Chapter 1

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

1.1 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 $a[HCO_3^-]/CO_2$ 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 ($a[HCO_3^-]$) 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 form north-west to south-east [11, 12].

However, a major disadvantage of most acid-base representations, whether based on BE or $a[HCO_3^-]$, 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 $a[HCO_3^-]$/BE values 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.

1.2 Methods

1.2.1 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.

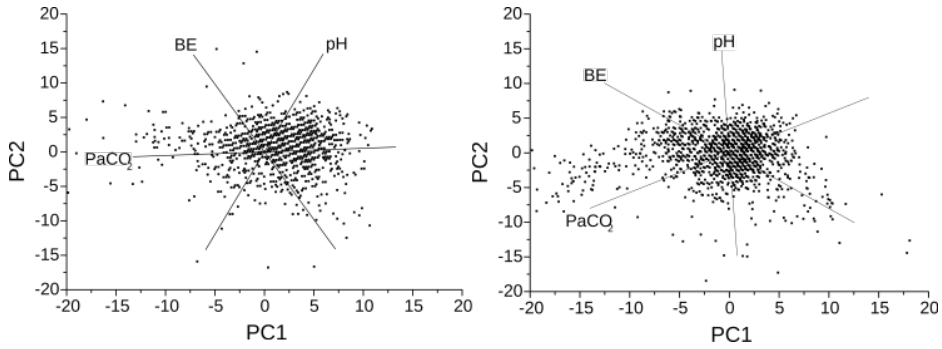

Figure 3–1. Original acid-base axes, projected on the PC1-PC2 subspaces for data sets AZRbe (left) and OLVGbe (right). The placement of the original acid-base axes is different as a result of different covariances in the original distributions.

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 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.

$$T = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} *\mathrm{U} \qquad (3\text{–}1)$$

in which 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 coordinate values of the acid-base observation in the proposed acid-base

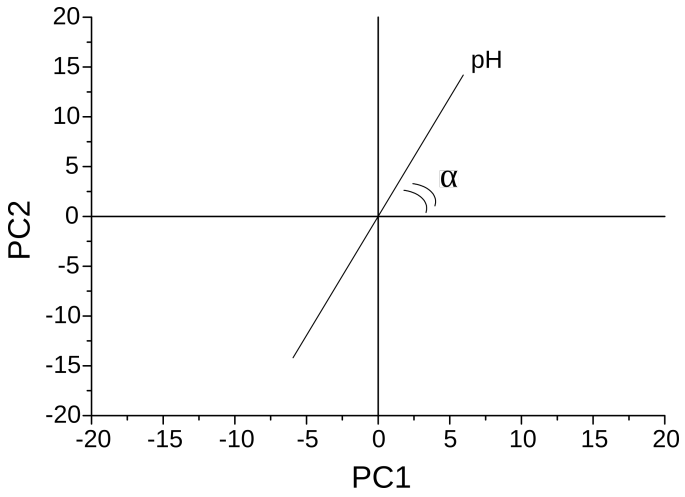

Figure 3–2. The angle between the projected original pH-axis and the PC1-axis in the PC1-PC2 subspace of data set AZRbe.

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$[HCO_3^-]$ (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];

$e_{31}\ x + e_{32}\ y + e_{33}\ z = 0$, (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 Error: Reference source not found. 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.

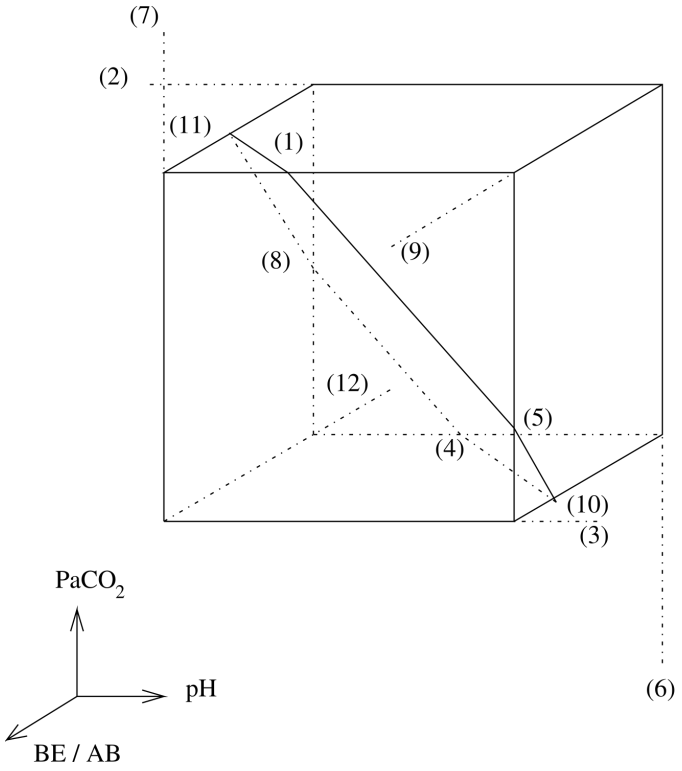

Figure 3–3. Intersection of a PC1-PC2 subspace with a 95% univariate reference cube. Arrows point into the direction of increasing values.

1.2.2 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 existence 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 PC1' and 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.

1.3 Results

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 $a[HCO_3^-]$ 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 in other familiar acid-base charts. In Figure 3–4, the pH axis runs along the abscissa, while the $PaCO_2$ and $a[HCO_3^-]$ axes are placed at an angle close to 60°, resulting in a tri-axial placement of the original acid-base axes.

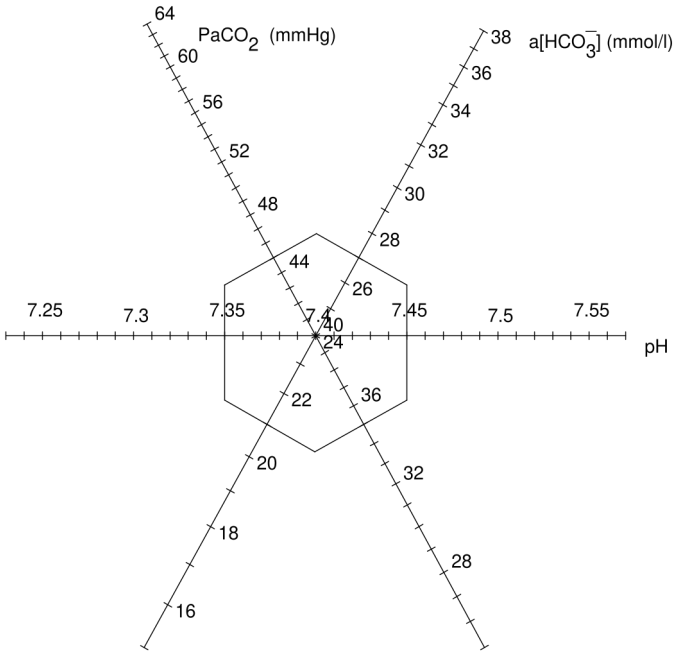

Figure 3–4. The proposed chart for pH, $PaCO_2$ and $a[HCO_3^-]$, as built from the data in the ELIab data set.

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 PC1' and 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.

1.3.1 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:

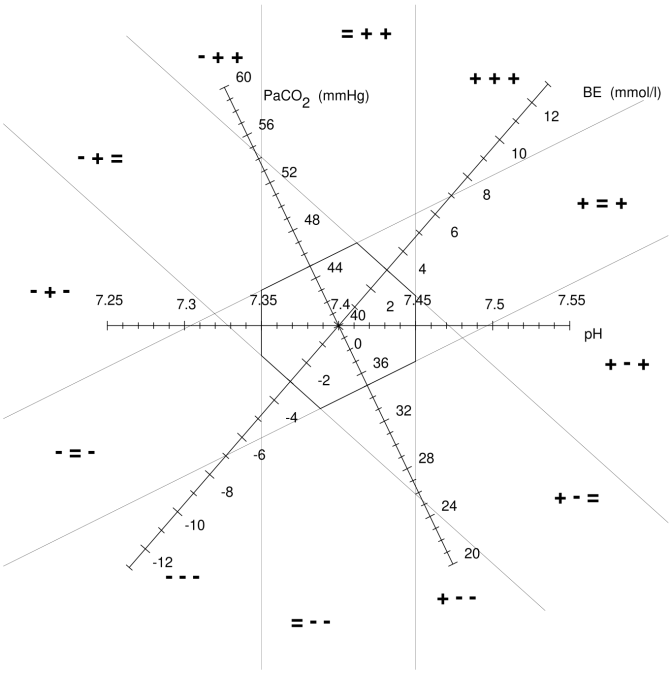

Figure 3–5. The proposed chart for pH, $PaCO_2$ and BE as built from the data in the ELIbe data set. The classical acid-base classification from Table 1–2 are plotted in (dotted lines).

- 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 regions, 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:

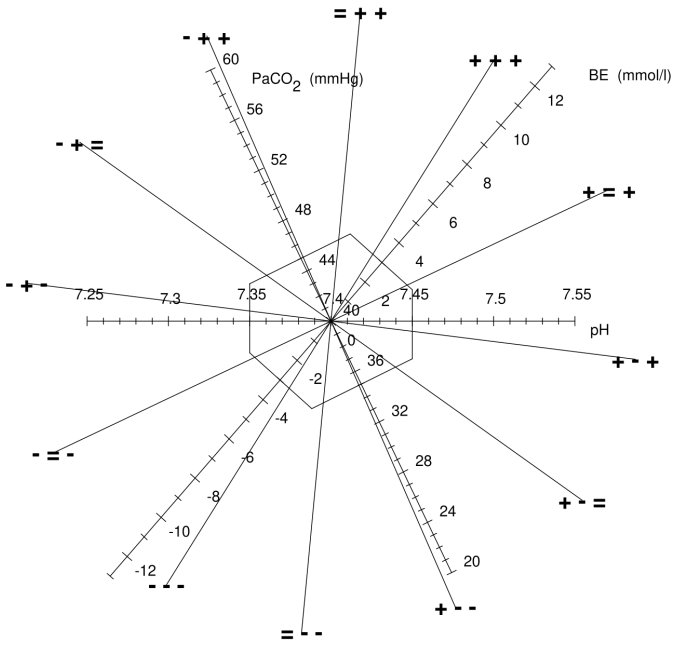

Figure 3–6. The ELIbe chart with the new vector classification scheme plotted in (dotted lines).

- 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 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.

1.4 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.

1.5 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, Jörgensen 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.

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. 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.