

# Intensive care monitoring: past, present and future

Neil McIntosh

**ABSTRACT** – Monitoring is the serial evaluation of time-stamped data, and the volume of such data in an intensive care unit is huge. Clinical and biochemical data may be available at hourly or more frequent intervals but physiological data are 'continuous'. Although sophisticated monitors display the physiological data in multiple and varied combinations, staff are challenged by the frequency of the false alarms and a lack of knowledge of the patterns from which they could predict problems. All these data, together with large amounts of clinical data, lead to information overload. In this paper, the case is made for the development of automatic decision-support systems based on statistical and probabilistic analysis of data patterns appropriate for the level of cognition of the user (nurses and juniors at the bedside rather than consultants). Such decision support could both reduce the false-positive alarms that frustrate clinical staff, and improve the early detection of pathophysiological events. We have used the development of a pneumothorax as our paradigm. Our data indicate that the clinical diagnosis of pneumothorax takes a median of 127 minutes, but using short decision algorithms based on routinely available monitoring data, most can be detected within 10–15 minutes of occurrence.

**KEY WORDS:** alarms, artefact, cognition, decision support, intensive care, monitoring, multidimensional space, newborn, pattern recognition, pneumothorax

## Introduction

Monitoring can be defined as the serial evaluation of time-stamped data. This type of evaluation can be helpful for both diagnosis and prediction of pathophysiological events. Such data are particularly common in the intensive care unit where there are rich streams of clinical data, eg the hourly volume of urine; repeated abdominal circumference measurements; biochemical or haematological data such as the sequential measurements of blood gases or platelets; and physiological data on monitors at the patient's bedside. Physiological monitors may give data at relatively low rates of transmission, eg with

heart rate, blood pressure, blood carbon dioxide etc occurring at one-second intervals, or there may be high rates of transmission at many times a second to give the electrocardiogram (ECG) or blood pressure waveform.

Historically, the first monitoring of patients using medical equipment was almost certainly with the clinical thermometer. This proved to be highly successful as certain diseases have specific temporal temperature patterns – thus relapsing fever presents in a patient with a high fever for a few days followed by a crisis when the temperature falls to normal for three to four days before a further feverish period. In contrast, patients with malaria may have a temperature of 102–104°F each evening with a normal morning temperature. Such fever patterns may be quite predictive of the underlying pathology. The first use of monitoring in intensive care was carried out by the neurosurgeon Harvey Cushing. In a classic example shown in Fig 1, he monitored blood pressure and heart rate before, during and after an operation on a ship's carpenter who had been crushed between a beam and the ship he was working on<sup>1</sup>. The carpenter was admitted to the local hospital in Boston with high blood pressure and slow pulse and was comatose. The clinical and neurological picture was that of an extradural haemorrhage causing raised intracranial pressure, so he was taken to theatre and a craniotomy was performed. The stages of the operation can be seen on the figure, where before and during the operation the time points are at 10-minute intervals. The already high blood pressure rises even higher with the introduction of the ether anaesthetic; the skull is opened and when the clot is evacuated the blood pressure falls and the pulse rate rises. As the patient comes round from the anaesthetic, the pulse rate rises further and the blood pressure falls further and then both are stable in the postoperative period.

We now know that this pattern of high blood pressure and low pulse rate is classical of raised intracranial pressure. There are other classical patterns that are taught in medical school – the rising pulse rate and falling blood pressure associated with haemorrhage, the fall-off in linear growth after birth in the pituitary dwarf, or the fall-off in weight or rate of weight gain with the introduction of gluten in a child who has coeliac disease. What has never been

This article is based on the Teale Lecture given at the Royal College of Physicians on 29 October 2001 by **Neil McIntosh** DSc(Med) FRCP, Professor of Child Life and Health, University of Edinburgh

*Clin Med JRCPL*  
2002;2:349–55

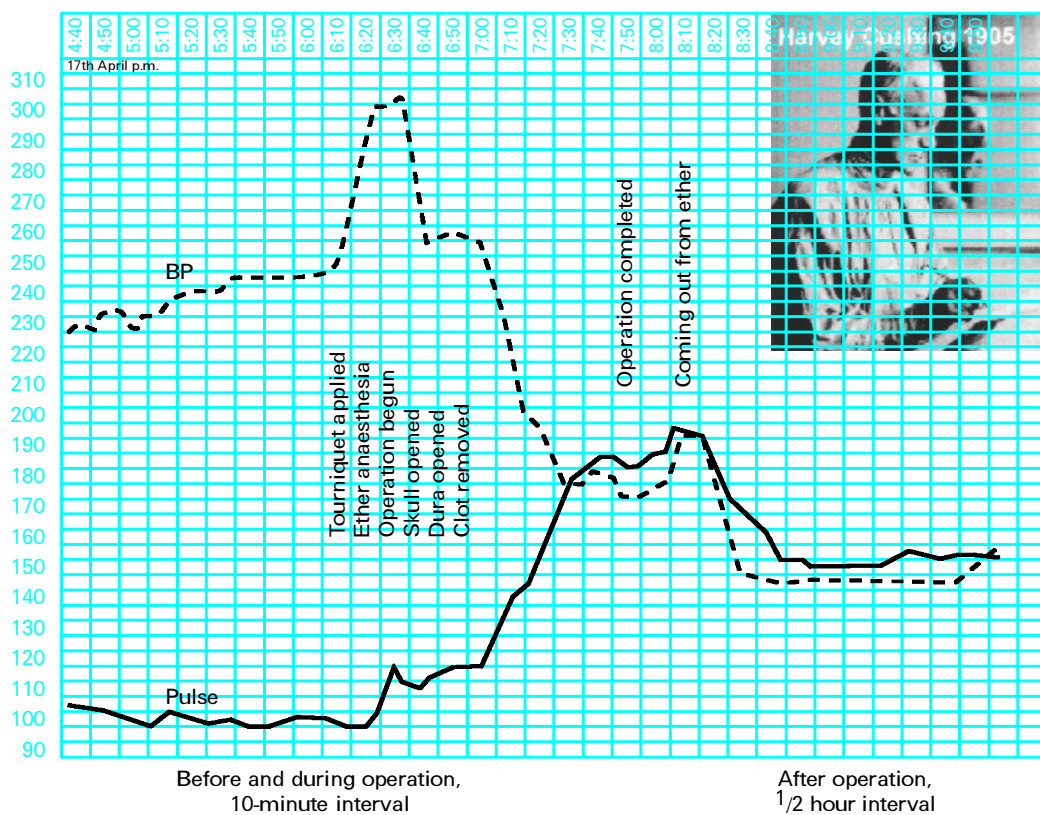


Fig 1. Operative record of craniotomy for extradural haematoma by Harvey Cushing (1905).

defined is the sensitivity, specificity, positive predictive value and negative predictive value of these patterns.

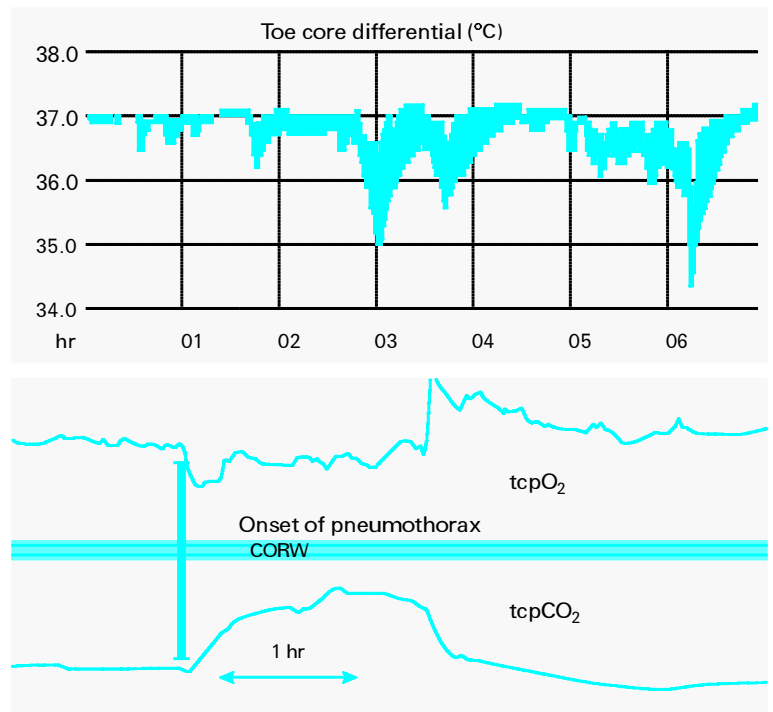
If these experiential heuristics taught in medical school are to be based on evidence, we need to know the reproducibility of these patterns. And that is where our own work has begun, in relation to high density time-stamped data in a neonatal intensive care unit.

### Why do we monitor in the intensive care unit?

If the digital read-out on a monitor presents us with a value (such as a mean blood pressure of 60 mmHg), in a full-term newborn infant we can be reassured that this variable is normal, and if the monitor has limit or threshold alarms we can also presume that we will have some warning of abnormality when the blood pressure exceeds or falls below the set limits. If, on the other hand, we know from a presentation of trend that the mean blood pressure is 60 mmHg and is slowly increasing, we have much more information. First, we know not only that the baby is normal but also that s/he is stable or even improving. We may also have the potential for earlier warning by seeing that there is a steady fall of the blood pressure and we can be alerted well before it reaches the lower threshold alarm limits. We can also see the effectiveness of treatment with the dopamine infusion bringing the low blood pressure up to the upper end of the normal range.

### Trend patterns in physiological data

For 15 years we have been trying to ascertain trend patterns in physiological data. Thus on the cot-side screen we may have the heart rate, the mean blood pressure, the transcutaneous oxygen, the transcutaneous carbon dioxide, and the temperatures. We have been attempting to identify patterns within this data. Figure 2a shows the pattern of cold stress<sup>2</sup>, seen within days of instituting our trend monitoring system. More difficult, but fairly consistent, has been the picture of the occurrence of a pneumothorax where the transcutaneous carbon dioxide level sharply increases usually at the same time as a sharp decrease in the transcutaneous oxygen<sup>3</sup> (Fig 2b). We believed that our system helped in management but wish to see whether there was any difference in outcome. We designed a pragmatic trial, to assess whether the availability of trends at the cot side improved patient outcome. We randomly allocated into the trial 600 consecutive neonates requiring intensive care over the first seven days of life, either with a trend monitor display next to the baby or without one<sup>4</sup>. We had a number of outcome measures – short-term, medium-term and long-term – but we were not able to show any important differences between the groups. In retrospect, we believe that because the randomisation only applied during the first seven days of life, that the many days and sometimes weeks after this that the babies spent in intensive care on our unit saturated out any differences between the groups.



**Fig 2. Physiological data patterns seen in a neonatal ICU. (a)** Pattern of cold stress in 600 g baby with diverging peripheral and central temperatures when incubator doors are opened for necessary care procedures. **(b)** The development of a pneumothorax in a ventilated preterm infant, with successful drainage after 2 hours when the diagnosis is made.

However, other reasons why trend monitoring used in this way does not make a difference to outcome could be data overload, an inability to ascertain which data are important, or simply that staff are ignorant of the patterns that exist.

To test the last, we set up a project to explore the cognitive processes of doctors and nurses in interpreting physiological monitoring data, testing how junior doctors, senior doctors and nurses used, viewed and understood data.

Fourteen trend events, each lasting 2 hours, were displayed in 7-minute blocks to each member of staff. An event might be: the effect of endotracheal suction, the effect of dopamine infusion, or the occurrence of a pneumothorax etc. Staff were asked to review the data, to describe them, pointing out on the screen any abnormality detected. The staff member and the computer screen were videoed from behind and audiotaped. The tapes were then transcribed and each comment labelled as: a description, an interpretation, an uncertainty, a hypothesis, a relationship amongst different parameters, or an artefact. Overall, the junior doctors and nurses gave more descriptions, whereas senior doctors noted more relationships and generated more hypotheses. The results are shown in Table 1. The seven senior doctors generated many more statements and these statements tended to be hypotheses to explain the data. The senior doctors were able to make a diagnosis in real time in 68% of cases, the junior doctors in 58%, and the nurses in only 25%. When the full 2-hour trace was revealed, senior doctors made correct final diagnoses in 80% of events, junior doctors 78%, but nurses only 32% – the last being the group of staff permanently with the babies. Our conclusions were that the patterns in multichannel trended physiological data (which should give early warning of impending problem) are poorly recognised even by senior doctors but particularly by nurses at the cot side<sup>5</sup>.

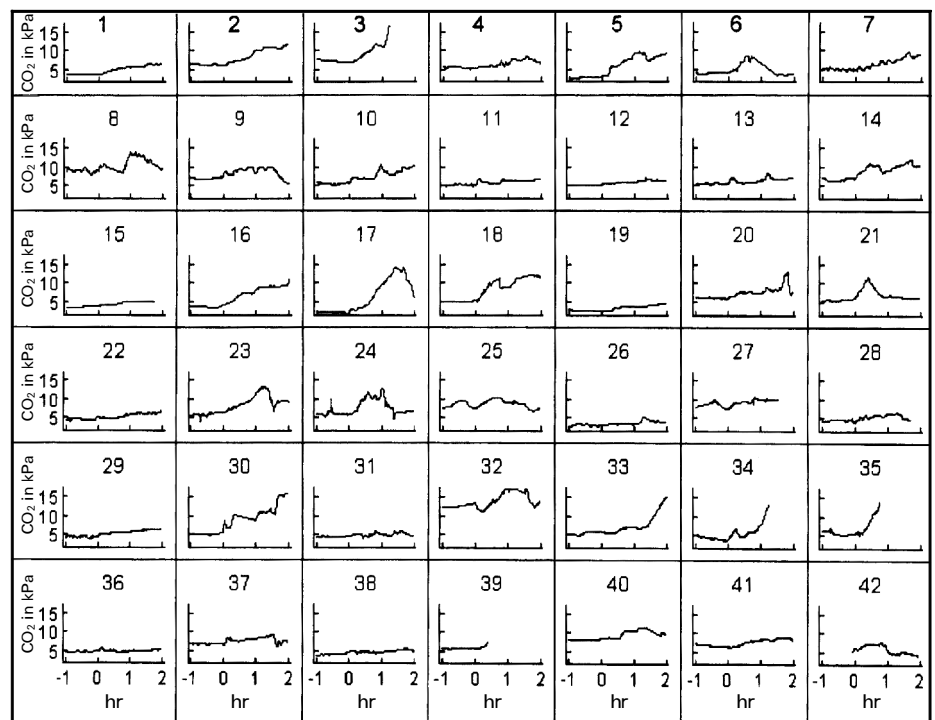
How can we then improve trend monitoring? Better visual display such as the portrayal of toe/core temperature difference instead of peripheral or central temperature singly is one way (Fig 2a). Others have described alternative data portrayals that alert to abnormality such as frequency polygons<sup>6</sup> and Metaphor Graphics<sup>7</sup>. The involvement of temporal considerations in such portrayals is more difficult<sup>8,9</sup>.

There is undoubtedly data overload in an intensive care unit and it is difficult to discriminate the data on the monitors because different features require different time compressions for their display. Thus, depending on the time axis, events may be more or less easy to identify. If an infant is having a number of apnoeic attacks or bradycardias, these will be revealed if the time axis has a period of 10 to 20 minutes, whereas if a dopamine infusion is started the blood pressure may rise slowly over a period of an hour and a time axis of 10 minutes will reveal little. Often it will be important to be aware of short-term and long-term trends at the same time, but the greater the quantity

**Table 1. Recognition of recorded events from physiological monitors.**

	Senior doctors	Junior doctors	Nurses
Number	7	8	10
Statements	11,921	6,610	9,465
Live diagnoses	68%	58%	25%
Eventual diagnosis	80%	78%	32%

Statements = total number of comments made about events by staff group. Live diagnosis = percentage of events diagnosed correctly by that staff group as event unfolded in simulated real time. Eventual diagnosis = percentage of events recognised by staff group when all data was displayed.



**Fig 3. Transcutaneous carbon dioxide levels in 42 cases of pneumothorax – each panel displays one case with 1 hour of data prior to onset and 2 hours following the occurrence.**

of data displayed, the more likely it is that the important data will remain unseen. We believe that, if trended data are going to be of any value, intelligence must be developed within the monitoring systems themselves. Computer algorithms will be required to produce intelligent warnings of impending problems utilising different time bases. These may also prompt certain actions if there is high sensitivity and specificity for the problem (decision support).

An early example that we attempted was to define pneumothorax by a small computer program (macro) of three lines:

```
<start pneumothorax programme
<15 min slope CO2 in range 0.04–0.4 kPa/min
<end pneumothorax programme
```

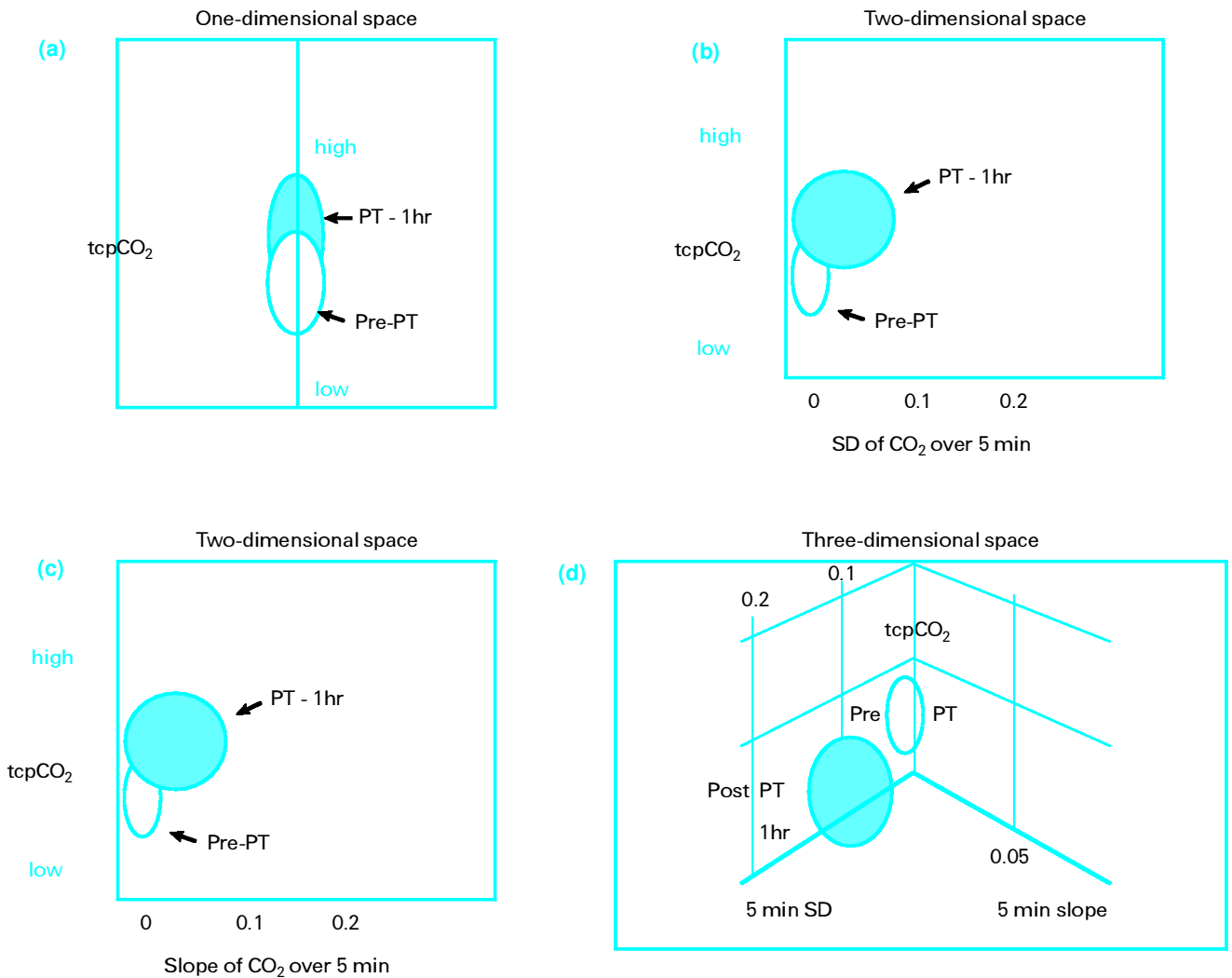
This simply looked for a particular slope in the trended carbon dioxide value. The sensitivity was excellent but the specificity was poor and related to the time spent monitoring<sup>10</sup>.

We have used the pneumothorax as our paradigm. This is a significant problem in the ventilated newborn infant. The incidence is variable and is decreasing with better ventilators and ventilation techniques. The mortality is historically as high as 40% and the associated morbidity related to pulmonary haemorrhage, chronic lung disease and periventricular haemorrhage is also significant.

In our first pneumothorax study we investigated whether, as in older patients, pneumothorax might be asymptomatic or whether rapid clinical deterioration immediately followed the onset<sup>3</sup>. If there was a significant preclinical period, diagnosis before clinical decompensation might reduce the mortality and morbidity. We examined 909 level 1 intensive care admissions representing over 90,000 intensive care days and found

42 infants who developed a pneumothorax on our neonatal care unit so that we had monitoring data before, during and after the occurrence. We needed gold standards for the time of onset and the time of diagnosis. We took as the time of diagnosis the time of the first x-ray confirming a pneumothorax. The gold standard for the time of onset was determined by the author who reviewed in detail the monitoring charts. The occurrence of pneumothorax is accompanied by a sharp rise in the transcutaneous carbon dioxide trace accompanied frequently by a rather shorter but equally rapid fall in the transcutaneous oxygen as illustrated in Fig 3b. In the 42 cases of pneumothorax, we found the median time between onset and diagnosis was 127 minutes with a range of 45 to 660 minutes. Ten percent died within 2 hours of the onset and 45% died before discharge. We concluded that the clinical diagnosis of pneumothorax was late and mortality is still very high<sup>3</sup>.

Our second study asked the question, is early warning of pneumothorax possible using expert knowledge of the carbon dioxide slopes? Our aims were to establish reference centiles for both the transcutaneous carbon dioxide levels but also for the slopes of carbon dioxide. To develop the normal centiles for both the levels and the slopes we used very premature infants who were stable and had low risk of subsequent problems at discharge<sup>11</sup>. The 97th centiles for the transcutaneous carbon dioxide slopes vary over the period when the trace is being examined. Over 5 minutes, the 97th centile is 0.06 kilopascals per minute; over 10 minutes 0.045; and over 15 minutes 0.037. Using these slopes, the pneumothorax was diagnosed in 24 cases within 5 minutes, 23 within 10 minutes and 18 within 15 minutes. In most cases, the inability to diagnose the pneumothorax was a lack of slope when the nurse in charge of the infant



**Fig 4. Multidimensional space.** (a) One dimension: the transcutaneous carbon dioxide level in the hour before and after onset of a pneumothorax. (b) Two dimensions: y axis the level, and x axis the slope, of the carbon dioxide. (c) Two dimensions: y axis the level, and x axis the variability, of the carbon dioxide. (d) Three dimensions: y axis the level, x axis the slope, and z axis the variability, of the carbon dioxide.

removed the probe to recalibrate it, not believing that the rising slope was a problem of the baby, but rather an equipment failure. Overall, 29 cases out of the total of 42 would have been diagnosed within 15 minutes. This indicated that the slope of the trended transcutaneous carbon dioxide can identify the onset of pneumothorax in a significant proportion of cases and thus might give extra warning for orderly intervention.

In a third pneumothorax study, we matched babies with pneumothorax with controls for the gestation of birth and the day of life. We used receiver-operator characteristic curves and found that the area under the curves for carbon dioxide trend slopes greater than the 90th centile was 82%. If we stipulated that the slope should be greater than the 90th centile for 5 consecutive minutes, the area under the curves was 89% with good specificity and sensitivity<sup>3</sup>.

The broad conclusions of our work up to this point indicated that medical and nursing staff were unused to the patterns on

the monitors, but that decision support using knowledge of both centiles of reference physiological data and derivatives of the data could provide automatic early warning with good positive predictive and negative predictive capacity.

### Multidimensional space

When considering a number of different variables together, each can be viewed as one dimension of a multidimensional space<sup>12</sup>. I would like to consider multidimensional space by comparing our normal data on carbon dioxide in the hour prior to a pneumothorax with the data in the one hour following a pneumothorax (Fig 3). One-dimensional space simply considers the values of carbon dioxide in the hour before and the hour after the occurrence of the pneumothorax (Fig 4a). Secondly, two-dimensional space considers either the carbon dioxide level together with the carbon dioxide variability (measured by stan-

## Key Points

### In intensive care units:

The most junior nursing and medical staff manage the patients at the bedside where detailed physiological data is displayed on multichannel monitors

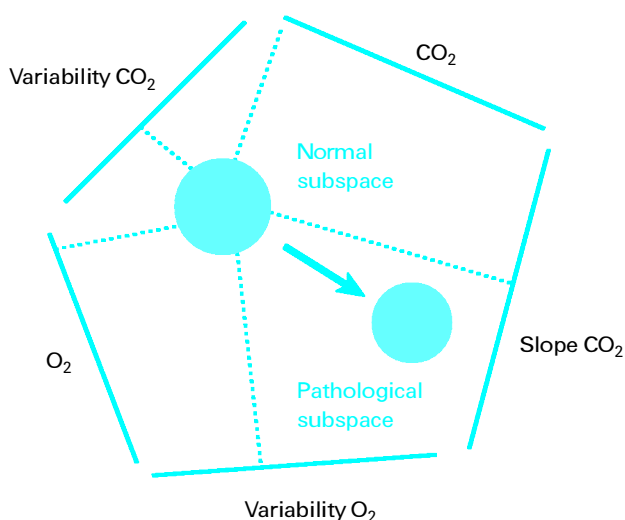
Lack of experience in pattern recognition and data overload lead to poor data discrimination and late diagnosis

Artificial intelligence techniques can identify important data patterns

Automatic decision aids utilising routinely collected data can support staff by giving early warning of key conditions

Building decision aids needs knowledge of where support is required and the consistency of data patterns associated with pathophysiological conditions and events

standard deviation) (Fig 4b), or alternatively the carbon dioxide level and the trend slope of the carbon dioxide level (Fig 4c), again comparing the hour before with the hour after the onset of the pneumothorax. Three-dimensional space portrays the carbon dioxide level, the slope and the standard deviation (Fig 4d). In the figures you can see how the overlap of carbon dioxide levels in one-dimensional space can be spread by viewing in two-dimensional space and completely separated by viewing in three-dimensional space. As the dimensions increase, the data from before and after the pneumothorax separate more clearly. Multidimensional space is a mathematical concept that can be applied when considering a large number of variables. In



**Fig 5.** 'Five-dimensional space' in two dimensions, with a normal subspace and a subspace referable to a particular pathology. Displayed in real time, the movement of data towards the pathological subspace indicates development of that particular problem.

five-dimensional space, one might consider the carbon dioxide level, the carbon dioxide slope, the carbon dioxide variability, the oxygen level and the oxygen slope.

Five-dimensional space can only be viewed in two dimensions. One can envisage that the coordinates from five separate variables will meet in a subspace of this five-dimensional space. When the individual is unwell this subspace will differ in a particular way depending on the underlying problem (Fig 5). By viewing this dynamically, it is possible to identify the development of the clinical problem at a point when the data from the well patient moves away from the 'well subspace' as he becomes unwell, to a subspace indicating a particular illness. In this way, we are able to see the development of pneumothorax very clearly within 5 minutes of its occurrence.

## Conclusions

The practice of medicine is largely a matter of pattern recognition. The signs and symptoms of a disease are presented in our classical teaching as patterns, eg the occurrence of a coryzal illness with white spots in the mouth (Koplik spots) indicates clearly that measles is the diagnosis. The recognition of symptom complexes is pattern recognition and similarly the responses to therapy should take a particular pattern. These patterns are usually not visual data patterns – we have had no training in these and are uncertain even when we see them as to how specific they are. However, pattern recognition in time series is what monitoring is all about. Can it be left to clinical staff to recognise these patterns? Clinicians have difficulty interpreting data for many reasons. For example, the vast amounts of data to be assimilated in an intensive care unit lead to data overload. Also, the distractions from the clinical environment lead to clinical overload. A multitude of alarms occur in intensive care unit usually indicating minor technical problems: the 'crying wolf' situation<sup>13</sup>. There are different levels of both nursing and medical expertise in the unit and overall it is the most junior staff who are at the bedside with the greatest diagnostic responsibility. There is also tiredness and human error in this highly fraught environment.

So pattern recognition in time series is what monitoring is all about, but since we cannot leave it to the clinical staff we can use pattern recognition techniques from the fields of artificial intelligence and computer science to build intelligent decision support into future monitors. A plan of action might be:

- 1 Patterns need to be recognised – this can be done with machine learning techniques<sup>14,15</sup> or by dedicated clinicians in the field going over many traces by eye<sup>3</sup>.
- 2 Patterns need to be described with artificial intelligence techniques<sup>10</sup>.
- 3 These descriptions need to be tested on large databases of such time series data<sup>3</sup>.
- 4 The monitors need to be programmed to recognise these patterns by artificial intelligence techniques.
- 5 Real-time testing then needs to be used to ensure that specificity and sensitivity is adequate for clinical usage.

## Acknowledgements

The work described above has been the result of collaboration with many people. Notably these have been Chris Bass and Peter Badger for their computer system design excellence; clinical colleagues in neonatal medicine (Drs Laing, Lyon, Midgeley, Menon, Stenson, Cunningham, and Becher) and colleagues in artificial intelligence and electrical engineering (Drs Murray, Dripps, Williams and Wilshaw) at the University of Edinburgh; psychologists (Professors Logie, Gilhooly and Alberdi) and computer scientists (Professor Hunter) at the University of Aberdeen; Dr I Kohane, Dr C Tsien and K Bradshaw at Boston Children's Hospital and Harvard University; Drs Miksch and Popov at the University of Vienna and Dr Cao in China.

## References

- 1 Cushing HW. The blood pressure reaction of acute cerebral compression illustrated by cases of intracranial haemorrhage. *Am J Med Sci* 1903;125:1017–44.
- 2 Mok Q, Bass CA, Ducker DA, McIntosh N. Temperature instability during nursing procedures in preterm neonates. *Arch Dis Child* 1991; 66:783–6.
- 3 McIntosh N, Becher J-C, Stenson BJ, Cunningham S *et al*. The clinical diagnosis of pneumothorax is late: use of trend data and decision support might allow preclinical detection. *Pediatr Res* 2000;48:408–15.
- 4 Cunningham S, Deere S, Symon A, Elton RA, McIntosh N. A randomized, controlled trial of computerized physiologic trend monitoring in an intensive care unit {see comments}. *Crit Care Med* 1998;26:2053–60.
- 5 Alberdi E, Becher J-C, Gilhooly K, Hunter J *et al*. Expertise and the interpretation of computerised physiological data: Implications for the design of computerised physiological monitoring in neonatal intensive care. *Int J Human Comput Studies* 2002;in press.
- 6 Green CA, Logie R, Gilhooly KJ, Ross DG, Ronald A. Aberdeen polygons: computer displays of physiological profiles for intensive care. *Ergonomics* 1996;39:412–28.
- 7 Horn W, Popow C, Uterasinger L. Support for fast comprehension of ICU data: visualisation using metaphor graphics. *Methods Inf Med* 2001;40:421–4.
- 8 Kosara R, Miksch S. Metaphors of movement: a visualisation and user interface for time orientated, skeletal plans. *Artif Intell Med* 2001;22: 111–31.
- 9 Kosara R, Miksch S. Visualizing complex notions of time. *Med Inf* 2001;10:211–5.
- 10 Giuffre M, Lyon AJ, Badger P, McIntosh N. The use of slopes in trend data for preclinical detection of pneumothorax (PT). *Pediatr Res* 1998; 43:174A.
- 11 McIntosh N, Becher J-C, Stenson BJ, Lyon AJ, Laing IA. The 5-minute slope characteristics of transcutaneous carbon dioxide tension can be useful for the early detection of neonatal pneumothorax. *Pediatr Res* 2000;47:368A.
- 12 Bellamy, J EC. Medical diagnosis, diagnostic spaces and fuzzy systems. *J Am Vet Ass* 1997;210:390–96.
- 13 Tsien, CL and Fackler, J. Poor prognosis for existing monitors in the intensive care unit. *Crit Care Med* 1997;25:614–19.
- 14 Tsien CL, Kohane IS, McIntosh N. Effect of data granularity on building models for multi-signal artifact detection. *Methods Inf Med* 2002; **submitted**.
- 15 Tsien CL, Kohane IS, McIntosh N. Multiple signal integration by decision tree induction to detect false alarms in the intensive care unit. *Artif Intell Med* 2000;19:189–202.