in the initialisation file to standardise, calculate rotated principal components, squared Mahalanobis distances (Equation 4–1) and to graphically display the tri-axial chart for new acid-base observations entered into the ABCHART program. The following information is stored in the initialisation file:

- means and standard deviations used for the standardisation as defined in step 3;
- the background model parameters (mean vector and variance-covariance matrix);
- elements of the transformation matrix T;
- coordinate values of the unit vectors representing the original acid-base axes;
- coordinate values of the hexagon representing the 95% univariate reference cube;
- coordinate values of the unit vectors of the 12 disorders (see Chapter 3);
- the metabolic variable (BE or $a[\text{HCO}_3^-]$) on which the model is based;
- the unit of measurement for the partial pressures (mmHg or kPa);
- whether or not the logarithm was used for PaCO_2 and $a[\text{HCO}_3^-]$.

Each model can be given a unique name. When using the ABCHART program, these model names can be listed so that a specific model can be chosen.

5.4 The ABCHART program

The ABCHART program is a database program in which patient information and acid-base measurements can be stored. ABCHART uses the Microsoft Access[®] relational database engine of VB to create and use MS-Access[®] 1.1 databases. The ABCHART database consists of three tables. The table *Patient* stores patient related information such as name, initials, date of birth and the patient identification. The table *Lab* holds information on the patient's acid-base analyses such as date and time of analysis, pH, PaCO_2 , $a[\text{HCO}_3^-]$, BE, PaO_2 and oxygen saturation of arterial blood. The table *Model* holds the model name that refers to one of the model names in the initialisation file. Referential integrity between the tables is enforced within the code of the program.

ABCHART can function as a stand-alone program in which data is entered manually. However, when installed on an ICU, a direct link can also be established with a clinical chemistry laboratory for automatic data-entry. This automatic data-entry facility is described in further detail in section 5.4.5. The ABCHART program has four basic windows: start-up window, patient data editing window, laboratory data editing window and the chart window. All four windows will now be briefly described.

5.4.1 Start-up window

At start-up, a window is displayed in which the following actions can be taken:

- creating or opening a specific ABCHART database with patient data;
- choosing a model from the initialisation file;
- setting model options;
- setting general options such as font and font size of the charts;
- switching to automatic data-entry mode.

Models and associated information can be selected and changed in a separate model selection window. In this window, all model names from the initialisation file will be listed and a choice can be made. Also, in this section, the percentile of the reference ellipse to be displayed in the tri-axial chart can be set, as well as the unit in which the partial pressures PaCO_2 and PaO_2 are measured (mmHg or kPa).

5.4.2 Patient data editing window

In the patient selection and editing window, patients can be added to, or deleted from, the database and existing data can be changed. When in automatic data-entry mode, however, no data can be changed or deleted. When installed in an ICU, only one person (typically someone from the central clinical chemistry department) would be authorised to switch from manual to automatic data-entry mode. From a list of patients in the database, a patient can be selected to retrieve and display patient information (see Figure 5–2). Users can determine which patients in the database are actually displayed in the patient list. In this way the list of patients can be limited to those patients currently present on the ICU, without actually removing patients from the database.

5.4.3 Laboratory data editing window

For a selected patient, the laboratory-editing window can be shown (see Figure 5–2). This window shows a list of all laboratory values of the selected patient included in the database at that moment. Laboratory values that can be saved in the database include pH, PaCO_2 , actual bicarbonate concentration ($a[\text{HCO}_3^-]$), Base Excess, PaO_2 and O_2 saturation. Laboratory values can be added or deleted from the database or changed

Acid-Base Charting System v3.0

Patient

Green-2,124,435

Laboratory values

ID: 46,299 G

Date (dd-mm-yy): 27-11-96

Time: 00.15

pH: 7.28

PaCO₂: 38.6 mmHg

HCO₃⁻: 18.3 mmol/l

BE: -8.7 mmol/l

SAT: .93

PaO₂: 114.9 mmHg

Arterial ☒

General information

Patient ID: 2,124,435

Last name: Green

First name: Peter J.J.

Birthdate: 03-04-55

☐ Hide

Patients

Brown, Elizabeth H.

Green, Peter J.J.

Smith, John H.

	Date	ID	pH	PCO ₂	AB
1	27-11-96 00.15	46,299 G	7.28	38.6	18.3
2	27-11-96 03.00	46,301 G	7.275	37.1	17.1
3	27-11-96 06.00	46,302 G	7.245	37.3	15.7
4	27-11-96 10.00	46,295 G	7.259	33.5	14.5
5	27-11-96 15.00	46,312 G	7.245	35.5	14.9
6	27-11-96 22.00	46,322 G	7.228	36.1	14.5
7	27-11-96 23.45	46,323 G	7.219	37.6	14.0
8	28-11-96 05.00	46,313 G	7.214	33.6	15.1
9	28-11-96 14.30	46,339 G	7.266	34.9	15.3
10	28-11-96 06.00	46,330 G	7.275	36.5	16.4
11	29-11-96 14.15	46,337 G	7.305	35.1	17

Figure 5–2. Patient data editing window (foreground) and laboratory data editing window (background) of the ABCHART program.

(only in the manual mode). Also, for each set of laboratory values, the source (arterial or non-arterial) can be indicated.

5.4.4 Charts window

In this window, both the tri-axial chart with the data of a selected patient is displayed (see Figure 5–3) as well as the monitoring plot of the Mahalanobis distance over time. Furthermore, the vectorial classification of the selected acid-base observation is displayed. The user can define the layout of the chart by setting several options. The acid-base trajectory of a selected patient can be followed exactly by advancing and re-treating through the set of acid-base observations pertinent to this patient.

5.4.5 Automatic data-entry facility

An important issue in the acceptance and usability of a computer system in an environment such as an ICU is that manual entry of the data must be kept to a minimum. Therefore, special care has been given to the design and programming of an automatic data-entry facility for the ABCHART program when installed on an ICU.

This facility is based on the daily reporting routine of the laboratories in the OLVG hospital in Amsterdam and the St. Elisabeth hospital in Tilburg. In these hospitals, an arterial blood sample is sent from the intensive care unit to a central clinical chemistry laboratory in another part of the hospital, where the actual blood gas analysis takes

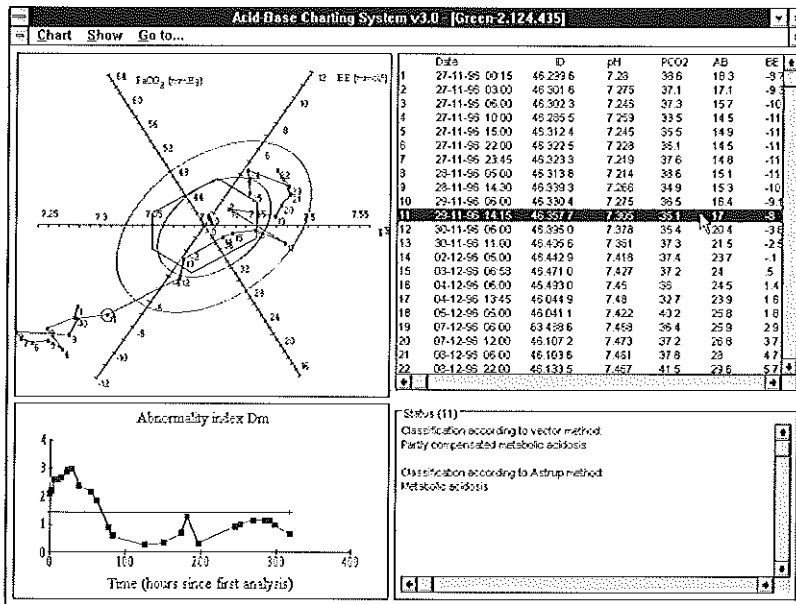


Figure 5-3. Charts and classification window of the ABCHART program. In the upper left corner the tri-axial chart is displayed with an acid-base path of a patient. Acid-base measurements are numbered 1 to 25. In the lower left corner, the time trend of the Mahalanobis distance (here called 'Abnormality index Dm' for clarity of the user) is shown. Note that the horizontal line in this plot represents the 30% equal probability ellipse, displayed as the inner ellipse in the tri-axial chart. Acid-base measurements falling below the line are inside this ellipse, while measurements above the line are located outside the ellipse. In the lower right corner, classifications according the Astrup and Siggaard-Andersen method and the vector method are given for acid-base observation number 11.

place (see Figure 5-4). The results are then sent from the central laboratory back to the intensive care unit via the Hospital Information System on a serial communication line, and printed on a dedicated laser printer residing at the intensive care unit. Typically, these blood gas reports are in plain text (ASCII) format.

For the automatic data-entry facility, a line-sharing device is used that receives a signal and redistributes it to multiple slave devices (DB25 Data Broadcast TL158AE-R3, Black Box Corporation, Utrecht, The Netherlands). At the ICU, the serial line coming from the HIS is plugged into this device while two serial lines coming from the slave ports of the device are redirected to both the ICU report printer and to a serial port of a personal computer installed at the desk. On this PC, a communication program is active that receives all incoming blood gas ASCII reports, extracts the relevant information from the reports, and sends it to the running ABCHART program. The ABCHART program then updates its open database and windows accordingly. This set-up ensures that malfunction of the communication program, the ABCHART pro-

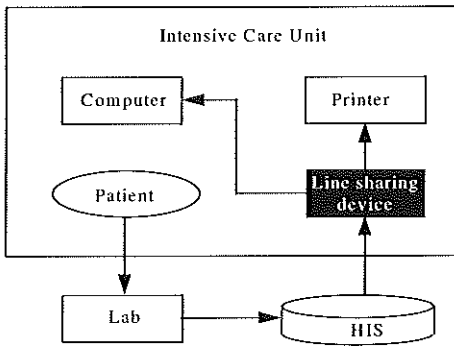
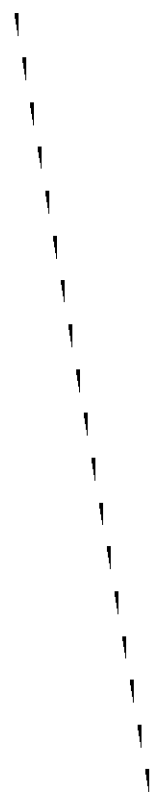


Figure 5-4. Set-up for automatic data-entry of an ABCHART program running at an ICU. An arterial blood sample of an ICU patient is sent to a central clinical chemistry laboratory (Lab) elsewhere in the hospital. Results of the analysis are sent back to the ICU through the Hospital Information System (HIS). At the ICU, the incoming serial signal from the HIS is duplicated by a line-sharing device and sent to both a printer and a personal computer (PC) installed at the ICU. On the PC, a running ABCHART program processes the incoming data and performs an update of its database and windows accordingly.

gram or any hardware problems do not have any effect on the standard reporting procedures at the ICU, since the line-sharing device will always send the ASCII report to the ICU printer. In this way, normal printing procedures in the intensive care department are not interfered with, while patient information is automatically updated in the database of the ABCHART program.

5.5 References

1. Microsoft. Visual Basic. Professional Edition 3.0 ed. Redmont, USA: 1993;
2. Press WH, Teukolsky SA, Vetterling WT, et al. Numerical recipes in C. The art of scientific computing. (2nd Edition,) : Cambridge University Press, 1992.
3. Borland. Borland C++ compiler. 4.5 ed. Scotts Valley, USA: 1994;



A re-appraisal of the tri-axial chart for monitoring
arterial acid-base values — three case reports

Marcel Hekking, Herman J.L.M. Ulenkate, Ben Speelberg,
Marianne J.E. Van Puyenbroek, Henk M.J. Goldschmidt, Edzard S. Gelsema

6.1 Abstract

Objective: To demonstrate the practicability of a tri-axial chart for the graphical and quantitative monitoring of arterial pH, arterial carbon dioxide partial pressure (PaCO_2) and actual arterial bicarbonate-ion concentration ($\text{a}[\text{HCO}_3^-]$) in intensive care patients.

Design: Case reports.

Setting: The general intensive care unit (ICU) of the St. Elisabeth hospital in Tilburg, The Netherlands.

Methods: Using a standard mathematical transformation, a data set of pH, $\log \text{PaCO}_2$ and $\log \text{a}[\text{HCO}_3^-]$ values can be transformed in such a way that a graphical display of all three variables is possible, while being faithful to their linear relationship. Remarkably, this graphical display closely resembles the plot on tri-axial coordinate paper that Hastings and Steinhaus described in 1931 for displaying changes in the acid-base balance during animal experiments. Two new monitoring parameters based on the chart and transformation are described. One of them monitors the abnormality of the acid-base status while the other monitors the rate of acid-base changes.

Conclusions: With the tri-axial acid-base chart, the complete acid-base status can be faithfully monitored. Moreover, the proposed monitoring parameters provide extra information about the arterial acid-base status that otherwise would remain hidden.

6.2 Introduction

The monitoring of laboratory values plays a central role in the intensive care unit (ICU). In general, monitoring involves the comparison of observed values with lower and upper limits for the variable in question. When monitoring the acid-base status using arterial pH, PaCO_2 and actual $\text{a}[\text{HCO}_3^-]$, however, the simultaneous use of lower and upper limits assumes these variables to be independent of each other which may lead to false positive and false negative observations [1]. Moreover, the linear relationship between the variables makes separate monitoring of all three variables illogical [2].

Recently, we developed a model that addresses both problems [3]. The model determines a bivariate reference region on the two dimensions as obtained after a mathematical data reduction of a pH, $\log \text{PaCO}_2$ and $\log \text{a}[\text{HCO}_3^-]$ distribution. After its development it was found that the associated graphical representation closely resembles the plot on tri-axial coordinate paper that Hastings and Steinhaus described in 1931 to study acid-base displacements (see Figure 6-1) [4].

In their coordinate paper, the three original acid-base axes are placed in the same typical tri-axial configuration. Original axes are not shown in Figure 6-1 but they can be reconstructed as follows. The *pH axis* is the hypothetical line through the origin, per-

pendicular to the *Constant pH* line. The CO_2 tension axis is the line through the origin, perpendicular to the *Constant CO_2 tension* line. Likewise, the BHCO_3 axis is the line through the origin, perpendicular to the *Constant BHCO_3* line. The angle between the CO_2 tension axis and the BHCO_3 axis is 60° .

In the present paper, we confirm the tri-axial configuration of the coordinate paper of Hastings and Steinhaus with specific statistical methods. Furthermore, the practicability of the proposed tri-axial chart and model as a visual and quantitative monitoring tool is illustrated with three intensive care patient examples.

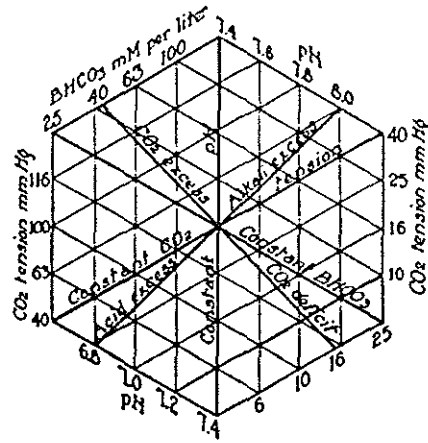


Figure 6-1. The tri-axial coordinate paper of Hastings and Steinhaus for representing pH, CO_2 tensions (PaCO_2) and BHCO_3 ($a[\text{HCO}_3^-]$) values (reprinted with permission of The American Physiological Society [4]).

6.3 Materials and methods

6.3.1 Patient cases

Arterial acid-base measurements of three intensive care patients were recorded. The measurements included arterial pH, PaCO_2 in mmHg, PaO_2 in mmHg and actual $a[\text{HCO}_3^-]$ in mmol/l. Measurements were performed in duplicate on an OMNI-AVL-6 (E-Merck Nederland bv, The Netherlands) and an ABL-520 (Radiometer Nederland, The Netherlands) at the central clinical chemistry laboratory.

6.3.2 Construction of the chart

The tri-axial chart and the multivariate arterial acid-base reference model were constructed from a large patient

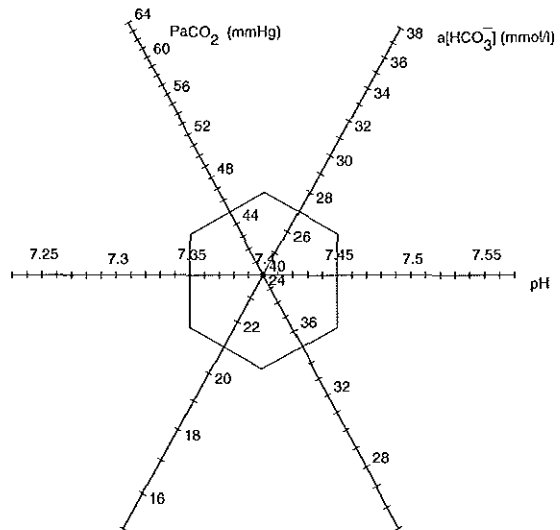


Figure 6-2. The tri-axial chart for the general ICU of the St. Elisabeth hospital.

acid-base data set consisting of 1500 cases (data set *ELIab* from Chapter 2), according to the methods described in the preceding chapters. Figure 6-2 shows the resulting tri-axial chart. The angle between the original PaCO_2 and $\text{a}[\text{HCO}_3^-]$ axes is close to 60° as is the case in the coordinate paper of Hastings and Steinhaus. Note that original acid-base values of a point in the chart can be read from the original acid-base axes by projecting this point perpendicularly onto the respective acid-base axes. The hexagon figure in the chart represents the region defined by the three standard 95% reference intervals for pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ (see Table 1-3).

6.3.3 Monitoring parameter 1: abnormality of acid-base status

In Chapter 4, the rationale and the calculation of squared Mahalanobis distances for multivariate acid-base observations was described. In statistical terms, a squared Mahalanobis distance indicates the relative abnormality of an acid-base combination in reference to a specific distribution. Figure 6-3 shows the location of the reference distribution for the ICU of the St. Elisabeth hospital in its corresponding tri-axial chart. The definition of the squared Mahalanobis distance is such that observations on the same ellipse will have equal squared Mahalanobis distances. The Mahalanobis distance is in fact the distance between an observation and the mean of a multivariate distribution, weighed for possible interrelationships between the variables. Thus, the Mahalanobis distance for a single acid-base observation can be used as a multivariate index of its relative abnormality. By plotting these indices in time, the relative severity of the

acid-base status of an intensive care patient may be monitored.

6.3.4 Monitoring parameter 2: rate of acid-base change

Another monitoring parameter which can be derived from the tri-axial chart of Figure 6-2 is the rate of acid-base change. This is the distance between two points corresponding to two consecutive analyses as plotted in the chart, divided by the time elapsed between the two analyses. Values close to 0 mean slow acid-base changes, an indication for a stable acid-base status.

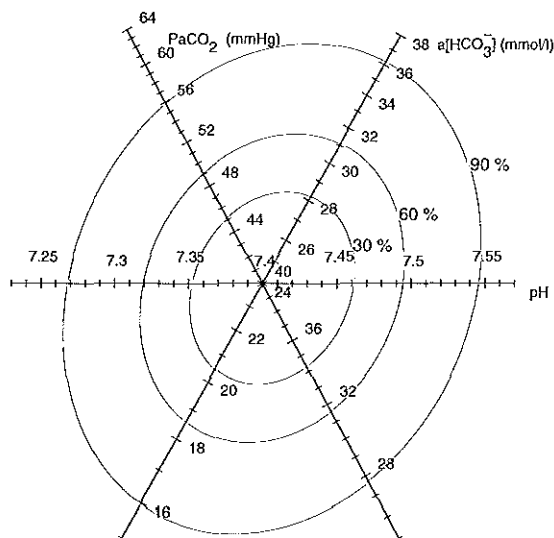


Figure 6-3. The multivariate patient-based reference model for the general ICU of the St. Elisabeth hospital. The 30%, 60% and 90% equal density ellipses each connect points of equal distribution densities.

6.4 Patient reports

In this section, the clinical courses of three patients, as retrospectively reconstructed from their hospital records, of the general ICU of the St. Elisabeth hospital are presented to demonstrate the use and practicability of the tri-axial acid-base chart and the monitoring parameters. Where relevant in these narratives, the user is referred to the charts and trend plots.

The tri-axial charts and trend plots for the three patients are displayed in Figure 6-4 to Figure 6-6, respectively. The upper part of each figure is the tri-axial chart, while the lower part represents the trend plots of both monitoring parameters. Original data can be found in Table 6-1 to Table 6-3. The numbers in brackets in the narratives refer to the numbers in the tri-axial charts, the trend plots and the tables. They indicate the measurement sequence number of the blood gas analyses.

6.4.1 Patient 1 (Figure 6-4, Table 6-1)

Patient 1 is a 67-year-old male, who was hospitalised because of respiratory failure due to pneumonia. On arrival at the ICU, an arterial blood gas analysis showed severe hypoxia and also a combined respiratory alkalosis and metabolic acidosis (1). The hyperventilation (respiratory alkalosis) was triggered by hypoxia and the metabolic acidosis was also due to hypoxia. The patient was intubated and ventilated: volume controlled

Table 6-1. Data for patient 1.

measurement order	hours after first analysis	pH	PaCO ₂ (mmHg)	a[HCO ₃ ⁻] mmol/l	Abnormality index	Rate of acid-base change
1	0	7.424	28	17.9	2.07	
2	3	7.335	38.3	19.9	1.18	2.118
3	6	7.399	31.3	18.9	1.47	1.396
4	12.5	7.367	33.8	18.7	1.31	0.273
5	26	7.241	46.3	19.3	2.51	0.528
6	34.08	7.187	51.4	18.3	3.42	0.352
7	36	7.214	46.5	18.2	2.91	0.995
8	49	7.282	39.5	18.1	1.95	0.289
9	63	7.243	44.6	18.6	2.47	0.180
10	74	7.237	39.4	16.2	2.68	0.269
11	98	7.129	54.7	17.5	4.27	0.287
12	104.27	7.135	51	16.6	4.11	0.226
13	122	7.108	45.4	13.8	4.54	0.205

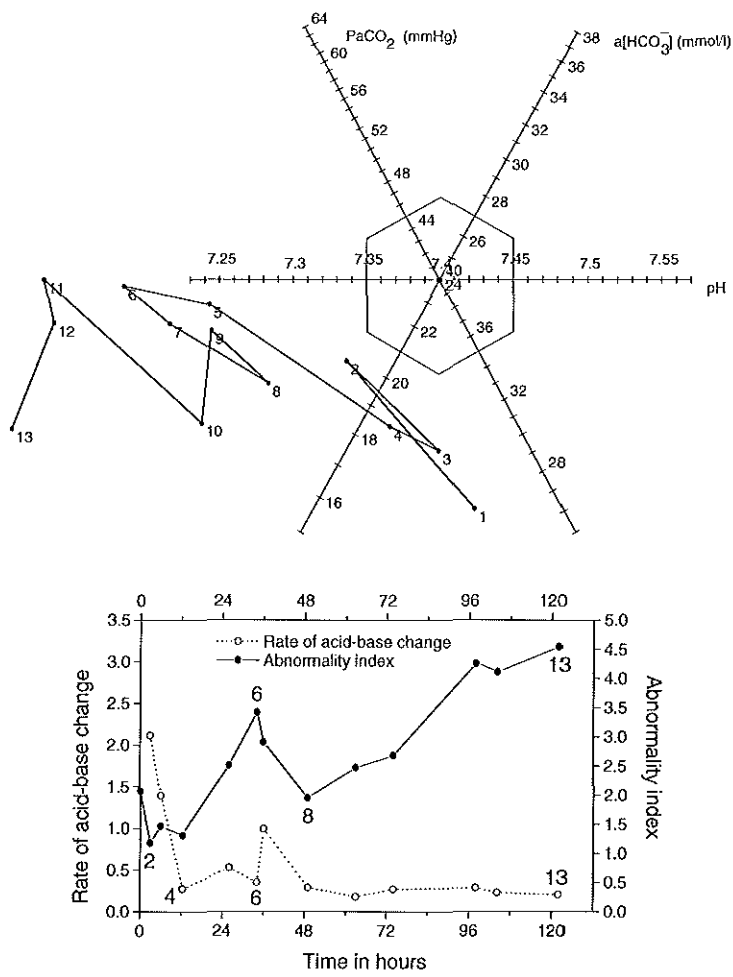


Figure 6-4. The tri-axial chart and the trend plots of the abnormality index (solid line, •, scale on the right) and the rate of acid-base change (dotted line, o, scale on the left) for patient 1.

(VC) with a breathing frequency of 16/min and a tidal volume of 450 ml, FIO_2 0.6 and positive end expiratory pressure (PEEP) of +5 cmH_2O . Three hours after starting artificial ventilation, PaO_2 increased to 102.3 mmHg and the acid-base status (2) showed persistence of metabolic acidosis and less hyperventilation, resulting in a slight uncompensated metabolic acidosis. Three hours later (3), the arterial blood gas analysis showed a respiratory compensated metabolic acidosis. The compensating respiratory alkalosis was due to stimulation of central and peripheral chemoreceptors resulting in agitation of the patient on ventilator support, therefore sedation was increased. Twelve and a half hours after starting the ventilation (4), the hyperventilation diminished and the slight metabolic acidosis persisted.

The next day (5), the patient was quiet and PaCO_2 was 46.3 mmHg, indicative of a combined respiratory and metabolic acidosis. In the afternoon (6), the patient was again agitated and was fighting the ventilator in the volume-controlled mode. The combined respiratory and metabolic acidosis became more serious. After an initial decrease in the abnormality index of 2.07 (1) to 1.18 (2), the acid-base status became increasingly abnormal with an abnormality index of 3.42 (6). The rate of acid-base change decreased from 2.12 (2) to 0.27 (4). It is apparent that the lines in the tri-axial chart from (1) to (6) are all roughly perpendicular to the $a[\text{HCO}_3^-]$ axis, indicating that the change in arterial pH from 7.424 (1) to 7.187 (6) can predominantly be attributed to respiratory problems.

The artificial ventilation was adjusted to the SIMV mode (Synchronised Intermittent Mandatory Ventilation) in combination with Pressure Support. The metabolic acidosis was assumed to be the result of hypotension and hypoxia. Therefore, dopamine was instituted in a dose of 5 $\mu\text{g/kg}$ per minute. The patient was sedated and the VC mode of the ventilator was instituted again. After adjustment of the ventilation, the respiratory acidosis had improved (7) although the metabolic acidosis remained. The next two arterial blood gas analyses (8,9) showed normal PaCO_2 values but the metabolic acidosis remained. As a result of the ventilation adjustments and the dopamine therapy, the acid-base status became slightly more normal. Consequently, the abnormality index decreases from 3.4 (6) to 1.91 (8). Laboratory analysis showed a lactate level of 2.1 mmol/l, while chloride was 114.5 mmol/l, indicating the presence of a hyperchloraemic acidosis. The patient showed a severe uncompensated metabolic acidosis (10). Oxygenation problems developed again, showing a PaO_2 of 57.3 mmHg. The patient was ventilated with an increased inspiration time and increasing levels of PEEP.

Between (10) and (11), bicarbonate therapy was started because of a hyperchloraemic metabolic acidosis. CT scan of liver and thorax as well as liver biopsy showed metastasis of small cell anaplastic lung carcinoma. There were increasing respiratory problems (11) showing respiratory acidosis, as a result of CO_2 overproduction by infusion of bicarbonate, in combination with persistent metabolic acidosis. Institution of higher tidal volumes resulted in acid-base disorders with a small respiratory component and a large metabolic component (12, 13). No further therapy was instituted and the patient died. The trend plots of the monitoring parameters illustrate the gradual deterioration of the patient; the abnormality index increases from 1.91 (8) to 4.54 (13). However, the rates of acid-base change remain constantly low with a small peak at (7), indicating a stable though severely abnormal acid-base status.

6.4.2 Patient 2 (Figure 6-5, Table 6-2)

A 63-year-old female (height: 1.55 m, weight: 51 kg) was admitted to the ICU after resuscitation for cardiac arrest in the Emergency Ward. The patient had diabetes for which no medication was necessary. The medical history mentioned pneumonia and psychosis. On admission, she was hypotensive and had aspirated gastric contents. She was intubated and ventilated in pressure regulated volume-controlled mode (Siemens Servo 300) with an FIO_2 of 1.0, a breathing frequency of 14/min with tidal volumes of 550 ml and +6 cmH₂O of positive end-expiratory pressure (PEEP).

The first blood gas analysis (1) showed a severe uncompensated metabolic acidosis as a result of hypotension (blood pressure 70/40 mmHg). She was treated with dopamine i. v. in a dose of 15 µg/kg per min in combination with noradrenaline in a dose of 0.1 µg/kg per min and intravenous infusions of NaCl 0.9% in order to support circulation. The PaO_2 was 63.7 mmHg and the O_2 saturation was 93%. Metabolic acidosis

Table 6-2. Data for patient 2.

meas- urement order	hours after first analysis	pH	PaCO_2 (mmHg)	$\text{a}[\text{HCO}_3^-]$ mmol/l	Abnor- mality index	Rate of acid-base change
1	0	7.257	38.8	17.3	2.30	
2	1	7.288	36.9	17.1	2.03	1.45
3	3.5	7.254	38	16.8	2.39	0.572
4	5	7.345	31.5	17.1	1.84	3.143
5	6.5	7.445	27	18.8	2.19	3.285
6	8.5	7.356	33.2	18.2	1.48	2.434
7	12.25	7.316	29.6	15.5	2.41	0.932
8	14.93	7.288	33.5	15.6	2.36	0.827
9	16.5	7.257	33.4	14.4	2.84	1.135
10	21.5	7.297	29.6	14.1	2.81	0.505
11	31.5	7.392	31.9	19	1.38	0.620
12	35.5	7.412	28.2	19	1.86	0.500
13	39.5	7.164	44	15.3	3.68	3.162
14	42.5	7.223	51.8	20.7	2.90	1.985
15	47.5	7.33	52.8	27.1	1.89	1.213
16	54.5	7.372	52.2	29.6	1.70	0.314
17	61.5	7.421	46.7	29.8	1.19	0.378
18	79.5	7.451	43	29.4	1.05	0.100
19	107.68	7.433	39.2	25.7	0.44	0.996

(2, 3, 4) remained and respiratory compensation was initiated. An acute respiratory distress syndrome developed (Murray Score of 3.2), probably as a result of aspiration and the PEEP level was increased to +10 cmH₂O.

Arterial blood gas analysis (5) showed a respiratory alkalosis due to hyperventilation and metabolic acidosis. The lactate level was 0.9 mmol/l with an anion gap of 4 mmol/l, indicating that the metabolic acidosis was hyperchloraemic (6). The abnormality index remained relatively high from (1) to (6). Moreover, there was a high rate of acid-base change from (4) to (6), indicating an unstable acid-base status as a result of a progressive deterioration of her circulatory collapse. A thermodilution catheter was introduced in the left subclavian vein. Haemodynamic support was adjusted according to cardiac output and pulmonary capillary wedge pressure measurements, ultimately resulting in a blood pressure of 130/70 mmHg. The patient herself was unable to compensate, resulting in a partially compensated metabolic acidosis (7).

Thereafter, the tidal volume was reduced since the peak inspiratory pressure was 37 cmH₂O. The patient developed a serious metabolic acidosis (8, 9, 10) as a result of diminished respiratory compensation, resulting in an increase of the abnormality index to 2.81 (10). With higher blood pressure, metabolic acidosis recovered in the end (11, 12). The patient was becoming increasingly stable as evidenced by a rate of acid-base change decreasing from 3.29 (5) to 0.5 (12). A severe pure metabolic acidosis, combined with a slight respiratory acidosis (increase in PaCO₂ of 28.2 to 44.0 mmHg), suddenly appeared (13). This change in acid-base status can be clearly seen in both the tri-axial chart as well as in the trend plots of the monitoring parameters.

The lactate level was normal, but there was a hyperchloraemic acidosis (chloride 128.9 mmol/l). A 150-ml sodium-bicarbonate infusion (8.4%) was given at 39.5 hours, just after (13). The hyperchloraemic acidosis was probably due to renal HCO₃⁻ loss or due to renal dysfunction as a result of inadequate renal perfusion. The metabolic acidosis decreased, but there was an accumulation of CO₂, resulting in a combined respiratory and metabolic acidosis (14). Ventilation was adjusted and bicarbonate supplementation was continued. Bicarbonate was discontinued (just after 15), since a pure respiratory acidosis developed and the ventilation was further adjusted.

The administration of bicarbonate leads to CO₂ and H₂O. Therefore, alveolar ventilation has to be increased when infusing bicarbonate, in order to eliminate the newly generated CO₂. In unconscious mechanically ventilated patients, the respiratory settings have to be increased accordingly. However, in this situation tidal volumes could not be adjusted sufficiently since there was already a high peak of inspiratory pressure. Minute volume can be increased by increasing the tidal volumes and, secondarily, the respiratory rate. In this case of severe ARDS there was no benefit from increasing the respiratory rate above 18/min. Finally, the respiratory acidosis was compensated (16) and pH returned to normal. After bronchial toilet, the PaCO₂ further decreased (17).

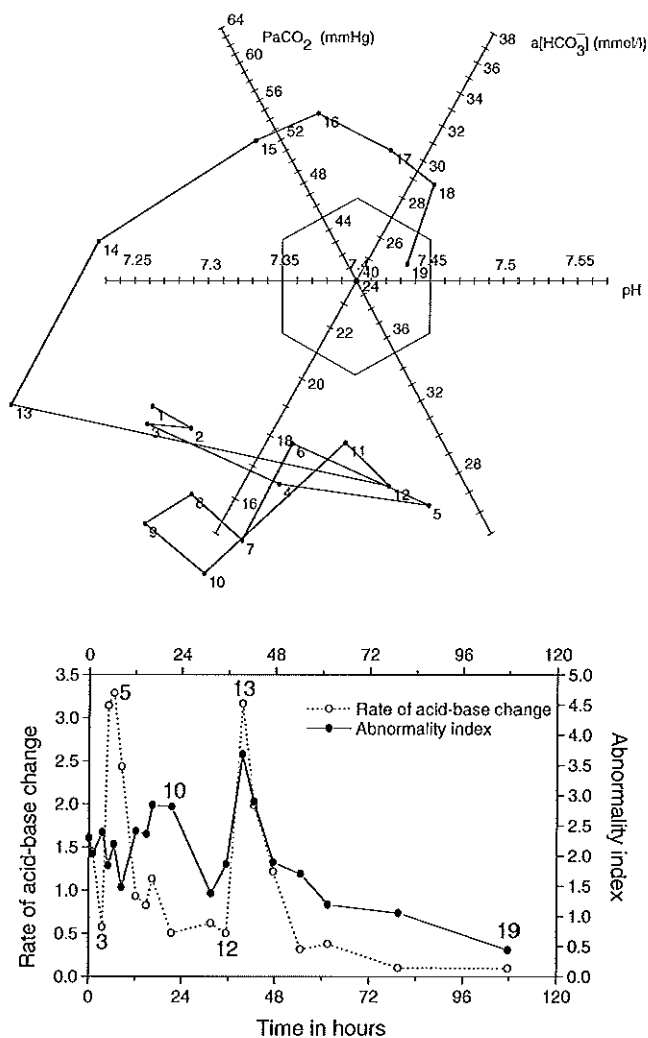


Figure 6-5. The tri-axial chart and the trend plots of the abnormality index and the rate of acid-base change for patient 2.

A slight uncompensated metabolic alkalosis then occurred (18) that may have been the result of dehydration or the massive bicarbonate infusion.

The patient died with normal arterial blood gas values due to renal failure and hypoxic cerebral injury as a result of the cardiac arrest (19). Oxygen delivery was determined several times during hospitalisation and showed values above 700 ml/min. Note that the bicarbonate therapy and the subsequent adjustments to the ventilation after (13) resulted in an acid-base path in the tri-axial that is not the shortest route to the centre of the ICU reference distribution. From (13) to (19) the acid-base status of the patient

gradually became more stable and normal, since the abnormality index decreased from 3.68 (13) to 0.44 (19), while the rate of acid-base change decreased from 3.16 (13) to 0.1 (19).

6.4.3 Patient 3 (Figure 6-6, Table 6-3)

A 50-year-old female was admitted to the ICU because of renal failure and respiratory insufficiency with scleroderma. The patient was ventilated in VC mode with 40% of oxygen, a breathing frequency of 16/min, a tidal volume of 550 ml and + 6 cm H₂O of PEEP. She received a Swan-Ganz catheter and inotropic support. The first arterial blood gas analysis showed a non-severe uncompensated metabolic acidosis due to renal failure (1) that deteriorated (2). Because renal failure was present, bicarbonate was given, resulting in a normal arterial blood gas analysis (3). Later (4), there was a slightly elevated PaCO₂ and tidal volume was increased. The patient developed an uncompensated metabolic acidosis (5) that became progressively severe (6). This can be seen in the trend plot as an increase of the abnormality index from 0.5 (3) to 2.5 (6). Haemodiafiltration (HDF) was applied for 3 hours and resulted in a less severe metabolic acidosis (7, 8). The metabolic acidosis remained (9) and bicarbonate therapy was instituted again. Also, ventilation was increased with higher tidal volumes, resulting in a decreased PaCO₂ (10). Metabolic acidosis persisted and HDF was applied again for

Table 6-3. Data for patient 3.

measurement order	hours after first analysis	pH	PaCO ₂ (mmHg)	a[HCO ₃ ⁻] mmol/l	Abnormality index	Rate of acid-base change
1	0	7.344	38	20.2	1.05	
2	15.75	7.289	38.8	17.6	1.95	0.198
3	48.83	7.375	38.5	22	0.53	0.149
4	62.75	7.351	42.8	23.1	0.87	0.150
5	87.75	7.293	42.4	19.9	1.71	0.133
6	111.75	7.233	42.4	17.3	2.64	0.136
7	124.75	7.32	43.2	21.7	1.31	0.386
8	135.75	7.279	44.9	20.4	1.94	0.183
9	138.75	7.29	41.7	19.5	1.76	0.482
10	140.75	7.326	39.1	20.4	1.20	0.947
11	144.08	7.393	38.7	23.1	0.24	0.998
12	169	7.417	35.6	22.3	0.61	0.069
13	175.72	7.338	46	24	1.22	0.788

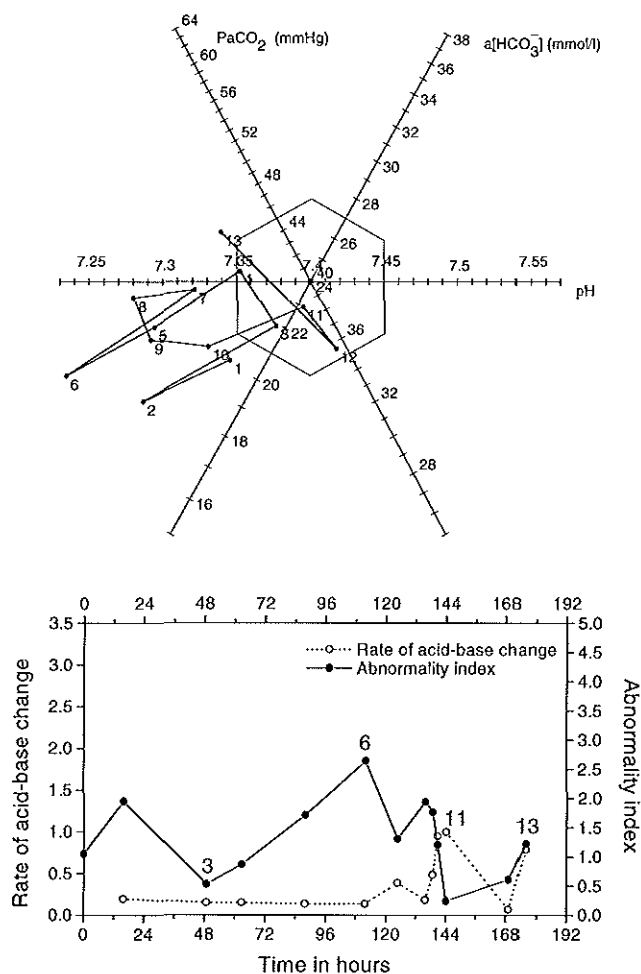


Figure 6-6. The tri-axial chart and the trend plots of the abnormality index and the rate of acid-base change for patient 3.

4 hours. After HDF, the arterial acid-base status returned to normal (11). For 24 hours the arterial blood gas values remained normal (12). This is illustrated by a decrease of the abnormality index from 2.64 (6) to 0.24 (11). Six hours later, a respiratory acidosis occurred (13), possibly the result of air trapping because of high PEEP levels, due to ARDS. Increasing levels of PEEP were necessary with prolonged inspiration time. Ventilation was further adjusted and it was possible to treat the scleroderma and wean the patient from the ventilator.

The rate of acid-base change remained low from (2) to (8). The tri-axial chart shows two spikes; one at (2) and one at (6). The first spike (2) was caused by an uncompensated metabolic acidosis due to renal failure. The acidosis was treated with bicarbonate

therapy resulting in normal acid-base values (3). The second spike (6) shows a recurrent metabolic acidosis. Subsequent HDF resulted in a slight amelioration of the acid-base values (7). From the tri-axial chart it can be deduced that all major acid-base changes are of metabolic origin since the projection of the complete acid-base path onto the PaCO_2 axis shows that all values fall within the standard 95% reference interval for PaCO_2 .

6.5 Discussion

In the present paper, a method for monitoring acid-base values of patients in an intensive care setting is described and demonstrated. The method is based on a tri-axial acid-base chart and two new quantitative acid-base monitoring parameters. One parameter monitors the abnormality of the acid-base status and the other monitors the rate of acid-base changes. The underlying mathematical model justifies the tri-axial coordinate paper of Hastings and Steinhaus [4]. In 1953, Cassels and Morse [5] used this coordinate paper of Hastings and Steinhaus to display acid-base reference values for children. Rispens *et al.* [6] made a case for the coordinate paper in 1974, while in 1980 Van Kampen [7] proposed to use the coordinate paper as a quality control instrument in the clinical chemistry laboratory. However, the Hastings and Steinhaus tri-axial coordinate paper never became popular in daily clinical practice and Hastings reports: 'I still think that ... tri-axial coordinate paper ... is ideal for plotting acid-base paths, but I have never been able to persuade anyone else that this is so, so I've stopped trying to do so.' [8].

The tri-axial coordinate paper of Hastings and Steinhaus was solely used as a graphical tool to visualise acid-base values and acid-base paths. A major advantage of the method presented in the present paper, however, is that the tri-axial chart results from an underlying mathematical model that enables the quantification of the chart. An extra advantage of the method is that values can be displayed against a background of reference data. The method is flexible in the sense that any reference distribution can be used. An ICU can thus build its own reference model and associated chart.

Use of the tri-axial acid-base chart in daily clinical practice may well be advantageous, since it has been shown that a graphical display of data yields a faster and more accurate interpretation as compared to numerical displays [9-11]. Moreover, the tri-axial configuration of the chart ensures that the effects of both respiratory (PaCO_2) and metabolic ($\text{a}[\text{HCO}_3^-]$) components on the acid-base status of arterial blood, are simultaneously and equally assessed. This is more difficult in other acid-base charts [12-15]. For instance, in a plot of PaCO_2 versus pH, the emphasis lies on the respiratory component while in a plot of $\text{a}[\text{HCO}_3^-]$ versus pH (the Davenport diagram), PaCO_2 is missing. In fact, only a tri-axial configuration is faithful to the intrinsic two-

dimensional relationship between all three acid-base values, as dictated by the Henderson-Hasselbalch equation.

The two monitoring parameters (acid-base abnormality and rate of acid-base change) extract information from raw acid-base measurements that is contained within the data but normally remains hidden. The monitoring parameters may well be useful in future acid-base monitoring systems. Single cut-off values could be set so that alarms are generated in cases of abnormal or unstable acid-base situations. They must, however, be seen as an addition to, and not a replacement of, the original acid-base values. Therapeutic actions after an alarm can then still be based on the original values.

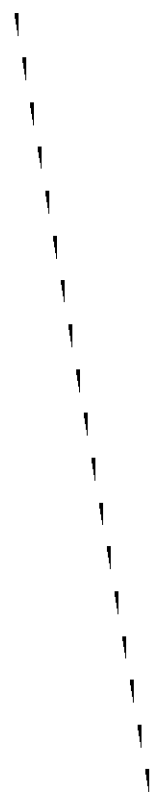
Particularly with the advent of continuous intra-arterial blood gas measurement devices, such a screening system may be very valuable, and even indispensable, in pre-processing the vast amount of data that will be generated by such devices [16].

6.6 Acknowledgement

The authors wish to thank Dr. R. N. M. Weijers for carefully reading the manuscript.

6.7 References

1. Stamhuis IH, Bezemer PD, Kuik D. Evaluation of univariate ranges with a multivariate standard. *J Clin Epidemiol* 1988; 41:359-366.
2. Madias NE, Adrogué HJ, Horowitz GL, et al. A redefinition of normal acid-base equilibrium in man: Carbon dioxide tension as a key determinant of normal plasma bicarbonate concentration. *Kidney Int* 1979; 16:612-618.
3. Hekking M, Lindemans J, Gelsema ES. Design and representation of multivariate patient-based reference regions for arterial pH, PaCO₂ and base excess values. *Clin Biochem* 1995; 28:581-585.
4. Hastings AB, Steinhaus AH. A new chart for the interpretation of acid-base changes and its application to exercise. *Am J Physiol* 1931; 96:538-540.
5. Cassels DE, Morse M. Arterial blood gases and acid-base balance in normal children. *J Clin Invest* 1953; 32:824-836.
6. Rispens P, Zijlstra WG, Van Kampen EJ. Significance of bicarbonate for the evaluation of non-respiratory disturbances of acid-base balance. *Clin Chim Acta* 1974; 54:335-347.
7. Van Kampen EJ. Throwing a curve at laboratory error. *Diagn Med* 1980; March/April:55-61.
8. Hastings AB, Part I: acid-base measurements *In Vitro*. Introductory remarks. In: Whipple HE, ed. Current concepts of acid-base measurement. New York: The New York Academy of Sciences, 1964; 1-274; vol 133.
9. Cole WG, Stewart JG. Human performance evaluation of a metaphor graphic display for respiratory data. *Meth Inform Med* 1994; 33:390-396.
10. Elting LS, Bodey GP. Is a picture worth a thousand medical words? A randomized trial of reporting formats for medical research data. *Meth Inform Med* 1991; 30:145-150.
11. Bennett KB, Flach JM. Graphical displays: implications for divided attention, focused attention, and problem solving. *Hum Factors* 1992; 34:513-533.
12. Cohen ML. A computer program for the interpretation of blood gas analysis. *Comput Biomed Res* 1969; 2:549-557.
13. Goldberg M, Green SB, Moss ML, et al. Computer-based instruction and diagnosis of acid-base disorders. *J Am Med Assoc* 1973; 223:269-275.
14. Bernauer J, Bender HJ, Hartung HJ, et al. Graphic presentation of blood gas data. *Int J Clin Monit Comput* 1984; 1:93-95.
15. Siggaard-Andersen O. An acid-base chart for arterial blood with normal and pathophysiological reference areas. *Scand J Clin Lab Invest* 1971; 27:239-245.
16. Roupie EE, Brochard L, Lemaire FJ. Clinical evaluation of a continuous intra-arterial blood gas system in critically ill patients. *Crit Care Med* 1996; 22:1162-1168.



Predicting hospital mortality of ICU patients from
arterial acid-base measurements

Marcel Hekking, Rob J. Bosman, Durk F. Zandstra,
Rob N.M. Weijers, Edzard S. Gelsema

7.1 Abstract

Objective: to develop and compare new mortality prediction models (MPM) that are based on arterial acid-base variables.

Design: logistic regression analysis of prospectively collected data.

Setting: a 20 bed medical-surgical intensive care unit (ICU) in the 'Onze Lieve Vrouwe Gasthuis', an inner-city teaching hospital in Amsterdam, The Netherlands.

Patients: consecutively admitted patients in the period of January 1997 to May 1998.

Measurements and main results: for 2430 consecutive ICU admissions, hospital mortality was predicted using four MPMs: APACHE II, SAPS II, MPM₀ II, and MPM₂₄ II. Discriminating power was assessed with Receiver-Operator-Characteristics (ROC) curves while calibration was assessed with Hosmer-Lemeshow's C*g values. Although discriminating power was acceptable (area under the curve [AUC] of 0.809, 0.835, 0.810, and 0.746 for APACHE II, SAPS II, MPM₀ II, and MPM₂₄ II, respectively), calibration was poor as demonstrated by large C*g values with *p*-values smaller than 0.001.

Using multiple logistic regression analysis on data coming from 1770 consecutively admitted patients, new MPMs were developed based on each of the three arterial acid-base variables: pH, carbon-dioxide partial pressure (PaCO₂), and base excess (BE). A fourth model was based on the Mahalanobis distance *d* of Chapter 4 (MPM_{*d*}). The single-variable based MPMs showed good discriminating power but calibration was poor. MPM_{*d*} however, showed good discrimination (AUC of 0.863) and excellent calibration (C*g of 8.87, *p*-value is 0.54).

Conclusions: when predicting mortality of ICU patients on the aggregate level, MPM_{*d*} could be useful at our ICU. MPM_{*d*} awaits prospective validation at our ICU and other ICUs.

7.2 Introduction

Predicting hospital mortality for patient groups has become an integral part of critical care practice. Most of the mortality prediction models (MPM) use prognostic indices (PI) that are mathematical transformations of severity scores. With a severity score as input, such models yield a probability of dying in the hospital – a value between zero and one – for the individual patient. By dividing observed hospital mortality by predicted hospital mortality for groups of patients, standardised mortality ratios (SMR) are obtained. SMRs are widely used to compare the performance of intensive care units, either in time or to other units [1].

Most of the common MPMs, such as the Acute Physiology And Chronic Health Evaluation II (APACHE II) [2], the Simplified Acute Physiology Score II (SAPS II) [3], and the Mortality Prediction Model II (MPM II) [4], use physiological measure-

ments. The most deranged values of these variables during the first hours after Intensive Care Unit (ICU) admission are extracted from the patient charts. These MPMs perform adequately at the aggregate level, but lack the calibration and discrimination to be used in the individual patient. They are very sensitive to items like sampling rate, sampling method and artefacts [5, 6]. Moreover, the whole process of extracting, interpreting, and managing the necessary data items is time-consuming and is a source for both inter-observer and intra-observer variability [7].

Haemodynamic variables describe physiologically important abnormalities in critically ill patients and often form the basis of an MPM [8, 9]. The haemodynamic, respiratory and/or metabolic status of the patient is most effectively reflected in arterial acid-base measurements. The objective of the present study is to develop and compare various acid-base based MPMs using the arterial blood gas analyses of the first 24 hours of ICU treatment.

7.3 Materials and methods

Data for the classical MPMs (APACHE II, SAPS II, MPM₀ II, and MPM₂₄ II) was collected over a 17-months period: January 1997 to May 1998. Data for the development of the acid-base based MPMs was collected over a 13-months period; April 1997 to May 1998. If a patient was admitted more than once to the ICU, only the first admission was used. Data for the classical MPMs was routinely extracted for all ICU admissions. Data for the acid-base based MPMs (arterial blood gas measurements) was automatically collected using a microcomputer installed at the ICU with a direct link to the laboratory of Clinical Chemistry and Haematology of the 'Onze Lieve Vrouwe Gasthuis' (see Section 5.4.5 for set-up). Blood gas measurements included: arterial pH, PaCO₂ in mmHg, base excess (BE) in mmol/l, a[HCO₃⁻] in mmol/l, PaO₂ in mmHg, and arterial oxygen saturation. Measurements were performed in duplicate on a Corning 288 (Chiron Diagnostics, The Netherlands) blood gas analyser. Specific approval of the Ethics Committee was not required since data was primarily collected for clinical purposes. No extra blood samples were taken.

For APACHE II, SAPS II and MPM₀ II, the following types of patients were excluded: burn patients, post-operative cardio-surgical patients, and patients with a length of stay (LOS) in the ICU of 8 hours or less. For MPM₂₄ II, patients with a LOS of 24 hours or less were excluded. For the acid-base based MPMs no exclusion criteria were applied.

Mortality predictions according to APACHE II, SAPS II, MPM₀ II and MPM₂₄ II were calculated for individual patients according to the respective guidelines [2-4]. Charting of the necessary variables was performed according to standard procedures at our ICU: the nursing staff entered at least one value for each variable per hour. The attending physician performed the necessary data-extraction (for calculation of

APACHE II, SAPS II and MPM II) 24 hours after ICU admission or on discharge from ICU if discharge was within 24 hours. The physicians had explicit instructions not to record values induced by suctioning, changes of medication pumps, etc. Only sustained, clinically relevant physiological changes were recorded. Only arterial blood gas analyses were used in the data collection. Continuous percutaneous saturation monitoring information was not extracted for use in the scoring systems. Data extraction as described above is routinely performed at our ICU for all ICU admissions (approximately 1900/year).

Acid-base based MPMs were developed using multiple logistic regression analysis. This technique aims to find a linear relationship between *n* predictor variables *x_i* and the logit, which is the natural log of the ratio *p* / (1 - *p*):

$$\text{logit} = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots \beta_nx_n, \tag{7-1}$$

where *p* is the probability of *y* = 1, *y* being a dichotomous (0,1) outcome variable [10-12]. The weights *β_i* are iteratively optimised in such a way

Table 7-1. Coding of age

age	after coding
[0 – 10) ^b	0
[10 – 20)	1
[20 – 30)	2
[30 – 40)	3
[40 – 50)	4
[50 – 60)	5
[60 – 70)	6
[70 – 80)	7
[80 – 90)	8
[90 – 100)	9
[100 – 110] ^c	10

^bfrom and including
^cto and excluding
^dto and including

that the predictor variables explain a maximum of variation in the outcome variable *y*. Having found the optimal weights, the probability of *y* = 1 for an individual case can be determined as: *e^{logit}* / (1 + *e^{logit}*), where *e* is the base of the natural logarithm.

For each single arterial acid-base variable, an MPM was designed using multiple logistics regression analysis: MPM_{pH}, MPM_{PaCO₂} and MPM_{BE}. The outcome variable in each model was hospital death (*y* = 1). Predictor variables included: 1) type of admission, (surgical or non-surgical), 2) age on a continuous scale (see Table 7-1 for scale), and 3) the largest absolute *z*-score *|z|* (*|x - m|/s*) in the first 24 hours after ICU admission as calculated with the means and standard deviations of Table 7-2. Typically, *|z|* runs from zero upwards with larger values indicating a more severe deviation of the acid-base measurement at hand.

Table 7-2. Means (*m*) and standard deviations (*s*) used for calculating absolute *z*-scores.

	<i>m</i>	normal range	<i>s</i>
pH	7.40	7.35 – 7.45	0.025
PaCO ₂	40 mmHg	35 – 45	2.5
BE	0 mmol/l	-3 – 3	1.5

Standard deviations *s* are derived from the normal range by assuming this range to be 4 standard deviations wide.

The absolute z -score $|z|$, however, only indicates the degree of deviation in a single acid-base measurement. We recently developed a method for expressing the deviation in pH, PaCO₂ and BE simultaneously with a single multivariate statistical index: the Mahalanobis distance d (Equation 4-1). This index takes into account the interrelationships between the variables as observed in a multivariate reference distribution of acid-base measurements [13-19]. In this case, the multivariate reference distribution *OLVGbe* was used. Based on the Mahalanobis distance d , we developed a fourth mortality prediction model (MPM_d). Outcome and predictor variables remained the same as in the single-variable MPMs with the exception that the largest $|z|$ was replaced with the largest d as found in the first 24 hours after ICU admission.

According to good modelling practice [20] we performed a 10-fold cross-validation on the acid-base based MPMs. We divided the total data set with n cases into 10 partitions of $n/10$ cases with an equal ratio of non-survivors and survivors. The mortality predictions of the patients in each partition were then calculated using the coefficients as estimated with a multiple logistic regression analysis of the data in the union of the remaining nine partitions. Multiple logistic regression analyses were performed with SPSS for Windows, release 7.5.

To quantify and compare the discriminating power (*i.e.* the ability to discriminate between survivors and non-survivors) of the resulting MPMs, non-parametric Receiver-Operator-Characteristics (ROC) curves were used. These curves were generated by plotting pairs of true-positive ratio (sensitivity) and false-positive ratio (1-specificity) as they appear using each mortality prediction in a data set as a cut-off point between survival and non-survival. The area under the curve (AUC) is a measure of discriminating power and is expressed in the range of 0.5 to 1. An AUC of 0.5 means zero discriminating power while an AUC of 1 indicates an optimal discrimination between survivors and non-survivors. AUCs were calculated according to the method of Hanley and McNeil [21]. Comparisons were made with the method of DeLong for comparing multiple AUCs derived from the same data [22]. DeLong's test statistic is χ^2 -distributed with $k-1$ degrees of freedom (*dff*), where k is the number of AUCs. The hypothesis of equal AUCs was rejected if a p -value was smaller than 0.05.

The goodness-of-fit (*i.e.* the ability to correctly predict the probability of death or survival) of the respective MPMs was assessed with a calibration plot and with the C*_g test statistic as proposed by Hosmer and Lemeshow [10, 23]. A calibration plot is constructed by sorting the mortality predictions in the data set and plotting the actual mortality rate against the predicted mortality rate in adjacent strata of 10% predicted mortality. For the calculation of C*_g, the sorted mortality predictions were grouped into 10 deciles, each containing $\approx n/10$ cases. C*_g was then calculated as:

$$\sum_{k=0}^1 \sum_{l=1}^{10} (O_{kl} - E_{kl})^2 / E_{kl}, \quad (7-2)$$

where O_{kl} is an observed number of persons, E_{kl} is an expected number of persons, k indicates the survivor (0) and non-survivor (1) group, and l the decile. This test statistic is χ^2 -distributed with 10 degrees of freedom [10]. The hypothesis of good fit was rejected if a p -value was smaller than 0.05.

7.4 Results

There were 2430 ICU admissions in the 17-months data acquisition period, 1770 occurred in the 13-months period. Patient characteristics can be found in Table 7-3. Figure 7-1 shows the ROC curves and the calibration plots for the classical MPMs (AUCs in brackets). MPM₂₄ II has the worst discriminating power (AUC of 0.746). SAPS II is the best discriminator, having a statistically significant larger AUC (0.835) than both APACHE II (DeLong's $\chi^2 = 5.57$, $df = 1$, p -value = 0.018) and MPM₀ II (DeLong's $\chi^2 = 3.96$, $df = 1$, p -value = 0.047). Goodness-of-fit, however, is poor for each model as demonstrated by the significantly large C*g values in Table 7-4. The calibration plot in Figure 7-1 confirms this finding: all models overestimate mortality in almost every 10% predicted mortality stratum.

Figure 7-2 presents the ROC curves for the acid-base based MPMs. The smallest AUC (MPM_{PACO₂}) is larger than the largest AUC as found among the classical MPMs (SAPS II). MPM_{BE} shows the largest AUC (0.879). The second-largest AUC is that of MPM_d (0.863). The difference between both AUCs is statistically significant (DeLong's $\chi^2 = 8.459$, $df = 1$, p -value = 0.004). The AUC for MPM_d is larger than the AUC of MPM_{PACO₂} (DeLong's $\chi^2 = 10.014$, $df = 1$, p -value = 0.002) and larger than

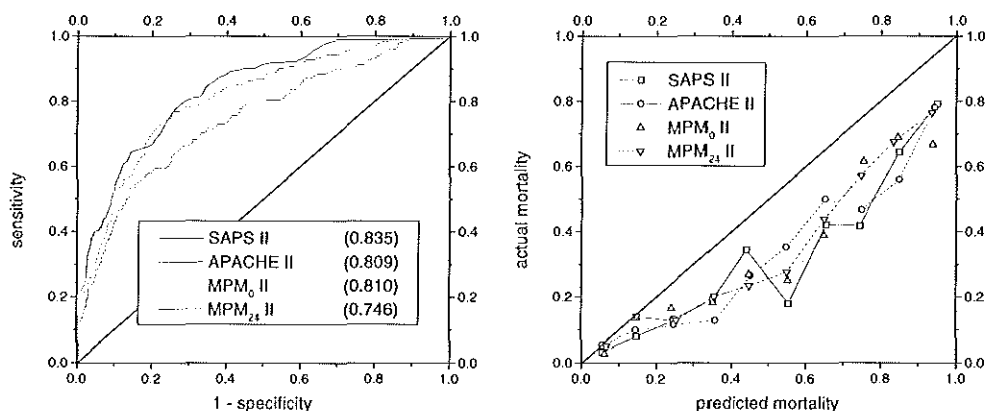


Figure 7-1. ROC curves (left) and calibration plots (right) for the classical MPMs. Between brackets the AUCs of the ROC. The line of identity in the calibration plot (line with slope 1) indicates the ideal situation: predicted mortality rate is equal to actual mortality rate.

the AUC of MPM_{PH} although the latter difference is not statistically significant (De-Long's $\chi^2 = 3.698$, $df = 1$, p -value = 0.054).

In Table 7-5, the goodness-of-fit results for the acid-base based MPMs are presented;

Table 7-3. Patient characteristics.

	total		classical MPMs				acid-base MPMs	
	<i>n</i>	%	$MPM_{24}II$		other ^b		<i>n</i>	%
			<i>n</i>	%	<i>n</i>	%		
Number of patients	2430	100	498	100	811	100	1770	100
Died in hospital	263	10.8	148	29.7	183	22.6	189	10.7
Age in years ^a	34, 66, 83		30, 67, 85		29, 66, 86		36, 67, 83	
LOS in days ^a	0.38, 0.96, 7.08		1.04, 3.13, 20.8		0.42, 1.63, 13.0		0.46, 0.96, 7.06	
Surgical								
Cardiovascular	49	2.0	15	3.0	46	5.7	38	2.1
Multiple trauma	36	1.5	16	3.2	29	3.6	21	1.2
Heart valve surgery	273	11.2	-	-	-	-	206	11.6
Peripheral vasc. surg. ^c	64	2.6	22	4.4	61	7.5	55	3.1
Haemorrhagic shock	19	0.8	9	1.8	16	2.0	14	0.8
Chronic cardiov. dis. ^d	1241	51.1	-	-	-	-	918	51.9
Gastrointestinal	104	4.3	48	9.6	94	11.6	78	4.4
Haematological	2	0.1	2	0.4	2	0.2	1	0.1
Renal	11	0.5	3	0.6	11	1.4	8	0.5
Metabolic	2	0.1	0	0.0	2	0.2	2	0.1
Neurological	4	0.2	1	0.2	4	0.5	3	0.2
Respiratory	126	5.2	31	6.2	107	13.2	93	5.3
<i>Total surgical</i>	1931	79.5	147	29.5	372	45.9	1437	81.2
Non-surgical								
Cardiovascular	270	11.1	192	38.6	230	28.4	183	10.3
Gastrointestinal	23	0.9	16	3.2	21	2.6	17	1.0
Haematological	2	0.1	2	0.4	2	0.2	-	-
Renal	14	0.6	12	2.4	14	1.7	10	0.6
Metabolic	10	0.4	6	1.2	10	1.2	7	0.4
Neurological	47	1.9	17	3.4	38	4.7	23	1.3
Respiratory	133	5.5	106	21.3	124	15.3	93	5.3
<i>Total non-surgical</i>	499	20.5	351	70.5	439	54.1	333	18.8

^a5th, 50th and 95th percentile, respectively

^bAPACHE II, SAPS II, and $MPM_{24}II$

^cPeripheral vascular surgery

^dChronic cardiovascular disease

LOS, length of stay.

MPM_a shows the best fit with a small C*g value of 8.87 (p -value = 0.54). MPM_{BE} has the largest discrepancy between actual and predicted frequencies for survival and non-survival with a p -value of 0.05 for the Hosmer-Lemeshow's C*g value.

In Figure 7-2, the calibration plot for MPM_a can be found. The calibration line more closely follows the line of identity than for any of the classical MPMs.

Table 7-4. Hosmer-Lemeshow's C*g values for the classical MPMs.

predicted mortality ^a	died		survived		<i>n</i>	predicted mortality	died		survived		<i>n</i>
	O ^b	E ^c	O	E			O	E	O	E	
APACHE II						SAPS II					
[0-4.2)	1	2.0	80	79.0	81	[0-4.2)	2	1.8	74	74.2	76
[4.2-7.6)	5	4.8	76	67.2	81	[4.2-7.2)	0	4.2	78	73.8	78
[7.6-10.9)	7	7.4	75	74.6	82	[7.2-10.6)	6	7.2	79	77.8	85
[10.9-17.7)	8	11.4	72	68.6	80	[10.6-16.7)	7	10.2	72	68.8	79
[17.7-26.2)	7	17.4	74	63.6	81	[16.7-26.6)	7	17.1	77	66.9	84
[26.2-37.8)	12	25.9	70	56.1	82	[26.6-39.2)	15	27.6	71	58.4	86
[37.8-51.6)	18	35.7	63	45.3	81	[39.2-53.0)	25	36.2	58	46.8	83
[51.6-65.3)	29	46.5	52	34.5	81	[53.0-70.0)	27	53.1	60	33.9	87
[65.3-81.6)	44	59.3	37	21.7	81	[70.0-86.1)	39	62.6	42	18.4	81
[81.6-100]	52	72.2	29	8.8	81	[86.1-100]	55	67.0	17	5.0	72
Total	183	282.5	628	528.5	811	Total	183	287.0	628	524.0	811
C*g = 118.5, <i>df</i> = 10, <i>p</i> << 0.001						C*g = 130.6, <i>df</i> = 10, <i>p</i> << 0.001					
MPM ₀ II						MPM ₂₄ II					
[0-5.3)	0	3.0	80	77.0	80	[0-14.5)	3	5.1	47	44.9	50
[5.3-7.8)	2	5.6	82	78.4	84	[14.5-22.6)	10	9.2	40	40.8	50
[7.8-12.4)	9	8.0	73	74.0	82	[22.6-29.5)	6	13.2	44	36.8	50
[12.4-18.7)	9	12.2	72	68.8	81	[29.5-36.5)	10	16.4	40	33.6	50
[18.7-24.9)	11	17.5	71	64.5	82	[36.5-42.4)	11	20.0	40	31.0	51
[24.9-34.9)	18	23.6	63	57.4	81	[42.4-50.4)	10	23.1	40	26.9	50
[34.9-47.3)	18	33.1	63	47.9	81	[50.4-58.6)	15	27.3	35	22.7	50
[47.3-60.9)	21	44.1	60	36.9	81	[58.6-70.5)	20	32.3	30	17.7	50
[60.9-79.9)	41	57.8	40	23.2	81	[70.5-80.3)	28	37.7	22	12.3	50
[79.9-100]	54	68.4	24	9.6	78	[80.3-100]	35	41.1	12	5.9	47
Total	183	273.3	628	537.7	811	Total	148	225.3	350	272.7	498
C*g = 91.6, <i>df</i> = 10, <i>p</i> << 0.001						C*g = 73.16, <i>df</i> = 10, <i>p</i> << 0.001					

^aas a percentage

^bO = observed number of subjects

^cE = expected number of subjects

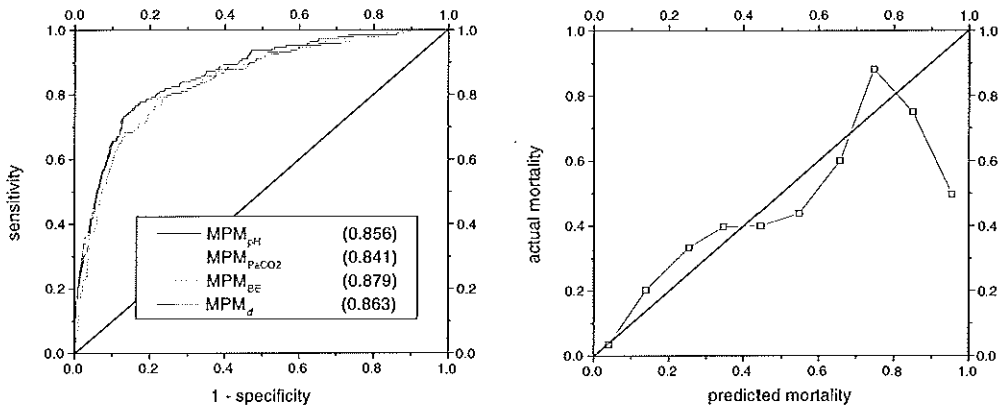


Figure 7-2. ROC curves for the acid-base based MPMs (left) and the calibration plot (right) for MPM_d .

It should be stressed that the deviation from the line of identity in the upper part of the graph is due to the limited number of subjects in the upper strata of mortality prediction. This occurs when a well-calibrated model is used on an ICU where the mortality overall is very low since, per definition, there will be only a limited number of subjects in the strata with high probabilities of dying.

Hence, considering both discriminating and predicting power, the model based on the Mahalanobis distance (MPM_d) seems to perform better than any other acid-base based MPM and better than any of the classical MPMs. MPM_d shows excellent calibration (small C^*g value) and good discrimination (large AUC)

Table 7-6 shows the results of the multiple logistic regression analysis for MPM_d as performed on the entire data set without cross-validation. For each predictor variable, the table presents an estimation of the coefficient β , its standard error (SE), and an associated odds ratio (OR). For the predictor variable *admission type* the OR indicates the decrease in the probability of dying for a surgical patient versus a non-surgical patient. For the continuous variable *age* the OR indicates the increase in the probability of dying for every 10 years of increase in age. There is a 1.5 fold increase in the probability of dying the hospital for every increase of 10 years. For the Mahalanobis distance the OR indicates the increase in probability for every unit increase of the index value, which is about 1.6 fold.

7.5 Discussion

Much of the discussion around mortality prediction models seems to be centred around the main question 'What is the model going to be used for?'. Three major areas of application can be distinguished: prediction of individual patient outcome, quality assessment and comparison of ICUs, and the stratification of ICU patients for clinical trials [20, 24]. It is generally agreed upon that the ICU scoring systems now

Table 7-5. Hosmer-Lemeshow's C^*g values for the acid-base based MPMs.

predicted mortality						predicted mortality					
died		survived		<i>n</i>		died		survived		<i>n</i>	
O	E	O	E			O	E	O	E		
MPM _{pH}						MPM _{p_aCO₂}					
[0-1.5)	1	1.8	177	176.2	178	[0-2.2)	0	2.6	177	174.4	177
[1.5-2.2)	2	3.2	176	174.8	178	[2.2-2.7)	5	4.4	172	172.6	177
[2.2-2.9)	4	4.4	173	172.6	177	[2.7-3.6)	3	5.5	173	170.5	176
[2.9-3.5)	6	5.6	171	171.4	177	[3.6-4.0)	5	6.7	172	170.3	177
[3.5-4.3)	7	6.9	170	170.1	177	[4.0-5.2)	10	7.9	168	170.1	178
[4.3-5.3)	10	8.4	166	167.6	176	[5.2-5.8)	4	9.8	175	169.2	179
[5.3-6.9)	9	10.6	169	167.4	178	[5.8-6.6)	13	10.8	164	166.2	177
[6.9-11.0)	10	15.1	167	161.9	177	[6.6-9.7)	20	14.1	157	162.9	177
[11.0-33.7)	51	37.0	126	140.0	177	[9.7-38.3)	48	41.1	129	135.9	177
[33.7-100]	89	96.0	86	79.0	175	[38.3-100]	81	86.7	94	88.3	175
Total	189	189.1	1581	1580.9	1770	Total	189	189.5	1581	1580.5	1770
C*g = 11.18, <i>df</i> = 10, <i>p</i> = 0.34						C*g = 14.03, <i>df</i> = 10, <i>p</i> = 0.17					
MPM _{BE}						MPM _a					
[0-1.4)	2	1.7	175	175.3	177	[0-1.6)	1	1.9	176	175.1	177
[1.4-2.0)	2	2.9	176	175.1	178	[1.6-2.2)	2	3.4	175	173.6	177
[2.0-2.5)	0	4.0	177	173.0	177	[2.2-2.9)	2	4.5	175	172.5	177
[2.5-3.2)	5	5.1	172	171.9	177	[2.9-3.6)	5	5.7	172	171.3	177
[3.2-4.0)	3	6.4	174	170.6	177	[3.6-4.4)	7	7.0	170	170.0	177
[4.0-4.9)	13	7.8	164	169.2	177	[4.4-5.4)	9	8.7	168	168.3	177
[4.9-6.6)	4	10.1	173	166.9	177	[5.4-6.9)	9	10.7	168	166.3	177
[6.6-10.7)	19	14.5	158	162.5	177	[6.9-10.7)	15	14.8	162	162.2	177
[10.7-32.4)	44	35.7	133	141.3	177	[10.7-34.3)	50	37.6	127	139.4	177
[32.4-100]	97	101.0	79	75.0	176	[34.3-100]	89	94.8	88	82.2	177
Total	189	189.1	1581	1580.9	1770	Total	189	189.1	1581	1580.9	1770
C*g = 18.17, <i>df</i> = 10, <i>p</i> = 0.05						C*g = 8.87, <i>df</i> = 10, <i>p</i> = 0.54					

available are not suited for predicting individual outcome [24-26]. Even if an MPM were to be designed that could adequately predict individual patient outcome, there would still be ethical discussions about whether, and if so, how to use this model as a decision aid [24].

Hence, MPMs at the present moment are most useful for comparing groups of patients within a single ICU. This comparison of patient groups is especially useful when doing clinical research. For instance, when comparing different treatments or drugs on the ICU, one has to have a way to ensure that patients receiving the placebo

Table 7-6. Results of a multiple logistic regression analysis on the total data set of 1770 patients for MPM_d

	β^a	SE(β)	OR	95% CI
Admission type (1 = surgical, 0 = non-surgical)	-2.0624	0.1934	0.13	0.09 – 0.19
Age on a 10 year scale	0.4249	0.0697	1.53	1.33 – 1.75
Largest d^b in first 24 hours after ICU admission	0.4727	0.0637	1.60	1.42 – 1.82
Constant	-4.6069	0.5189		

^aestimated weight

^bMahalanobis distance

SE=standard error; OR=odds ratio, and CI=confidence interval

As an example, a 30-year-old surgical patient with a d of 10 has a 16-fold (10×1.60) higher chance of dying in the hospital than a similar patient with a d of 0. Predicted mortality for this patient is: $e^{(-4.6242 + (2.0137 \times 1) + (0.4352 \times 3) + (0.4728 \times 10))} / (1 + e^{(-4.6242 + (2.0137 \times 1) + (0.4352 \times 3) + (0.4728 \times 10))}) = 0.35$.

and patients receiving the actual treatment or drugs are comparable with respect to their severity of illness. The second use of MPMs is comparison of well-defined groups of patients between ICUs (case-mix adjusted). When MPMs are used for such purposes, the emphasis must be on good calibration rather than on good discrimination [20].

We demonstrated that APACHE II, SAPS II, MPM_0 II, and MPM_{24} II show acceptable degrees of discrimination (large AUCS) for our ICU, but a poor calibration (large Hosmer-Lemeshow's C^*g values). Over the whole spectrum of mortality predictions, each model predicted a higher mortality than was actually observed (see Figure 7-1). This can be explained in two ways: the scoring systems are not calibrated for our patient population, or our ICU performs better than expected. Moreover, after applying the exclusion criteria inherent to the respective models, only 811 of the initial 2430 patients remained for which APACHE II, SAPS II, and MPM_0 II predictions could be calculated. The bulk of the 1611 discarded records were post-operative cardio-surgical patients. As post-operative cardio-surgical patients comprise the majority of our ICU admissions, exclusion of these patients severely hampers the ability of the models to describe the nature of our ICU.

There are also more general disadvantages of the classical MPMs. The number of data items necessary to predict mortality ranges from 95 to 130 items [27, 28]. The manual extraction of these data is a very time-consuming process. The issue of intra-observer and inter-observer variability in data-extraction has to be explored further [29]. Even if the problem of observer variability is solved, problems concerning the manual charting method still remain. Errors in manual vital signs recording occur in approximately 20% of the nursing shifts [30]. The most elegant solution to this problem seems to be the implementation of a Patient Data Management System. However, none of the commercially available Patient Data Management Systems has a solid extraction method for these data [31]. The classical MPMs are very sensitive to

items like sampling rate, sampling method and artefacts [5, 32, 33]. Problems with lead-time bias, case-mix differences and external (international) validation of the models keep cluttering the critical care literature [34-38].

MPM_d seems to be an attractive model for the prediction of mortality at our ICU: discriminating power is good and calibration is excellent. MPM_{BE} has a significantly better discrimination but its calibration, however, is less than MPM_d . It is not surprising that models based on invasive haemodynamic variables perform well in predicting hospital mortality outcome of ICU patients. They reflect the physiological status of a patient and especially physiological instability is the primary determinant of mortality [9]. In this light, it is interesting to observe that MPM_{24} II has the worst discriminating power (AUC of 0.746) in this study. MPM_{24} II hardly uses any physiological data, since it is mainly based on acute and chronic diagnoses. The reason why MPM_d is superior over the single-variable acid-base MPMs may be that it incorporates information of all three separate acid-base variables. Furthermore, since the Mahalanobis distance is based on a reference distribution of data coming from ICU patients of our own ICU, it may more closely represent relevant deviations in arterial acid-base status than an absolute z -score for a single acid-base variable.

MPM_d has several advantages over the classical MPMs. Collection of data is simple and straightforward: age, type of admission (surgical, non-surgical), and the arterial blood gas analysis with the largest d during the first 24-hour after ICU admission. No exclusion criteria were used: all 1770 consecutive admissions were included. The Mahalanobis distance based MPM has been developed on all ICU admissions. Selker [39] defined a number of desirable features of an ICU mortality prediction model: 1) sufficiently accurate and easy to use in real-time clinical settings, 2) based on the first minutes of hospital presentation, 3) not affected by whether a patient is hospitalised, 4) based on data collected in the usual care of patients, 5) integrated into a computer system, 6) independent of the diagnosis related groups system, and 7) open for inspection and testing. Although MPM_d does not fulfil all requirements, some of the more important points are covered. The analysis of the arterial acid-base status has become an indispensable part of ICU patient care and arterial acid-base values are routinely measured (point four) for a majority of ICU patients, often right at the start of ICU admission (point two). As regards point five, for this study results of the arterial acid-base measurements were sent from the central clinical chemistry laboratory back to the ICU and automatically recorded on a microcomputer installed at the ICU. Therefore, no extra data extractions are necessary. Moreover, MPM_d is based on objectively measurable variables rather than on subjective diagnostic systems (point six). However, since arterial punctures are not common outside the hospital, MPM_d does violate point three of Selkes' list. Also, since reference distributions need to be constructed for ICUs separately, MPM_d cannot be unconditionally transferred from one

ICU to another. Point seven may therefore be difficult to adhere to if inspection and testing is to be performed in ICUs other than the one for which the MPM_d was designed.

It must be stressed that the number of patients for the development of MPM_d was limited (1770). Therefore, MPM_d awaits prospective validation in the same ICU. Other future research may focus on the extension of the multivariate reference model with other variables, for instance arterial PaO_2 . Another line of research is extension of the model to other ICUs.

7.6 Acknowledgement

We thank J. van Leeuwen (department of Clinical Chemistry and Haematology) for assistance in linking the microcomputer to the laboratory of Clinical Chemistry and Haematology.

7.7 References

1. Becker RB, Zimmerman JE: ICU scoring systems allow prediction of patient outcomes and comparison of ICU performance. *Crit Care Clin* 1996; 12:503-514
2. Knaus WA, Draper EA, Wagner DP, et al: APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13:818-829
3. Le Gall JR, Lemeshow S, Saulnier F: A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *J Am Med Assoc* 1993; 270:2957-2963
4. Lemeshow S, Teres D, Klar J, et al: Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *J Am Med Assoc* 1993; 270:2478-2486
5. Bosman RJ, Oudemans van Straaten HM, Zandstra DF: The use of intensive care information systems alters outcome prediction. *Intensive Care Med* 1998; 24:953-958
6. Suistomaa M, Ruokonen E, Takala J, et al: Lack of standardised data collection causes bias in APACHE II and SAPS scores. *Intensive Care Med* 1997; 23:545
7. Féry-Lemonnier E, Landais P, Loirat P, et al: Evaluation of severity scoring systems in ICUs--translation, conversion and definition ambiguities as a source of inter-observer variability in Apache II, SAPS and OSF. *Intensive Care Med* 1995; 21:356-360
8. Yeung HC, Lu MW, Martinez EG, et al: Critical Care Scoring System--new concept based on hemodynamic data. *Crit Care Med* 1990; 18:1347-1352
9. Sarmiento J, Torres A, Guardiola JJ, et al: Statistical modeling of prognostic indices for evaluation of critically ill patients. *Crit Care Med* 1991; 19:867-870
10. Hosmer DW, Lemeshow S: Applied logistic regression. New York: Wiley-Interscience, 1989, Barnett V, Bradley RA, Stuart Hunter J, Kadane JB, eds., *Wiley series in probability and mathematical statistics*
11. Kollef MH, Schuster DP: Predicting intensive care unit outcome with scoring systems. *Crit Care Clin* 1994; 10:1-18
12. Ruttimann UE: Statistical approaches to development and validation of predictive instruments. *Crit Care Clin* 1994; 10:19-35
13. Jolliffe IT, Morgan BJT: Principal component analysis and exploratory factor analysis. *Stat Meth Med Res* 1992; 1:69-95
14. Hekking M, Lindemans J, Gelsema ES: Design and representation of multivariate patient-based reference regions for arterial pH, PaCO₂ and base excess values. *Clin Biochem* 1995; 28:581-585
15. Albert A, Harris EK: Multivariate interpretation of clinical laboratory data. New York: Marcel Dekker Inc, 1987, Owen DB, Cornell RG, eds., *STATISTICS: Textbooks and Monographs*; vol 75
16. Boyd JC, Lacher DA: The multivariate reference range: an alternative interpretation of multi-test profiles. *Clin Chem* 1982; 28:259-265
17. Slotnick HB, Etzell P: Multivariate interpretation of laboratory tests used in monitoring patients. *Clin Chem* 1990; 36:748-751
18. Solberg HE: Multivariate reference regions. *Scand J Clin Lab Invest Suppl* 1995; 222:3-5
19. Winkel P, Lyngbye J, Jørgensen K: The normal region - a multivariate problem. *Scand J Clin Lab Invest* 1972; 30:339-344
20. Angus DC, Pinsky MR: Risk prediction: judging the judges. *Intensive Care Med* 1997; 23:363-365
21. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36

22. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44:837 - 845
23. Lemeshow S, Hosmer DW: A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; 115:92 - 106
24. Sherck JP, Shatney CH: ICU scoring systems do not allow prediction of patient outcomes or comparison of ICU performance. *Crit Care Clin* 1996; 12:515-523
25. Schäfer JH, Maurer A, Jochimsen F, et al: Outcome prediction models on admission in a medical intensive care unit: do they predict individual outcome? *Crit Care Med* 1990; 18:1111-1118
26. Lemeshow S, Klar J, Teres D: Outcome prediction for individual intensive care patients: useful, misused, or abused? *Intensive Care Med* 1995; 21:770-776
27. ICNARC Case Mix Program Dataset Specification by the Intensive Care National Audit & Research Centre. London, England, 1995
28. Bosman RJ: Dutch National Intensive Care Evaluation (NICE) datadictionary.
29. Ledoux D, McKinley S, Fisher M, et al: Assessment of interobserver correlation of the APACHE II score. *Intensive Care Med* 1998; 24:14
30. Hammond J, Johnson HM, Varas R, et al: A qualitative comparison of paper flowsheets vs a computer-based clinical information system. *Chest* 1991; 99:155-157
31. de Keizer NF, Stoutenbeek CP, Hanneman LA, et al: An evaluation of Patient Data Management Systems in Dutch intensive care. *Crit Care Med* 1998; 24:167-171
32. Taylor DEM, Whamond JS: Reliability of human and machine measurements in patient monitoring. *Eur J Intensive Care Med* 1975
33. Pollack MM, Patel KM, Ruttimann U, et al: Frequency of variable measurement in 16 pediatric intensive care units: influence on accuracy and potential for bias in severity of illness assessment. *Crit Care Med* 1996; 24:74-77
34. Rowan KM, Kerr JH, Major E, et al: Intensive Care Society's APACHE II study in Britain and Ireland-II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *Br Med J* 1993; 307:977-981
35. Beck DH, Taylor BL, Millar B, et al: Prediction of outcome from intensive care: a prospective cohort study comparing Acute Physiology and Chronic Health Evaluation II and III prognostic systems in a United Kingdom intensive care unit. *Crit Care Med* 1997; 25:9-15
36. Nouira S, Belghith M, Elatrous S, et al: Predictive value of severity scoring systems: comparison of four models in Tunisian adult intensive care units. *Crit Care Med* 1998; 26:852-859
37. Dragsted L, Jorgensen J, Jensen NH, et al: Interhospital comparisons of patient outcome from intensive care: importance of lead-time bias. *Crit Care Med* 1989; 17:418-422
38. Murphy-Filkins R, Teres D, Lemeshow S, et al: Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: how to distinguish a general from a specialty intensive care unit. *Crit Care Med* 1996; 24:1968-1973
39. Selker HP: Systems for comparing actual and predicted mortality rates: characteristics to promote cooperation in improving hospital care. *Ann Intern Med* 1993; 118:820-822

Clinical evaluation of a tri-axial chart, multivariate reference model and
vector classification method for arterial acid-base data

Marcel Hekking, Herman J.L.M. Ulenkate

8.1 Introduction

In the preceding chapters, a computer program (ABCHART) has been described for the quantitative and qualitative monitoring of arterial acid-base data in an intensive care setting. ABCHART displays acid-base trajectories of patients in a tri-axial acid-base chart, calculates Mahalanobis distances to assess multivariate abnormality, and classifies arterial acid-base observations according to the vector method presented in Chapter 3. In the present chapter, a study is described that aims at an evaluation of the ABCHART computer program in a clinical setting. This evaluation will comprise three separate investigations.

First, the effect of the introduction of the computer program ABCHART in an intensive care setting is assessed. Use of the graphical tri-axial acid-base chart and Mahalanobis distances in daily clinical practice provides the intensive care clinician with additional information. This extra information may well lead to changes in the way clinicians request and use acid-base laboratory analyses.

The second investigation concerns the comparison of the vector classification method with the Astrup and Siggaard-Andersen method. The vector method was developed as an alternative for the Astrup and Siggaard-Andersen method, since the latter has some unwanted features, as described in Chapter 1.

Finally, the discriminating power (*i.e.* the ability to discriminate 'normal' from 'abnormal' arterial acid-base observations) of the multivariate reference model will be compared with the discriminating power of the multiple univariate approach. Separating abnormal from normal arterial acid-base observations may be very useful in an intensive care setting where monitoring plays a key role. The classical approach is to evaluate laboratory data against relevant, fixed univariate reference intervals. In Chapter 4, it was demonstrated that this approach may lead to erroneous conclusions in cases where multiple variables are evaluated simultaneously. When assessing the abnormality of multiple laboratory measurements, the Mahalanobis distance is theoretically to be preferred. Small Mahalanobis distances are associated with acid-base normality, whereas larger values indicate acid-base abnormality. The use of a specific cut-off point for the Mahalanobis distance leads naturally to a dichotomous classification system. One of the goals of this study is to demonstrate a possible superiority of the Mahalanobis distance over the classical univariate approach in a clinical setting. In 1992, Wulkan [1] demonstrated the superiority of a trivariate reference model for pH, PaCO_2 , and base excess values over a univariate interval approach. Two senior intensive care clinicians were asked to independently classify acid-base observations coming from their ICU as being normal or abnormal. Wulkan demonstrated that using a cut-off value of 7.8 for the squared Mahalanobis distance (> 7.8 is abnormal and ≤ 7.8 is normal) resulted in a better agreement with the judgement of the clinicians than the

application of standard univariate reference intervals. A prerequisite, however, for a valid use of multivariate reference models is that there is no functional dependency between the variables included in the model. This prerequisite does not hold for a tri-variate pH, PaCO_2 and BE reference model, since BE is calculated from pH and PaCO_2 (with a fixed value for haemoglobin) [2-4]. In Chapter 2, the relationship between pH, PaCO_2 and BE was demonstrated to be almost linear. Moreover, in the study of Wulkan the degree of agreement of the multivariate model with the clinicians' judgement was only determined for a single squared Mahalanobis distance cut-off point (7.8), using Cohen's kappa statistic [5, 6]. However, the use of the kappa statistic does not provide an overall measure of a model's discriminating power, since the degree of agreement will be different for other cut-off points. A last drawback of Wulkan's study is that the clinicians performed their classifications in a non-clinical setting outside the ICU.

8.2 Materials and methods

8.2.1 Set-up and patient data

The study took place at the general ICU of the St. Elisabeth hospital (Tilburg, The Netherlands). At the time of the study the ICU harboured 10 beds of which five were located on the ward where the study took place. During the study period, in total five resident clinicians took care of the daily medical needs of the patients on the ward. The residents were supervised by two senior clinicians.

A microcomputer (66 MHz 80486 Laser Personal Computer) with a colour inkjet printer (Hewlett Packard Deskjet 600c) was installed at the central desk of the ward to run the ABCHART program. Typical screen-output of ABCHART consists of a tri-axial chart to display acid-base trajectories, pertinent acid-base values and Mahalanobis distances, and a graphical trend plot of the Mahalanobis distances. Typical printer-output consists of the tri-axial chart, pertinent acid-base values and Mahalanobis distances. A direct link with the department of clinical chemistry was established as described in Section 5.4.5. During the study, ABCHART automatically received patient and laboratory data from the department of clinical chemistry. Patient and laboratory data included: patient identification number, patient name, initials, date of birth, sample identification number, arterial pH, PaCO_2 , BE, HCO_3^- , PaO_2 , and oxygen saturation. Arterial acid-base measurements were performed in duplicate on a ABL 520 (Radiometer) and a AVL OMNI 6 (Merck) at the department of clinical chemistry. The mean values were reported to the ICU. The tri-axial chart and the multivariate reference model were based on the *ELIbe* data set and constructed according to the methods described in Chapters 2, 3, and 4.

8.2.2 Study design

Three data acquisition periods (I, II, and III) were held. The total study period comprised 11 months (October 1996 to August 1997). During period I (or baseline), ABCHART was connected to the department of clinical chemistry, but screen-output and printer-output was switched off. During period II, output to screen and printer was enabled. Residents were able to view the plots of acid-base trajectories of ICU patients in the tri-axial chart and to inspect the calculated Mahalanobis distances. For each incoming arterial acid-base report, ABCHART automatically generated printer-output. Nursing staff was instructed to attach the printer-output to the original acid-base report so that the tri-axial chart and Mahalanobis distances were available to the residents when visiting their patients. During period III, output to screen and printer was suppressed again.

During all three periods the following parameters were measured to demonstrate a possible effect of the introduction of the ABCHART computer program:

- the number of arterial acid-base analyses per patient,
- the time between consecutive analyses,
- the distance in the tri-axial chart between consecutive analyses,
- the percentage of arterial acid-base analyses with the patient being on artificial respiration.

Residents were asked to classify acid-base observations during period II into one of 17 categories: four primary acid-acid base disturbances with three degrees of compensation (no compensation, partly compensated and fully compensated), four combined primary acid-base disorders, and normal. Classifications were to be made on the printer-output of ABCHART.

To obtain a 'silver' standard data set, the two senior clinicians were asked to also classify the arterial acid-base observations of period II, using the same classification categories as used by the residents. For this purpose, a special computer program was built that allowed each senior clinician to independently classify arterial acid-base observations on the ward. The database of personal classifications of each senior clinician was protected with a password.

8.2.3 Data analysis

Possible changes between the three periods with regard to the parameters described in the preceding section were tested with a non-parametric Kruskal-Wallis (K-W) test or a χ^2 -test. An α of 0.05 was used in all tests.

Acid-base classification agreements were assessed using Cohen's kappa [5, 6]. Given a data set of 'true' acid-base classifications, Cohen's kappa is the proportion of agreeing classifications, taking into account the fact that agreement can also occur by chance. Kappa is zero when the agreement obtained can be explained by chance only and

unity when there is complete agreement. In general, kappa values above 0.75 are considered to represent excellent agreement, while kappa values of 0.40 and below represent poor agreement [7, 8]. The data set with 'true' classifications comprised all arterial acid-base observations classified identically by both senior clinicians.

To compare the discriminating power of the multivariate reference model with the univariate reference model, Receiver-Operator-Characteristics (ROC) curves were used [9, 10]. Given a data set of acid-base observations with their Mahalanobis distances and a 'true' normal/abnormal classification, a ROC curve may be obtained by plotting the pairs of *sensitivity* (y) and *1-specificity* (x) as they appear when using the Mahalanobis distance of each observation as a cut-off point to classify the entire data set. The resulting area under the ROC curve (AUC) expresses the degree of overall discriminating power without the need to select one specific Mahalanobis distance as a cut-off point. An AUC of 0.5 means zero discriminating power, while an AUC of 1 indicates optimal discriminating power.

The Mahalanobis distance can straightforwardly be applied as a classifier. The multiple univariate approach, however, does not provide such a classifier based on a single variable. Therefore, a univariate analog of the Mahalanobis-distance was developed, as shown in Figure 8-1. The data set with the 'true' normal/abnormal classifications was constructed from the arterial acid-base observations classified identically (normal or abnormal) by both senior clinicians. AUCs were calculated according to the method of Hanley and McNeil [10].

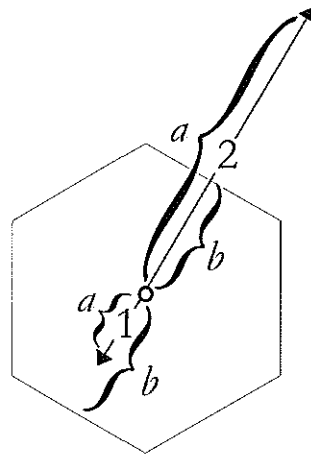


Figure 8-1. Calculation of the univariate index of abnormality. The vectors 1 and 2 represent two examples of acid-base observation vectors in the tri-axial acid-base chart. The hexagon represents the standard normal reference region (see Chapter 3). The index is calculated as 'alb'. Thus, observations outside the normal region will have values larger than 1 (vector 2), while observations within the region will have values smaller than 1 (vector 1). Values will be larger for observations located further away from the origin (o).

8.3 Results

8.3.1 Effect of the introduction of ABCHART

Table 8-1 gives a summary for the three data acquisition periods. The ABCHART program received a total of 570, 816, and 453 valid arterial acid-base observations during period I, II, and III, respectively. There are no statistically significant differ-

Table 8-1. Summary of data acquisition periods.

	period I	period II	period III
Acquisition period	14/10/96 – 31/12/96	14/01/97 – 25/5/97	5/6/97 – 20/8/97
Number of male patients	67 (67.2 %)	67 (67.2 %)	52 (51.9 %)
Total number of ABGs	570	816	453
Age mean	56	62	56
median	60	68	61

ences between the three periods regarding gender (p -value of χ^2 test statistic = 0.155) and age of patients (p -value of K-W test statistic = 0.235).

Table 8-2 gives the mean and median values for the four parameters defined in Section 8.2.2 during periods I, II, and III. There are no statistically significant differences in the number of arterial blood gas analyses per patient (p -value of K-W test statistic = 0.172), hours between consecutive analyses (p -value of K-W test statistic = 0.689), nor in the distance in the tri-axial chart between consecutive analyses (p -value of K-W test statistic = 0.185). A statistically significant difference between the three periods was found for the percentage of acid-base analyses with the patient being on artificial respiration (p -value of χ^2 test statistic = 0.001). However, this difference is predominantly caused by a deviating value during period III, indicating that also for this parameter there is no effect of the introduction of the computer program.

Table 8-2. Effect of the introduction ABCHART on arterial acid-base parameters

		period I	period II	period III
Number of acid-base analyses per patient.	mean	8.51	12.18	8.71
	median	5	7	4
Hours between consecutive acid-base analyses.	mean	13.5	13.2	16.1
	median	9	9	8.5
Distance in tri-axial chart between consecutive analyses.	mean	2.68	2.67	2.78
	median	1.92	2.18	2.05
Percentage of analyses with patient on artificial respiration.		84.5 %	85.0 %	70.1 %

8.3.2 Acid-base classification agreements

During period II, in total 798 arterial acid-base observations were classified by both senior clinicians. To be able to compare these classifications with the classifications of the vector method and the classifications of the Astrup and Siggaard-Andersen method, the initial 17 categories were regrouped into the 13 categories of Table 1-2.

Table 8-3. Classification matrices of the vector method and the Astrup and Siggaard-Andersen method versus the joint judgement of the senior clinicians. Columns represent the classifications of the clinicians, while rows represent the classifications of both methods. In a single column the numbers on the right in bold represent the Astrup and Siggaard-Andersen method, while the non-bold values on the left represent the vector method.

Class	1 res. aci.	2 part. comp. res. aci.	3 combi1	4 part. comp. met. alka.	5 met. alka.	6 resp. alka. + met. alka.	7 resp. alka.	8 part. comp. res. alka.	9 combi2	10 part. comp. met. aci.	11 met. aci.	12 resp. aci. + met. aci.	13 normal	Total
1	23 29											2 2		25 31
2	3	2 1	8										2	15 1
3			25 35										6 4	31 39
4			8 1		3								11	22 1
5					22 25	2 5							10 5	34 35
6					8 2	6 1								14 3
7					1 1	36 38			1				4 1	40 41
8							1		17 1				6	24 1
9									23 36	1			2 1	25 38
10									6	5 6	1 1		6	18 7
11										3 1	65 65	6		74 66
12	3											48 54	1	52 54
13													157 157	157 157
14		1	5			2	4		8				37	57
Total	29 29	2 2	41 41		25 25	10 10	43 43		46 46	8 8	66 66	56 56	205 205	531 531

Res. = respiratory, met. = metabolic, aci. = acidosis, alka. = alkalosis, comp. = compensated, part. = partly.

Combi1 and combi2 represent situations in which no clear single classification can be given without additional clinical information or information about the (medical) history of the patient. In these situations, three acid-base classification categories are in theory possible. See also Table 1-2.

Using these 13 categories, 531 of the 798 observations were given the same classification by both senior clinicians. These 531 observations were then classified using both the vector method and the Astrup and Siggaard-Andersen method. In both classification schemes, the standard univariate reference intervals were used to determine acid-base abnormality. The resulting classification matrices are given in Table 8–3. Cohen's kappa for the Astrup and Siggaard-Andersen method versus the clinicians' diagnoses was shown to be 0.81 with a standard error (SE) of 0.019, while Cohen's kappa for the vector method was 0.74 with an SE of 0.21. Both classification models show therefore good to excellent agreement with the joint judgement of the senior clinicians. The Astrup and Siggaard-Andersen method, however, seems to perform better than the vector method, even though observations are found in category 14 (unclassifiable) of the Astrup and Siggaard-Andersen method which is – per definition – not the case in the vector method.

Using the same 'silver' standard of 531 arterial acid-base observations, we were also able to assess the agreement of the senior clinicians' judgement with the judgement of the residents. From the 531 acid-base observations, 431 were also classified by the residents. Cohen's kappa for these 431 cases was shown to be 0.32 with an SE of 0.033 indicating a rather poor agreement.

8.3.3 Comparison of reference models

Of the 798 acid-base observations classified by both senior clinicians, 727 were given the same classification (normal or abnormal). Using these 727 observations as the 'silver' standard, the AUC for the multivariate reference model and for the multiple univariate reference model were both 0.994, with an SE of 0.004. Hence, both models perform equally well in separating normal from abnormal arterial acid-base observations and a superiority of the multivariate model over the multiple univariate model could not be confirmed.

The optimum cut-off point for the Mahalanobis distance (*i.e.* the point where the total accuracy is the greatest) was shown to be 1.01. Using this cut-off value for the Mahalanobis distance yielded a sensitivity of 0.964 and a specificity of 0.961 for this particular 'silver' standard data set. The optimum cut-off point for the univariate index of abnormality was shown to be 1.24 with an associated sensitivity of 0.977 and specificity of 0.941. This value of 1.24 indicates that the senior clinicians applied a slightly broader definition of normality than the one defined by the standard univariate normal reference intervals for the three arterial acid-base parameters pH, PaCO₂, and BE.

8.4 Discussion

One of the goals of this study was to investigate whether the introduction of the computer program ABCHART presented in Chapter 5 would have any influence on some

aspects of arterial blood gas analysis on an ICU. We did not find any significant influence of the computer program on the parameters investigated. In an evaluation with the residents during the study, it was indicated that the residents did look at the output of the computer program, but that actual clinical decisions were taken on the original acid-base data. The new acid-base chart was found to be interesting but its tri-axial configuration was still confusing for most of them. Furthermore, although the concept of the Mahalanobis distance was reasonably well understood by all residents, they never actually used it in their decision making process. The fact that we could not measure any significant difference in the parameters between the data acquisition periods does not mean, however, that the Mahalanobis distance and the tri-axial acid-base chart do not have any clinical value. First, this is only one study in one particular ICU and (larger) investigations should be conducted in other ICUs. Second, arterial blood gas analysis has become a routine operation on an ICU. Analyses are regularly performed in the mornings, evenings, and, depending on the status of the patient, at fixed time points during the day. It may be very hard to change that pattern of request behaviour with the introduction of a totally new and unfamiliar arterial acid-base graph and reference model.

Another goal of the study was to assess the performance of two acid-base classification methods; the vector method of Chapter 3 and the Astrup and Siggaard-Andersen method of Chapter 1. Use of computerised arterial acid-base data classification, especially in an intensive care setting, may be very beneficial [1, 11-14]. We confirmed this finding for this particular ICU, as we demonstrated that the degree of agreement between the senior clinicians' judgement and that of the residents was only poor.

Both the vector method and the Astrup and Siggaard-Andersen method performed well in classifying pH, PaCO_2 , and BE values as demonstrated with relatively large kappa values. However, the Astrup and Siggaard-Andersen method appeared better than the vector method. A regular use of a computerised Astrup and Siggaard-Andersen classification scheme may well enhance the performance of the residents with respect to the classification of arterial acid-base measurements. The performance of the Astrup and Siggaard-Andersen method may be increased by the development of specific classification rules for the category 'unclassifiable'. The performance may increase even more if a slightly larger normal region is applied than the one defined by the standard normal univariate reference intervals for pH, PaCO_2 , and BE.

Our last goal was to demonstrate a possible superiority of the multivariate reference model over the multiple univariate approach with regard to the discrimination between normal and abnormal acid-base observations. With the data from this ICU, we could not demonstrate such superiority but with other patient populations this may be different. For this particular ICU, the Mahalanobis distance closely resembles the univariate index of abnormality, since the centre of the multivariate reference region is

located near the origin of the tri-axial chart (see Table 4–1). Moreover, there is only a weak correlation between the two principal components in the reference data set. The location and shape of a patient-based multivariate acid-base reference region is very much dependent on the nature of an ICU, the kind of patients on which the reference region is based, and the operational standards on that particular ICU. A multivariate acid-base reference model for a totally different ICU (for instance the model for the neonatal ICU of Chapter 4) may indeed show larger differences between the Mahalanobis distance values and the values for the univariate index of abnormality. It then remains to be investigated whether these differences are in favour of the Mahalanobis distance. When such investigations are conducted it may be worthwhile to evaluate the models not only against ‘normal’/‘abnormal’ classifications but also against interpretations like ‘acid-base status is acceptable’/‘not acceptable’. An interpretation of the acid-base values in such a way more or less incorporates the nature of a particular ICU and the operational standards on that ICU. A difference between the multivariate and the univariate reference model may in this case indeed be in favour of the multivariate model. Thus, based on the results of this first clinical evaluation of the Mahalanobis distance it is concluded that further investigations need to be performed. Moreover, although for the evaluation of acid-base values on this ICU the univariate index is as useful as the Mahalanobis distance, the Mahalanobis distance is still to be preferred since it ensures that, if the model is to be extended with other variables, newly introduced correlations between variables are accounted for.

In conclusion, the value of the Mahalanobis distance needs to be further assessed with other patient populations and treatment regimens using different reference models. The most comprehensive and appropriate way to do so is to compare the classification performance of the Mahalanobis distance with that of the univariate index of abnormality, using ROC analysis as described in this chapter. With regard to the classification of arterial acid-base disorders, the Astrup and Siggaard-Andersen classification method was to be preferred over the vector method for this particular ICU. Moreover, for this ICU, acid-base abnormality may be assessed both with the Mahalanobis distance and the univariate index, using the optimum cut-off values of 1.01 and 1.24, respectively. Classification rules need to be generated for the ‘unclassifiable’ regions in the Astrup and Siggaard-Andersen classification method.

8.5 Acknowledgement

We thank the nursing staff and personnel of ward A5 of the St. Elisabeth hospital for their kind co-operation.

8.6 References

1. Wulkan RW. Expert systems and multivariate analysis in clinical chemistry. Rotterdam: Erasmus University Rotterdam, 1992; 111 pp.
2. Siggaard-Andersen O, Siggaard-Andersen M. The oxygen status algorithm: a computer program for calculating and displaying pH and blood gas data. *Scand J Clin Lab Invest* 1990; 203:29-45.
3. Siggaard-Andersen O, Engel K, Jørgensen K, et al. A micro method for determination of pH, carbon dioxide tension, base excess and standard bicarbonate in capillary blood. *Scand J Clin Lab Invest* 1960; 12:172-176.
4. Astrup P. A new approach to acid-base metabolism. *Clin Chem* 1961; 7:1-15.
5. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measurem* 1960; 20:37-46.
6. Schouten HJA. Statistical measurements of interobserver agreement for categorical data (Thesis). Rotterdam: Erasmus University, 1985; 131 pp.
7. Moorman PW. Towards formal medical reporting (Thesis). Rotterdam: Erasmus University, 1995; 138 pp.
8. Landis JR. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:671-679.
9. Zweig MH, Campbell G. Receiver-Operating Characteristics (ROC) plots. A fundamental Evaluation Tool in Clinical Medicine. *Clin Chem* 1993; 39:561-577.
10. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36.
11. Broughton JO, Jr, Kennedy TC. Interpretation of arterial blood gases by computer. *Chest* 1984; 85:148-149.
12. Horn DL, Radhakrishnan J, Saini S, et al. Evaluation of a computer program for teaching laboratory diagnosis of acid-base disorders. *Comput Biomed Res* 1992; 25:562-568.
13. Schreck DM, Zacharias D, Grunau CF. Diagnosis of complex acid-base disorders: physician performance versus the microcomputer. *Ann Emerg Med* 1986; 15:164-170.
14. Hingston DM, Irwin RS. A computerized interpretation of arterial pH and blood gas data: Do physicians need it? *Resp Care* 1982; 2:809-815.

9

Summary, general conclusions, and future research

9.1 Summary

This thesis deals with the evaluation and classification of arterial acid-base measurements in an intensive care unit (ICU) using chemometric techniques. In *Chapter 1*, some fundamental principles underlying human acid-base physiology are discussed and the present ways of evaluating arterial acid-base values in an intensive care unit are described. There are two distinct 'schools' of arterial acid-base analysis. One uses arterial pH, PaCO_2 and the bicarbonate ion concentration ($\text{a}[\text{HCO}_3^-]$), while the other school advocates the use of arterial pH, PaCO_2 and base excess (BE). A strict adherence to the classification scheme for arterial pH, PaCO_2 and BE as proposed by Astrup and Siggaard-Andersen leads to a category 'unclassifiable'. Moreover, fundamental problems exist with respect to the use of standard 95% reference intervals when evaluating acid-base normality. First, in both 'schools' there is an (almost) linear relationship between the three basic acid-base variables used. This means that one of the three 95% reference intervals is redundant. Second, from a statistical point of view the simultaneous use of multiple 95% reference intervals may lead to erroneous conclusions. This thesis presents a method for deriving valid multivariate acid-base reference models for both 'schools' of acid-base interpretation, to be used in an intensive care setting. Furthermore, it proposes a method for the classification of pH, PaCO_2 and BE values which has no 'unclassifiable' categories. Finally, this thesis presents the results of a first clinical evaluation of the proposed acid-base reference model and classification system.

The essence of the reference model is that a multivariate reference region is derived from a large distribution of arterial acid-base data (pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ or pH, PaCO_2 and BE) coming from ICU patients themselves. The first step in the process is to remove the redundancy that is present in such an acid-base data distribution. In *Chapter 2*, the use of a statistical technique called principal component analysis (PCA) is described to perform such a data reduction. It presents the results of PCA transformations of six acid-base data distributions coming from four different intensive care units. Results show that for each PCA transformed acid-base distribution more than 99% of the variance is concentrated in the first two principal component axes. PCA may therefore be a valuable tool for the validation of arterial acid-base measurements. Given a PCA transformed acid-base distribution, a new combination of acid-base measurements may be checked for consistency, by calculating the value of the third principal component. This value must be close to the observed mean of the third principal component in the reference distribution, for the observation to be valid.

Based on the PCA results of Chapter 2, *Chapter 3* proposes a new graphical representation for the combination arterial pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ and for the combination pH, PaCO_2 and BE. Since the plane spanned by the first two principal compo-

nent axes after a PCA transformation represents the major part of the variance in an acid-base data distribution, it is sufficient to plot new acid-base observations in this principal component subspace. In order to have a standard appearance, a rotation in the principal component subspace is applied in such a way that the projection of the pH-axis is always displayed horizontally with values increasing from left to right. Furthermore, this chapter proposes a vector classification method, based on the principal component subspace, as a solution to the problem of the category 'unclassifiable' in the Astrup and Siggaard-Andersen classification scheme presented in Chapter 1. The essence of this new classification method is that in the principal component subspace vectors may be defined, representing the classification categories of the Astrup and Siggaard-Andersen scheme. Twelve such specific acid-base disorder vectors are derived. A classification can be made based on the angle of a patient vector with each of the twelve classification vectors. The disorder vector yielding the smaller angle determines the classification of the patient vector.

Chapter 4 describes the techniques to define patient-based multivariate reference regions in a principal component subspace. The assumption underlying the proposed technique is that a patient-based multivariate distribution of rotated principal component values is composed of a core of Gaussian distributed normal observations with contaminating abnormal ones located in the outer regions of the distribution. A trimming procedure is described to determine this core of Gaussian distributed observations. Having found this core of observations, multivariate reference regions may be defined using its multivariate mean and corrected variance-covariance matrix. Typically, resulting multivariate reference regions appear as ellipses in the proposed chart of Chapter 3. Multivariate abnormality of individual observations may also be quantitatively expressed by calculating the Mahalanobis distance of the observation from the mean of the distribution. Results of the trimming procedure for each of the six PCA transformed distributions are presented. Shape and location of the multivariate reference regions are largely dependent on the nature of an ICU.

In *Chapter 5*, two computer programs are described. The first program, called ABTRANS, can be used to derive multivariate acid-base reference regions and acid-base charts as described in the Chapters 2, 3, and 4. The output of this computer program is stored in specific initialisation files to be read by the second computer program, called ABCHART. This program plots acid-base data observations in the acid-base chart, calculates Mahalanobis distances, displays them in a trend plot, and classifies acid-base observations according to both the vector method and the Astrup and Siggaard-Andersen method. The ABCHART program can be linked to the information delivery system of a clinical chemistry department.

The practicability of the acid-base chart presented in Chapter 3 for the quantitative and qualitative monitoring of arterial pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ in an intensive care

environment is demonstrated in *Chapter 6*. Acid-base trajectories of three patients coming from the general adult ICU of the St. Elisabeth hospital (Tilburg) are displayed in the proposed chart and their information content is described. Furthermore, the tri-axial coordinate paper proposed by Hastings and Steinhaus in 1931 is presented in this chapter to illustrate its striking resemblance with the proposed chart for arterial pH, PaCO_2 and $\text{a}[\text{HCO}_3^-]$ for the ICU of the St. Elisabeth hospital.

Chapter 7 shows that the largest Mahalanobis distance as calculated from pH, PaCO_2 and BE in the first 24 hours after ICU admission is a good predictor of hospital mortality. The performance of the hospital mortality prediction models APACHE II, SAPS II, MPM_0 and MPM_{24} in predicting the hospital mortality of patients coming from the general ICU of the 'Onze Lieve Vrouwe Gasthuis' hospital (Amsterdam) is questionable. Although these models display sufficient discriminating power, calibration is only poor. Using multivariate logistic regression and cross-validation techniques, four different models are developed for predicting hospital mortality of ICU patients coming from the OLVG hospital. Three models are based on the largest absolute value determined within 24 hours after ICU admission for the deviation from the mean of a single acid-base variable (pH, PaCO_2 or BE). The fourth model is based on the largest Mahalanobis distance as calculated within the first 24 hours of ICU admission using the *OLVGbe* model of Chapter 4. Results show that the mortality prediction model based on the Mahalanobis distance is to be preferred. This model shows good discriminating power as well as good calibration. The model based on BE shows better discriminating power but calibration is poor.

Chapter 8 presents the results of a clinical evaluation of the *ELIbe* (see Chapter 4) multivariate acid-base reference model and tri-axial acid-base chart as conducted on the general ICU of the St. Elisabeth hospital (Tilburg). There are no differences between a period of model and chart use and a preceding baseline period for the following parameters: 1) the number of arterial acid-base analyses per patient, 2) the time between consecutive analyses, 3) the distance in the tri-axial chart between consecutive analyses, and 4) the percentage of arterial acid-base analyses with the patient being on artificial respiration. Comparing acid-base disturbance classifications of the vector method with the classifications of a 'silver' standard data set (*i.e.* the joint judgement of two senior clinicians) shows good agreement. However, the method of Astrup and Siggaard-Andersen shows a better agreement. There is a low agreement between the classifications of the 'silver' standard data set and the acid-base classifications given by the residents working on the ICU at the time of the study. Finally, the area under the Receiver-Operator-Characteristics (ROC) curve is equal for both reference models when this particular 'silver' standard (categorised into 'normal'/'abnormal') is used. This indicates that there is no difference between the models regarding the ability to

separate abnormal from normal arterial acid-base observations in this study population for this 'silver' standard.

9.2 General conclusions and future research

Although the study described in Chapter 7 indicates that the Mahalanobis distance may be preferred over the classical univariate approach when evaluating arterial acid-base data for the purpose of predicting hospital mortality, a superiority of the multivariate reference model in a clinical setting could not be demonstrated. This does not imply, however, that the concept and approach described in this thesis for generating and using multivariate reference regions for arterial acid-base variables in an intensive care setting is wrong. One important reason why we failed to demonstrate a superiority of the multivariate model over the univariate one may be the fact that the outcome of both models did not differ very much for the ICU of the St. Elisabeth hospital. Differences between both types of reference models could well be found for other ICUs since the nature of an ICU determines for a large part the shape and location of a multivariate reference region.

Further research may develop in four directions. First, multivariate reference regions for acid-base data may be compared with the univariate approach in other ICUs. Second, multivariate acid-base reference regions may be developed for specific kinds of ICU patients. The multivariate reference distributions used throughout this thesis are all based on acid-base data coming from randomly selected patients. Multivariate reference regions may, for example, be made specifically for ICU patients with chronic obstructive pulmonary disease (COPD) or for patients with renal problems. Such reference regions can then no longer be used as general 'normal' reference regions but they should rather be used as regions indicating what the typical values are for this kind of patient. Third, the multivariate acid-base reference model may be extended with other relevant clinical chemistry variables like PaO_2 , haemoglobin, lactate, etc.. It is recommended that the approach described in this thesis is followed, *i.e.* building multivariate reference regions in appropriate principal component subspaces to avoid the problem of data redundancy. Finally, the hospital mortality prediction model of Chapter 7 based on the acid-base Mahalanobis distance may be further developed. In this prediction model, the largest Mahalanobis distance in the first 24 hours of ICU admission is used as one of the predictor variables. One could, for instance, include the aspects of time and recovery into the model by taking the area under the curve of all acid-base Mahalanobis distances within the first 24 hours of admission. Also, other variables like specific haemodynamic variables known to be good predictors of hospital mortality may be added to the reference model. It is interesting to note that for these models the Mahalanobis distance is particularly attractive. Multiple variables

may be included into a multivariate reference model, while the multiple logistic regression model remains based on one Mahalanobis distance.

Regarding the classification of arterial acid-base data, it seems to be worthwhile to further develop the Astrup and Siggaard-Andersen classification method by defining classification rules for the 'unclassifiable' regions. The rather low agreement of the residents' acid-base classification with the joint judgement of senior ICU clinicians as demonstrated in Chapter 8, clearly indicates the benefit that a computerised arterial acid-base classification system would have for this particular ICU. Based on the results of Chapter 8, the best results may be achieved using the tri-axial chart with the Astrup and Siggaard-Andersen classification areas drawn in. The normal area may be indicated in this chart with the optimal reference multivariate acid-base reference ellipse as determined in Chapter 8.

The experience obtained during the clinical evaluations described in this thesis may be of value when considering setting up similar studies in the future and it is therefore relevant to discuss this at this point. The primary goal of the research described in this thesis was to investigate whether a statistical and chemometric analysis of acid-base measurements of large groups of patients would yield general rules allowing an objective and reproducible analysis of the acid-base status of an individual patient. Preferably, this analysis should be equal to or even be better than the analyses of a well-trained experienced clinician so that it can be of practical use for the lesser experienced clinician during acute situations at the bedside of a patient. The results of the evaluations described in Chapters 7 and 8 generally confirm this supposition.

The implementation, however, of this theoretical statistical concept into daily clinical practice proved to be troublesome. Before a clinician is willing to replace existing and familiar techniques with newer ones originating from less familiar disciplines like clinical chemistry and medical informatics, he wants to be confident that these new techniques will not harm the individual patient in any way. This means, particularly in an experimental design, that the new model can only be tested in situations in which the results one wants to obtain, are already known and judged positively. Moreover, most clinicians feel that the alternative technique, *i.e.* the current method of interpreting acid-base measurements, performs quite adequately. The observations regarding the performance of residents, as described in Chapter 8, questions this belief.

The results of the research described in this thesis show that chemometric and statistical methods may indeed attribute useful information to basic clinical chemistry parameters. Whether this information will be used in the future depends, for a large part, on the success of convincing the clinician beforehand of the gain in quality that can be obtained using the presented methods and pertinent new parameters.

10

Samenvatting, algemene conclusies en verder onderzoek

10.1 Samenvatting

Dit proefschrift heeft als onderwerp de evaluatie en classificatie van arteriële zuur-base metingen op een afdeling van intensieve zorg (Intensive Care Unit, ICU) met behulp van chemometrische technieken. In *Hoofdstuk 1* wordt een aantal basisprincipes behandeld van de humane zuur-base fysiologie en uiteengezet hoe op dit moment zuur-base metingen worden geëvalueerd en geïnterpreteerd op een ICU. Er bestaan twee 'scholen' van zuur-base analyse en interpretatie. Volgens de ene 'school' vormen de arteriële zuurgraad pH, de arteriële zuurstofspanning PaCO_2 en de arteriële bicarbonaat-ion concentratie $\text{a}[\text{HCO}_3^-]$ de basis voor een zuur-base interpretatie. De andere 'school' propageert het gebruik van arteriële pH, PaCO_2 en een afgeleide parameter genaamd base excess (BE). Een strikte toepassing van de classificatie methode voor deze laatste 'school', zoals door Astrup en Siggaard-Andersen ontwikkeld, leidt in principe tot een categorie 'niet classificeerbaar'. Daarnaast stuit het gebruik van standaard 95% referentie intervallen voor de respectievelijke zuur-base variabelen op fundamentele problemen. Voor beide 'scholen' geldt ten eerste dat er een (nagenoeg) lineair verband bestaat tussen de drie gebruikte zuur-base variabelen. Dit betekent dat één van de drie variabelen overbodig is. Ten tweede kan, vanuit een statistisch oogpunt gezien, het gelijktijdig gebruik van meer dan één standaard 95% referentie interval leiden tot foute conclusies. Dit proefschrift beschrijft een methode voor het genereren van valide multivariate zuur-base referentie modellen welke te gebruiken zijn in een intensive care omgeving. Bovendien beschrijft het een methode voor het classificeren van arteriële pH, PaCO_2 en BE waarden waarin per definitie de categorie 'niet classificeerbaar' afwezig is. De laatste twee hoofdstukken van dit proefschrift beschrijven de resultaten van een eerste klinische evaluatie van het multivariate referentie model en de nieuwe classificatie methode.

De essentie van het referentie model is dat een multivariaat referentie gebied wordt afgeleid van een grote verdeling van arteriële zuur-base metingen (pH, PaCO_2 en $\text{a}[\text{HCO}_3^-]$ of pH, PaCO_2 en BE) van intensive care patiënten zelf. De eerste stap in dit proces is het verwijderen van de redundantie die in dergelijke zuur-base verdelingen aanwezig is. In *Hoofdstuk 2* wordt beschreven hoe met behulp van een principale component analyse (PCA) deze redundantie kan worden verwijderd. De resultaten van PCA transformaties op zes verschillende zuur-base verdelingen van vier verschillende intensive care afdelingen worden gepresenteerd. Deze resultaten laten zien dat voor elke PCA getransformeerde zuur-base verdeling meer dan 99% van de totale oorspronkelijke variantie beschreven wordt door de eerste twee principale component assen. PCA zou hiermee een waardevolle techniek kunnen zijn bij het valideren van arteriële zuur-base metingen. Combinaties van zuur-base metingen zouden, gegeven een PCA getransformeerde verdeling, gecontroleerd kunnen worden op interne consisten-

tie door de berekening van de derde principale component. Deze waarde moet dicht genoeg bij het waargenomen gemiddelde van de derde principale component van de referentie verdeling liggen om de combinatie als valide te kunnen beschouwen.

Gebaseerd op de PCA resultaten van Hoofdstuk 2 beschrijft *Hoofdstuk 3* een nieuwe grafische representatie voor de combinatie pH, PaCO_2 en $\text{a}[\text{HCO}_3^-]$ of de combinatie pH, PaCO_2 en BE. Aangezien het vlak beschreven door de eerste twee principale component assen na PCA nagenoeg alle variantie in de initiële data set verklaart, is het mogelijk zuur-base waarnemingen in dit vlak te presenteren zonder verlies van informatie. Om een standaard configuratie van de grafiek te verkrijgen wordt een zodanige rotatie op het vlak toegepast dat de pH-as altijd horizontaal staat met oplopende waarden van links naar rechts. In dit hoofdstuk wordt bovendien een vector classificatie methode gepresenteerd, gebaseerd op het geroteerde principale component vlak, als oplossing voor het probleem van de 'niet classificeerbare' categorieën in de methode van Astrup en Siggaard-Andersen zoals beschreven in Hoofdstuk 1. De essentie van deze nieuwe classificatie methode is dat in het geroteerde principale component vlak vectoren kunnen worden gedefinieerd welke overeenkomen met de classificatie categorieën in de methode van Astrup en Siggaard-Andersen. Er bestaan derhalve twaalf van zulke classificatie vectoren. Een classificatie is gebaseerd op de hoek die de patiënt-vector maakt met elk van de twaalf classificatie vectoren. De classificatie vector met de kleinste hoek bepaalt de uiteindelijke classificatie categorie van de patiënt-vector.

Hoofdstuk 4 beschrijft de technieken voor het afleiden van patiënt-gebaseerde multivariate referentie gebieden in het geroteerde principale component vlak. De assumptie bij de voorgestelde methodiek is dat de multivariate verdeling is opgebouwd uit een kern van Gauss-verdeelde normale waarnemingen met abnormale waarnemingen aan de buitenkanten van de verdeling. Een 'trim' procedure wordt geschreven waarmee deze Gauss-verdeelde kern van waarnemingen kan worden bepaald. Multivariate referentie gebieden kunnen vervolgens worden gedefinieerd met het multivariate gemiddelde en de gecorrigeerde variantie-covariantie matrix van de Gauss-verdeelde kern. Deze referentie gebieden hebben de vorm van een ellips wanneer ze in de grafiek van Hoofdstuk 3 worden afgebeeld. De multivariate abnormaliteit voor individuele observaties kan ook kwantitatief worden uitgedrukt middels het berekenen van de Mahalanobis afstand van de observatie tot het gemiddelde van de Gauss-verdeelde kern. De resultaten van deze procedure voor de zes PCA getransformeerde verdelingen worden gepresenteerd. De vorm en de locatie van de multivariate referentie gebieden zijn met name afhankelijk van de patiënten populatie op een ICU.

In *Hoofdstuk 5* worden twee computer programma's beschreven. Het eerste programma (ABTRANS) kan gebruikt worden voor het bepalen van multivariate zuur-base referentie gebieden en de constructie van de nieuwe grafische representatie zoals beschreven in de Hoofdstukken 2, 3 en 4. De resultaten van dit programma kunnen

worden bewaard in speciale initialisatie bestanden die vervolgens kunnen worden uitgelezen met het tweede computer programma (ABCHART). Dit laatste programma plot zuur-base observaties in de grafische representatie van Hoofdstuk 3, berekent Mahalanobis afstanden, toont deze afstanden in een tijd plot en classificeert zuur-base observaties volgens de vector methode en de methode van Astrup en Siggaard-Andersen. Het ABCHART programma kan gekoppeld worden aan het informatie systeem van een klinisch chemische afdeling of laboratorium.

De praktische bruikbaarheid van de tri-axiale zuur-base grafiek uit Hoofdstuk 3 voor het kwalitatief en kwantitatief monitoren van pH , PaCO_2 en $\text{a}[\text{HCO}_3^-]$ in een intensive care omgeving wordt beschreven in *Hoofdstuk 6*. Het zuur-base verloop van drie patiënten van de volwassenen ICU van het St. Elisabeth ziekenhuis te Tilburg wordt beschreven en afgebeeld in de tri-axiale grafiek voor deze ICU. In dit hoofdstuk wordt bovendien het tri-axiale coördinaat papier, zoals geïntroduceerd door Hastings en Steinhaus in 1931, beschreven om de sterke gelijkenis te tonen met de tri-axiale grafiek voor de ICU van het St. Elisabeth ziekenhuis.

Hoofdstuk 7 toont aan dat de grootste Mahalanobis afstand, berekend uit pH , PaCO_2 en BE in de eerste 24 uur na ICU opname, een goede voorspeller is van ziekenhuis mortaliteit. De prestaties van de mortaliteit predictie modellen APACHE II, SAPS II, MPM_0 en MPM_{24} om de mortaliteit te voorspellen van ICU patiënten van het 'Onze Lieve Vrouwe Gasthuis' te Amsterdam zijn matig. De modellen hebben een goed onderscheidend vermogen maar de calibratie is slecht. Met behulp van logistische regressie en cross-validatie technieken zijn vier nieuwe predictie modellen ontwikkeld en getest. Drie van de vier modellen zijn gebaseerd op de grootste absolute waarde gemeten binnen 24 uur na ICU opname voor een enkele zuur-base variabele (pH , PaCO_2 of BE). Het vierde model is gebaseerd op de grootste Mahalanobis afstand in de eerste 24 uur na ICU opname zoals berekend met het *OLVGbe* model van Hoofdstuk 4. De resultaten tonen aan dat het predictie model gebaseerd op de Mahalanobis afstand te prefereren is; zowel het onderscheidend vermogen als de calibratie is goed. Het predictie model gebaseerd op BE heeft weliswaar een beter onderscheidend vermogen maar de calibratie is slecht.

Hoofdstuk 8 presenteert de resultaten van een klinische evaluatie van het *ELIbe* multivariate zuur-base referentie model (zie Hoofdstuk 4) en de tri-axiale grafiek, uitgevoerd op de ICU van het St. Elisabeth ziekenhuis te Tilburg. Er zijn geen verschillen tussen een periode van computer met grafiek gebruik en een voorafgaande 'baseline' periode voor de volgende parameters: 1) het aantal arteriële zuur-base observaties per patient, 2) de tijd tussen twee opeenvolgende arteriële zuur-base observaties, 3) de afstand in de tri-axiale grafiek tussen opeenvolgende arteriële zuur-base observaties, en 4) het percentage arteriële zuur-base observaties waarbij de patiënt kunstmatig beademd wordt. De zuur-base classificaties gemaakt met de vector methode komen goed

overeen met de classificaties van een 'zilveren' standaard, zijnde de gezamenlijke beoordeling van twee ervaren ICU artsen. Classificaties gemaakt met behulp van de As-trup en Siggaard-Andersen methode komen echter beter overeen met de 'zilveren' standaard. De overeenkomst tussen de classificaties van de 'zilveren' standaard en de classificaties van arts-assistenten is slecht. Tenslotte, de oppervlakte onder de Receiver-Operator-Characteristics (ROC) curve is voor beide referentie modellen even groot voor deze 'zilveren' standaard (gecategoriseerd naar 'normaal'/'abnormaal'). Dit betekent dat er geen verschil is tussen beiden modellen wat betreft het onderscheiden van normale en abnormale zuur-base observaties in deze studie populatie voor deze 'zilveren' standaard.

10.2 Algemene conclusies en verder onderzoek

Alhoewel het onderzoek beschreven in Hoofdstuk 7 aangeeft dat de Mahalanobis afstand te prefereren is boven de klassieke univariate benadering als het gaat om het evalueren van zuur-base gegevens ten behoeve van het voorspellen van ziekenhuis mortaliteit, kon in een klinische evaluatie het voordeel van een multivariate zuur-base referentie model boven een univariate benadering niet worden aangetoond. Dit betekent echter niet dat het concept en de aanpak zoals beschreven in dit proefschrift voor het genereren van multivariate zuur-base referentie gebieden fout is. Een van de belangrijkste redenen waarom het niet gelukt is om de meerwaarde van een multivariaat referentie model aan te tonen is het feit dat er geen groot verschil was tussen de multivariate Mahalanobis afstand en de univariate abnormaliteit index voor de ICU van het St. Elisabeth ziekenhuis (Hoofdstuk 8). Dergelijke verschillen kunnen echter wel op andere ICU's aanwezig zijn omdat de aard van een ICU voor een groot gedeelte de vorm en locatie van een multivariate patiënt-gebaseerde zuur-base referentie gebied bepaalt.

Eventueel verder onderzoek kan in principe in vier richtingen worden voortgezet. Allereerst, het multivariate zuur-base referentie model beschreven in dit proefschrift zou kunnen worden vergeleken met het multivariate model in andere ICU's. Een tweede richting zou kunnen zijn om modellen te bouwen en te evalueren voor speciale groepen ICU patiënten. De multivariate modellen die gebruikt zijn in dit proefschrift zijn gebaseerd op referentie populaties die willekeurig zijn samengesteld omdat er geen patiënt selectie criteria zijn toegepast. Een multivariaat zuur-base referentie model zou kunnen worden afgeleid van de zuur-base metingen van bijvoorbeeld ICU patiënten met chronische luchtweg problemen (COPD) of patiënten met nierproblemen. Uiteraard kunnen de resulterende referentie gebieden dan niet meer worden gezien als gebieden die 'normaliteit' aangeven maar eerder als gebieden die aangeven wat de typische zuur-base waarden zijn voor dit soort patiënten. De derde richting zou de uitbreiding van het multivariate zuur-base referentie model kunnen zijn met andere rele-

vante klinisch chemische variabelen zoals PaO_2 , hemoglobine, lactaat, e.d.. De benadering zoals in dit proefschrift is beschreven is hierbij aan te bevelen. Dat wil zeggen, referentie gebieden definiëren op sub-coördinaat stelsels van principale componenten zodat rekening wordt gehouden met eventuele redundantie in de oorspronkelijke data set. Als laatste richting zou men het mortaliteit predictie model zoals beschreven in Hoofdstuk 7 verder kunnen ontwikkelen. In dit model is de grootst waargenomen Mahalanobis afstand binnen 24 uur na ICU opname gebruikt als een voorspellende variabele. Men zou het aspect van tijd en herstel in het model kunnen incorporeren door de oppervlakte onder de curve van alle Mahalanobis afstanden in de eerste 24 uur na ICU opname als voorspellende variabele toe te voegen. Bovendien zouden andere variabelen waarvan bekend is dat ze een goed voorspellend vermogen hebben kunnen worden toegevoegd. Juist voor dit soort uitbreidingen is het multivariate referentie model geschikt. Immers, het aantal variabelen in een multivariaat referentie model kan variëren maar het mortaliteit predictie model blijft simpel omdat het nog steeds gebaseerd is op één enkele Mahalanobis afstand.

Wat betreft de classificatie van arteriële zuur-base variabelen lijkt de verdere ontwikkeling van de Astrup en Siggaard-Andersen methode door classificatie regels voor de 'onclassificeerbare' gebieden te ontwikkelen, nuttig. De lage overeenkomst tussen de zuur-base classificaties van arts-assistenten en de classificaties van twee ervaren ICU artsen zoals beschreven in Hoofdstuk 8, demonstreren het potentiële nut van een computer programma voor het interpreteren van zuur-base gegevens op deze specifieke ICU. Gebaseerd op de resultaten van Hoofdstuk 8 zouden de beste resultaten voor deze ICU gehaald kunnen worden door gebruik te maken van de tri-axiale grafiek met daarin de Astrup en Siggaard-Andersen gebieden. Het 'normaal' gebied voor zuur-base waarnemingen zou hierin kunnen worden gemarkeerd met de optimale multivariate referentie ellips zoals bepaald in Hoofdstuk 8.

Enkele ervaringen die zijn opgedaan tijdens de praktijkevaluaties zoals beschreven in dit proefschrift, zijn mogelijk van belang bij de opzet van vergelijkbare studies in de toekomst en daarom in dit bestek vermeldenswaard. Uitgangspunt voor het in dit proefschrift beschreven onderzoek was dat door statistische en chemometrische analyse van combinaties van zuur-base gegevens van grote groepen patiënten, het mogelijk zou zijn om daaraan wetmatigheden te ontleen die konden worden gebruikt voor een ge-objectiveerde en reproduceerbare analyse van de zuur-base toestand van de individuele patiënt. Die analyse moest idealiter van gelijke of hogere kwaliteit zijn dan die van de ervaren clinicus en in die hoedanigheid een betrouwbare en praktische steun voor de minder ervaren clinicus in de actuele situatie aan het bed van de patiënt. De uitkomsten van de evaluaties (Hoofdstuk 7 en 8) bevestigden in grote trekken die vooronderstelling.

De weg van dit theoretisch statistisch concept naar een praktische toepassing in de kliniek is echter zeer moeizaam gebleken en eigenlijk ook gebleven. Voor een clinicus bereid is om vertrouwde werkwijzen te verlaten en te vervangen door nieuwe die ontleend zijn aan wellicht onbekende statistische technieken en bovendien vanuit andere, voor de clinicus minder vertrouwde vakgebieden als de klinische chemie en de medische informatica, worden aangereikt, zal hij eerst overtuigd moeten zijn dat de individuele patiënt op wie de technieken worden toegepast, daar op geen enkele manier nadelige effecten van ondervindt. In feite kan men, juist in een experimentele opzet, het model pas in de klinische praktijk toepassen als de resultaten die men met het onderzoek wilt bereiken, al bekend en positief uitgevallen zijn. Daar komt bij dat in de beleving van veel clinici het alternatief, de gevestigde methode voor interpretatie van zuurbase gegevens, redelijk tot goed verloopt. De observaties met betrekking tot de prestaties van arts-assistenten, zoals beschreven in Hoofdstuk 8, plaatsen echter vraagtekens bij die veronderstelling.

De resultaten van het in dit proefschrift beschreven onderzoek laten zien dat met chemometrische en statistische analysemethoden wel degelijk waardevolle informatie toegevoegd kan worden aan basale klinisch chemische onderzoekparameters. Of die informatie in de toekomst benut gaat worden, zal in hoge mate afhangen van de mate waarin de clinicus op voorhand overtuigd kan worden van de winst in kwaliteit van de informatie die met de gepresenteerde analysemethoden en daaraan ontleende nieuwe parameters, te bereiken valt.

Dankwoord

Eindelijk mag ik 'm dan schrijven, het dankwoord. Allereerst dank ik mijn promotor prof.dr. E.S. Gelsema. Beste Edzard, tijdens het begin van mijn onderzoek hoorde ik collega's klagen over de schaarse contacten die zij hadden met hun begeleiders. Dat kwam mij geheel onbekend voor want in al die jaren is het niet één keer gebeurd dat je geen tijd voor me had. Altijd kon ik bij je binnenstappen om problemen te bepraten. Dat sommige problemen om een iets minder mathematische aanpak vroegen doet daar niets aan af. Mijn dank gaat ook uit naar mijn co-promotor dr. J. Lindemans. Beste Jan, ook jij had altijd wel even tijd als ik bij je aanklopte. Vooral gezien jouw tijdroevende functie is dat opmerkelijk te noemen.

Een deel van het onderzoek vond plaats in het St. Elisabeth ziekenhuis te Tilburg en het OLVG te Amsterdam. Het zijn met name dr. H.M.J. Goldschmidt en dr. R.N.M. Weijers geweest die daar de kar op de rails hebben gezet. Beste Henk en Rob, ik weet 100% zeker dat zonder jullie enthousiasme en inzet dit proefschrift er nooit was gekomen. Mijn dank gaat ook uit naar het verplegend personeel en de medische staf van IC5 van het St. Elisabeth ziekenhuis. Herman Ulenkate en Jeroen van Leeuwen mag ik hier ook zeker niet vergeten. Ik wil al mijn collega's van de afdeling Medische Informatica bedanken maar speciaal Bob Schijvenaars en Peter Rijnbeek.

Ook alle leden van het dispuut *Nico Bronsgeest Goes Latin* ben ik dank verschuldigd. Op onze bijeenkomsten heb ik altijd ruimschoots geestelijk kunnen bijtanken. Als voorzitter zijnde spreek ik de hoop uit dat we nog menig maal bij elkaar zullen komen.

Iemand die mij vooral tijdens de laatste jaren van mijn onderzoek heeft meegemaakt (en daarom niet te benijden is) is Natasja Bouwer. Lieve Tas, dat ik bezig was met promoveren trok niet alleen een zware wissel op mijzelf maar zeker ook op jou. Ik wil je bedanken voor je geduld en je begrip maar ook voor je aanmoedigingen en aansporingen om het vooral toch maar af te maken.

Tja, en dan natuurlijk de familie Hekking (hieronder reken ik ook Ben van der Lubbe en Rob en Cora Landman). Ik kan gerust zeggen dat ieder van jullie een steentje, weliswaar variërend in grootte en aard, hebben bijgedragen aan mijn persoonlijke ontwikkeling en daarom indirect aan de totstandkoming van dit proefschrift. Daarvoor dank. Extra dank gaat uit naar mijn zwager John Zuiderduin. Johnie-boy, je kent me als geen ander. Van de spelletjes memory tot de allerlaatste ontwikkelingen. Je hebt me bijgestaan – gevraagd, soms ongevraagd – op velerlei gebied en ik weet zeker dat je dat altijd zult blijven doen. Als laatste wil ik mijn vader en moeder bedanken. Ze hebben mij een onbekommerde jeugd gegeven en ik draag daarom dit proefschrift aan hen op.

september 1999



Curriculum vitae

Marcel Hekking was born on October 2, 1965 in 's Gravenhage. He attended VWO from 1979 to 1985 at the Edith Stein college in 's Gravenhage. In that year he started the study Healthsciences at the Rijksuniversiteit Limburg in Maastricht (presently known as Universiteit van Maastricht) with Biological Healthsciences as main subject. During his final period of practical training at the DNA-laboratory of the Biomedical Centre in Maastricht he became interested in medical informatics. In 1989, he graduated and in 1991 he started his PhD research at the department of Medical Informatics of the Erasmus University Rotterdam of which the results are described in this thesis. He is currently working as a medical writer and statistician at IMRO TRAMARKO, an international Contract Research Organisation (CRO).

