# AUSTRALASIAN ANAESTHESIA 2003

*Invited papers and selected continuing education lectures*


**Editors:**

John Keneally, Department of Anaesthesia
The Children's Hospital at Westmead, Sydney and University of Sydney

Michael Jones, Department of Anaesthetics
Royal Prince Alfred Hospital, Sydney

Published in 2003 by
Australian and New Zealand College of Anaesthetists
630 St Kilda Rd
Melbourne, Vic 3004

ISSN 1032-2515



Printed by Bridge Printery Pty Ltd
63-65 Burrows Road, Alexandria, NSW 2015

# Contents

# *Regional Sub-editors:*

| | |
|---|---|
| **Faculty of Intensive Care:** | Bala Venkatesh, Royal Brisbane Hospital, Brisbane |
| **Faculty of Pain Medicine:** | Paul Graziotti, Cambridge Pain Management, Perth |
| **Australian Capital Territory:** | Gerald Flynn, Canberra Hospital, Canberra |
| **New South Wales:** | Ed Loughman, Prince of Wales Hospital, Sydney |
| **New Zealand:** | Trevor Dobbinson, Dunedin |
| **Queensland:** | Peter Moran, Princess Alexandra Hospital, Brisbane |
| **Tasmania:** | Colin Chilvers, Launceston General Hospital, Launceston |
| **South Australia:** | Robert Singleton, Royal Adelaide Hospital, Adelaide |
| **Victoria:** | Rowan Thomas, St Vincent's Hospital, Melbourne |
| **Western Australia:** | Neville Gibbs, Sir Charles Gairdner Hospital, Perth |

# AUSTRALASIAN ANAESTHESIA 2003

# *Preface*

In the Preface of *Australasian Anaesthesia 2001*, we reported that it was hoped that the "Blue Book" would become an annual production. For a variety of reasons, that was not possible.

Despite this lapse, the popularity of this continuing education project has not waned. College Fellows often ask us when the next edition will appear. Trainees are even more demanding. They are convinced that a large proportion of examination questions derive from its contents.

We hope your expectations for this new edition will be met by the mix of papers you will find here. Many have had their origins in presentations at Faculty, Regional or Special Interest Group meetings. Some revisit old problems. Many reflect our increasing concern regarding medico-legal and ethical issues which trouble society as a whole, not just the members of our professions. All have been chosen by members of the Editorial group, whose interest was provoked when they first heard the authors talk about their topics. Without the support and enthusiasm of the sub-editors, the production of this book would be impossible.

We wish to thank all the authors, a number of whom are not members of our College or its Faculties, for the time and effort they have given to *Australasian Anaesthesia 2003* It is no easy matter to sit down and organise an essay of several thousand words. As usual, the papers have been subject to minimal peer review and their content remains the responsibility of the individual authors.

Abbott Australasia has again generously supported the production of Australasian Anaesthesia and we wish to acknowledge their readiness to assist in this endeavour.

John Keneally
Michael Jones

# Monitoring for Unconsciousness During General Anaesthesia

TIM McCULLOCH, MB, BS, BSc(Med), FANZCA
Staff Specialist, Dept of Anaesthetics, Royal Prince Alfred Hospital, Camperdown, NSW

*Tim McCulloch is a Staff Specialist Anaesthetist at the Royal Prince Alfred Hospital in Sydney, NSW. This paper was presented in part at a NSW Anaesthetic CME meeting.*

Reliable monitoring of unconsciousness during general anaesthesia has been termed the "Holy Grail" of anaesthetic monitoring.[1] The aim of this article is to briefly describe recent developments in this area and to critically review the current knowledge of available monitors with respect to prevention of awareness during general anaesthesia.

**Awareness During General Anaesthesia**

There is no generally accepted definition of awareness under anaesthesia. The term could be considered an oxymoron, given that unconsciousness is an essential feature of the state of general anaesthesia. Awareness, therefore, represents a failure to ensure absence of sensibility during general anaesthesia. An episode of awareness may or may not be associated with recall of the intraoperative event.

The consequences of awareness are variable. In the worst cases, patients suffer extreme fear, excruciating pain and overhear offensive or frightening remarks. Continuing with the worst-case scenario, they receive unsympathetic reactions from hospital staff, are not believed or have the experience trivialised, and may come to doubt their own sanity. They go on to suffer post-traumatic stress disorder with anxiety, depression and flashbacks. Not surprisingly, they may be extremely apprehensive about any future surgery. A disturbing series of case reports by Cobcroft and Forsdick make useful reading as an illustration of the devastating impacts of awareness.[2] At the opposite extreme, patients who do not experience pain may be euphoric from the anaesthetic agents despite their inadequate doses and, although experiencing awareness, may rate their anaesthetic care as entirely satisfactory.[3]

Recent estimates of the prevalence of awareness with recall vary between 0.1% and 1% of general anaesthetics, depending on the patient group. Practising anaesthetists tend to find these figures surprising; probably because most cases do not come to attention unless patients are specifically questioned postoperatively. Myles et al reported a large study in an Australian hospital in the 1990s that confirmed an incidence of over 0.1%.[4] In a prospective study designed to maximise the chance of detecting awareness with recall, Sandin et al reported an incidence of 0.18% in patients given neuromuscular blocking drugs and 0.1% in non-paralysed patients. The episodes of awareness in non-paralysed patients were reported as being brief and without sequelae.[5] A follow-up study, two years after the episodes of awareness, confirmed that

post-traumatic stress disorder is common after intraoperative awareness and that severe symptoms are confined to patients who received neuromuscular blocking drugs.[6]

Current trends in anaesthetic practice aimed at reducing times to emergence and, in day-surgery patients, hastening discharge could potentially increase the risk of awareness. For example, use of newer drugs with a more rapid offset of action (e.g. sevoflurane, desflurane and remifentanil) would be expected to speed the return of consciousness if there is an inadvertent interruption to drug-delivery. Hence, there may be a reduced window of opportunity for detecting a technical problem before awareness occurs. Also, the aim of speeding emergence creates a pressure to titrate anaesthetic drug doses to the minimum required for each patient. Without a method to determine with certainty that a paralysed patient is unconscious, reducing doses of anaesthetic agents could increase the risk of awareness.

**Monitoring General Anaesthesia**

*The Ideal Monitor*

There are fundamental problems in attempting to monitor for unconsciousness during anaesthesia. In contrast to the physiological variables routinely monitored during anaesthesia, there is not even a generally accepted definition of consciousness. Furthermore, we have virtually no concept of how conscious experience arises. An ideal monitor of general anaesthesia would directly detect the presence or absence of this thing called "consciousness" but, given the lack of knowledge of the physiological basis of consciousness, it seems improbable that such a monitor will be developed soon.

Recent developments in monitoring the central nervous system effects of general anaesthetic agents have not been derived from any particular hypothesis regarding the origin of consciousness. Nor are they based on any particular theory of the mechanism of action of general anaesthetic agents. Without an understanding of the mechanism of consciousness, the best we can hope for at this time is a measurement that *correlates* with consciousness. Research has focused on attempts to correlate the state of consciousness with:

1. various features of the surface electroencephalogram (EEG); or,
2. changes in cortical auditory evoked potentials (AEPs).

An ideal monitor would detect the presence or absence of consciousness in all circumstances. It would be independent of factors that could alter the EEG such as the choice of anaesthetic agent, age of the patient, body temperature, neurological disease and psychotropic medications. If a monitor is to be used to reduce the risk of a rare event such as awareness, it must have a very high sensitivity for consciousness. Ideally, there should be no overlap in the output of the monitor in conscious individuals with the output in unconscious individuals; i.e. there should be 100% sensitivity and specificity.

*Bispectral Index*

The bispectral index, BIS® (Aspect Medical Systems, Natick, USA), was the first of the new generation monitors of anaesthetic effect to be commercially released and is the most extensively studied and will therefore be the main focus of this review. Rampil has given a description of the development and derivation of the BIS.[7] The BIS is a score derived from the frontal EEG. The proprietary algorithm to determine this

score was developed empirically by a computer-aided search for statistical correlations between EEG characteristics and the state of consciousness. Hundreds of EEG recordings were collected from healthy volunteers during graded sedation/anaesthesia and various features of the EEG were examined for correlation with the clinical assessment of sedation. Several features of the EEG were identified and these features are fed into the BIS algorithm to derive the index. The unique feature of the BIS is that it includes analysis of the bicoherence of the EEG, which is a measure of the phase relationships between the component waves.

The output of the BIS algorithm is a score between 0 and 100. The BIS score is above 90 in awake subjects and falls with increasing sedation until the patient becomes unresponsive — usually around an index of 70. Once consciousness is lost, the BIS continues to fall with increasing doses of anaesthetic agent. The manufacturers of the BIS recommend that for surgical anaesthesia, doses of anaesthetic drugs be adjusted to maintain a BIS between 40 and 60. The BIS monitor responds appropriately to most anaesthetic drugs including volatile agents, propofol, thiopentone and benzodiazepines.[8,9] There are two notable exceptions: the BIS does not appear to track the anaesthetic effects of nitrous oxide[10,11] or ketamine.[12] Also, the BIS has been reported to correlate poorly with loss of responsiveness during induction of anaesthesia with fentanyl.[13]

To smooth out the number, the BIS is calculated as a rolling average of the previous 15 or 30 seconds of EEG recording. The index reflects the state of the EEG over the immediately preceding time period so, when the state of arousal changes suddenly, the index necessarily lags behind. When a patient responds to a stimulus, for example by movement or eye opening, it is common for the BIS to rise rapidly a few seconds *after* the clinical response.

The BIS monitor incorporates sophisticated and quite effective algorithms for detecting artefacts. Some types of artefact (electrocardiograph and blink artefacts) can be "repaired" but high frequency artefact from facial muscle activity or from diathermy causes rejection of the EEG epoch. During prolonged periods of diathermy, the monitor cannot calculate a BIS score. Occasionally, the algorithm fails to reject high frequency artefact, resulting in a falsely high BIS. The latest revision of the BIS monitor (BIS-XP) is much more resistant to high frequency artefact. This should reduce the problems of falsely high readings and loss of readings during diathermy.

Drummond has noted that the published information regarding the reliability of the BIS is limited as most studies have been on healthy young adults receiving only one or two drugs.[14] In contrast to highly controlled studies, real-world patients receive a multitude of different combinations of anaesthetic drugs. The reliability of the BIS in patients on psychotropic medications and patients with central nervous system diseases has not been established. With regards to elderly patients, Katoh et al studied the BIS scores at low doses of sevoflurane and compared patients over 65 years with younger adults.[15] They found that the BIS performed similarly well in predicting loss of consciousness in each of the age groups studied. However, there is a phenomenon in elderly patients whereby the EEG displays increased high frequency activity just prior to the onset of burst-suppression and this can cause a paradoxical increase in the BIS with increasing doses of anaesthetic agent.[16]

*Memory, Recall and the BIS*

If a subject is presented with information while being sedated with anaesthetic

agents, the chance of them recalling the information has been shown to correlate well with the BIS.[8,17] Glass et al investigated a variety of anaesthetic agents and calculated that the chance of remembering a picture or a word was less than 5% if the BIS was below 64 at the time of the stimulus.[8] In the year 2001, the manufacturer of the BIS claimed that the monitor had been used on over 2 million patients and that they had received 54 reports of possible awareness occurring while the patients were monitored with the BIS. In 28 of these cases, the trend recording of the BIS could be retrieved and in each of those cases the index was over 65 at the time of awareness. The latest information from the manufacturer is that BIS monitoring has been used in over 5.5 million cases and, at the time of writing, there has yet to be reported a definite case of intraoperative awareness with a trend record confirming a BIS below 60 at the time of the event.

It is now well known that doses of anaesthetic agents adequate to prevent explicit recall of events do not always suppress all evidence of learning. Implicit memory is said to occur if a stimulus presented during an anaesthetic, whilst not consciously recalled, is found to alter subsequent behaviour. In a study of implicit memory formation during general anaesthesia for trauma surgery, the likelihood of implicit memory correlated well with the BIS, but not with other indicators of anaesthetic effect such as spectral edge frequency, end-tidal volatile concentration, or haemodynamic variables.[18] In that study, there was still some evidence of implicit memory formation at BIS scores below 60 (but no incidents of explicit recall).

*Consciousness, Responsiveness and the BIS*

Sedation can be graded according to the strength of the stimulus required to elicit a response and the BIS has repeatedly been shown to correlate well with such a grading. However, the correlation is not absolute as there is always some overlap between the BIS scores of responsive versus non-responsive individuals.[8, 10, 15, 17, 19, 20] That is, there is no cut-off BIS value that enables detection of consciousness with both 100% specificity and 100% sensitivity.

The BIS is influenced by both the dose of anaesthetic drugs and the intensity of surgical stimulation. This means that, prior to the onset of surgery, it is not possible to use the BIS to determine if the dose of anaesthetic agent is appropriate. Röpcke et al found the average desflurane concentration required to achieve a BIS of 50 in healthy adults prior to skin incision was 2.2%, compared to 6.8% during laparotomy.[20] It is common to see the BIS rise suddenly after a stimulus such as incision or intubation.[19] The change in BIS after stimulation is unpredictable and can be obtunded by prior administration of opioids. The BIS gives information about the current state of the EEG — or, more accurately, the average state of the EEG over the preceding 15 or 30 seconds — but does not reliably predict the future response of a patient to a new stimulus. If general anaesthesia is defined as a state of unrousable unconsciousness, it could be said that the BIS correlates well with the state of unconsciousness, but correlates less well with the unrousability.

*Use of BIS monitoring to individualise anaesthetic dose*

With both intravenous and inhalational anaesthetic agents, there is significant variation between individuals in the dose required to achieve unconsciousness. A common approach to this problem, particularly in paralysed patients, is to administer

a dose that the anaesthetist judges will be adequate to prevent awareness in the vast majority of patients. The problem with this approach is that most patients receive a dose in excess of their requirement, sometimes resulting in untoward side effects and delayed emergence. The potential advantages of a monitor that could facilitate individualised dosing are obvious. By titrating anaesthetic administration to the BIS, both the average doses of anaesthetic drugs and the time to emergence can be reduced.[21] The cost savings could be significant, although in one report the drug and time savings were unimpressive.[22]

Monitoring the cortical effect of general anaesthetic agents has the potential to enable more rational treatment of haemodynamic disturbances. For example, if the blood pressure rises above acceptable limits, BIS monitoring may help resolve the dilemma of whether or not to increase the dose of hypnotic agents. If the BIS is already below 50, inadequate anaesthesia is highly unlikely and it may be more appropriate to administer opioids or antihypertensive drugs. Conversely, with hypotension the BIS monitor may be used to confirm that the patient requires lower than average doses of anaesthetic agents. On the other hand, if the hypotensive patient has a BIS above 60 it would not be advisable to reduce anaesthetic doses and pharmacological support of the cardiovascular system may be more appropriate.

*Does the BIS reduce the risk of awareness?*

The low prevalence of awareness under general anaesthesia creates great difficulties in finding a definitive answer to this question.[23] A randomised trial would require a prohibitively large number of patients to detect a clinically significant reduction in the incidence of awareness in all patients having general anaesthesia. To deal with this problem, a multicentre randomised trial currently underway in Australasia (the "B-Aware Trial") is confined to patients thought to be at higher risk of awareness. There are also two larger non-randomised prospective studies in progress internationally (one cohort study and one historically controlled) which will attempt to document the incidence of awareness in the presence of BIS monitoring.

The BIS monitor appears to track the effect of anaesthetic agents reliably in the vast majority of cases so that a failure of drug delivery (e.g. faulty vapouriser or propofol pump) should lead to an unexpectedly high BIS. Hopefully, this would alert the clinician before awareness occurs.

*Could Using the BIS Monitor Cause Awareness?*

The theoretical concern that use of the BIS monitor could increase the risk of awareness must be considered. As noted above, potential advantages of using the BIS to titrate anaesthetic doses include more rapid and predictable emergence times, reduced side effects of anaesthetic drugs and reduced drug costs. However, with regard to the risk of awareness, any strategy aimed at reducing anaesthetic agent dosage must be examined very carefully, as there is clearly an inherent tension between the two goals of preventing awareness and reducing drug use. If a monitor is used to guide a reduction of anaesthetic dosage to the bare minimum in every patient, then the risk of awareness could be increased. If anaesthetic doses are adjusted to maintain patients' BIS scores close to a certain value, then it is important that the chosen BIS value represents an extremely low probability of consciousness.

The recommended target for the BIS during surgery is between 40 and 60, but scores approaching 60 may in fact represent a very narrow safety margin in some patients. In

the study by Flaishon et al, two out of 40 patients obeyed commands at BIS scores of 58 or 59.[9] Kearse et al[10] reported responses to voice with BIS scores down to 57 and inspection of the of data in other studies indicate that some subjects responded to voice at BIS scores less than 60[8] and as low 54.[24] It is important to note that the BIS scores in these studies were recorded prior to any stimulation. With arousal, the EEG changes and the BIS subsequently rises[24] (the BIS necessarily lags behind the cortical changes). Utilising the isolated forearm test, Schneider et al found that some patients responded to command after the stimulus of tracheal intubation, despite a BIS below 60 prior to the stimulus.[19] The BIS scores rose to between 65 and 80 in those patients who responded (none of whom had any recall). These data indicate that BIS scores approaching 60, particularly if there is no stimulation at the time, can be associated with inadequate anaesthesia and a new stimulus, such as intubation or surgical incision, may then cause return of consciousness. Therefore, the use of BIS monitoring does not replace the clinical skill of the anaesthetist in anticipating the likely effects of intense stimulation and administering adequate drug doses in advance to prevent awareness. Use of BIS monitoring without bearing in mind the above considerations could be falsely reassuring and could theoretically contribute to cases of awareness.

It is important to note that the above speculations about the possibility of BIS monitoring leading to awareness are entirely theoretical. Although it is known that cases of awareness have occurred in the presence of BIS monitoring, details of most of these cases have not been published so it is not possible to say whether BIS monitoring could have contributed to the problem. If use of BIS monitoring to guide anaesthetic doses can lead to awareness, it would only take one or two cases per thousand for this effect to override any potential benefit of the monitor in detecting awareness.

*Movement*

The issue of unwanted patient movement is interesting. When the BIS algorithm was developed, it was hoped that it would predict movement in response to surgical stimuli. This did not succeed and the focus switched to the BIS as an indicator of the hypnotic effect of general anaesthetics. Suppression of movement by anaesthetic agents probably occurs at the level of the spinal cord, so it is perhaps not surprising that a monitor of cortical activity fails to predict movement. Administration of opioids during a general anaesthetic has little effect on the EEG, and therefore the BIS, but significantly reduces the likelihood of movement on skin incision.[25] Guignard et al found that giving remifentanil prior to intubation in patients anaesthetised with propofol had no immediate effect on the BIS but there was a dose-dependent inhibition of both the rise in BIS after intubation and the likelihood of movement after intubation.[26]

If a monitor could be used to guarantee immobility without giving muscle relaxants, the number of patients being paralysed could be reduced and, presumably, the risk of awareness reduced in some patients. However, if the BIS is used to minimise the dose of anaesthetic agents, this can lead to an increased incidence of unwanted movement, particularly if little intraoperative opioid is given. The response of the anaesthetist could be to paralyse patients whom they would not have felt obliged to paralyse at a higher anaesthetic dose. This combination — reduced anaesthetic dose and increased use of muscle relaxants — could cause awareness if using the BIS monitor failed to ensure unconsciousness.

**Other monitors of awareness**

Besides the BIS, the most extensively described monitor for consciousness is middle-latency auditory evoked potentials (mlAEPs). Headphones placed on the patient produce an auditory stimulus and the cortical EEG response is recorded. The term "latency" refers to the time lag between the auditory stimulus and the subsequent EEG response. Loss of consciousness with anaesthesia is associated with both an increased latency and decreased amplitude of the mlAEPs (the earlier brainstem potentials are preserved). A single number, the auditory evoked potential index, has been developed that reflects both latency and amplitude changes.[27] A monitor (A-Line®, DanmeterA/S, Odense, Denmark) is now commercially available; this incorporates a new method of extracting the auditory evoked potential which reduces the number of stimuli required to update the index and hence improves the response time to changes in arousal.[24] Struys et al compared this monitor to the BIS during administration of propofol as a sole agent and found it performs similarly well in distinguishing between conscious and unconscious subjects.[24] Both monitors showed a similar overlap in their output between conscious and unconscious subjects and therefore had similar sensitivity and specificity. They also noted that both the A-Line and the BIS were similarly poor at predicting response to a noxious stimulus.

Other EEG-derived indexes are in the process of development but at this stage there is no peer-reviewed information available regarding their performance. These include monitors that calculate the entropy of the EEG signal: with the onset of anaesthesia the EEG becomes more "organised" and therefore has reduced entropy.

**Could a monitor ever eliminate awareness under anaesthesia?**

Some instances of awareness may not be preventable despite the aid of a monitor of consciousness. For example:

- Awareness is more common in severely hypovolaemic trauma victims and other cardiovascularly unstable patients due to a deliberate reduction in anaesthetic dose. Even with a monitor warning of awareness, it will not always be possible to give adequate anaesthesia.
- Awareness due to the accidental administration of muscle relaxant prior to induction of anaesthesia would not be prevented by monitoring.
- Intraoperative monitoring would not be expected to prevent spurious reports of awareness. Human memory is never entirely reliable and could be even less so with the emotional stress of surgery, along with the effects of illness and anaesthetic drugs. Awareness is unequivocal if the patient can relate explicit details of intra-operative events that they could not have known about without having been conscious at the time. However, the experience described by the patient is often less clear-cut, making it difficult to be certain whether true intraoperative awareness occurred. Experiences during the early recovery period, either real or dreamt, can be remembered by the patient as having occurred intraoperatively. Also, it is conceivable that the rare patient might fabricate a report of awareness.
- Even if a monitor indicates possible consciousness, the response of the anaesthetist may not always prevent recall. For example, two cases have been reported of awareness despite the BIS monitor warning of inadequate anaesthesia.[28] In the first case, the anaesthetist increased the propofol dose but it was only later discovered that the propofol infusion was not reaching the patient. In the second case, there was a technical failure in the delivery of volatile agent. The BIS was only moderately

raised (65-70) and the anaesthetists, believing this to be acceptable, did not discover the technical problem for some time.

**Conclusions**

The BIS and other more recently developed monitors offer the possibility of rational administration of agents aimed separately at suppressing consciousness and suppressing reflex responses to surgical stimuli. Use of the BIS has been found to reduce average anaesthetic drug use and reduce times to emergence. However, the BIS is not an ideal monitor, as it does not perform equally well for all anaesthetic agents and there is some overlap between the index in conscious and unconscious patients. A BIS within the recommended range is not an absolute guarantee of adequate anaesthesia as the index can increase suddenly and unpredictably with a new stimulus. It remains to be demonstrated whether use of the BIS, or any other monitor, reduces the risk of awareness under anaesthesia and there remains a theoretical possibility that it could have the opposite effect.

**References**


1. Todd MM. EEGs, EEG processing, and the bispectral index. *Anesthesiology* 1998; 89:815-817.
2. Cobcroft MD, Forsdick C. Awareness under anaesthesia: the patients' point of view. *Anaesth Intens Care* 1993; 21:837-843.
3. Miller DR, Blew PG, Martineau RJ, Hull KA. Midazolam and awareness with recall during total intravenous anaesthesia. *Can J Anaesth* 1996; 43:946-953.
4. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM. Patient satisfaction after anaesthesia and surgery: results of a prospective survey of 10,811 patients. *Brit J Anaesth* 2000; 84:6-10.
5. Sandin RH, Enlund G, Samuelsson P, Lennmarken C. Awareness during anaesthesia: a prospective case study. *Lancet* 2000; 355:707-711.
6. Lennmarken C, Bildfors K, Enlund G, Samuelsson P, Sandin RH. Victims of awareness. *Acta Anaesthesiol Scand* 2002; 46:229-231.
7. Rampil IJ. A primer for EEG signal processing in anesthesia. *Anesthesiology* 1998; 89:980-1002.
8. Glass PS, Bloom M, Kearse L, Rosow C, Sebel P, Manberg P. Bispectral analysis measures sedation and memory effects of propofol, midazolam, isoflurane, and alfentanil in healthy volunteers. *Anesthesiology* 1997; 86:836-847.
9. Flaishon R, Windsor A, Sigl J, Sebel PS. Recovery of consciousness after thiopental or propofol. Bispectral index and isolated forearm technique. *Anesthesiology* 1997; 86:613-619.
10. Kearse LA, Jr., Rosow C, Zaslavsky A, Connors P, Dershwitz M, Denman W. Bispectral analysis of the electroencephalogram predicts conscious processing of information during propofol sedation and hypnosis. *Anesthesiology* 1998; 88:25-34.
11. Barr G, Jakobsson JG, Owall A, Anderson RE. Nitrous oxide does not alter bispectral index: study with nitrous oxide as sole agent and as an adjunct to i.v. anaesthesia. *Brit J Anaesth* 1999; 82:827-830.
12. Sakai T, Singh H, Mi WD, Kudo T, Matsuki A. The effect of ketamine on clinical endpoints of hypnosis and EEG variables during propofol infusion. *Acta Anesthesiol Scand* 1999; 43:212-216.
13. Barr G, Anderson RE, Owall A, Jakobsson JG. Effects on the bispectral index during medium-high dose fentanyl induction with or without propofol supplement. *Acta Anesthesiol Scand* 2000; 44:807-811.
14. Drummond JC. Monitoring depth of anesthesia: with emphasis on the application of the bispectral index and the middle latency auditory evoked response to the prevention of recall. *Anesthesiology* 2000; 93:876-882.
15. Katoh T, Bito H, Sato S. Influence of age on hypnotic requirement, bispectral index, and 95% spectral edge frequency associated with sedation induced by sevoflurane. *Anesthesiology* 2000; 92:55-61.
16. Detsch O, Schneider G, Kochs E, Hapfelmeier G, Werner C. Increasing isoflurane concentration may cause paradoxical increases in the EEG bispectral index in surgical patients. [see comments.]. *Brit J Anaesth* 2000; 84:33-37.
17. Liu J, Singh H, White PF. Electroencephalographic bispectral index correlates with intraoperative recall and depth of propofol-induced sedation. *Anesth Analg* 1997; 84:185-189.

18. Lubke GH, Kerssens C, Phaf H, Sebel PS. Dependence of explicit and implicit memory on hypnotic state in trauma patients. *Anesthesiology* 1999; 90:670-680.
19. Schneider G, Wagner K, Reeker W, Hanel F, Werner C, Kochs E. Bispectral Index (BIS) may not predict awareness reaction to intubation in surgical patients. *J Neurosurg Anesthesiol* 2002; 14:7-11.
20. Ropcke H, Rehberg B, Koenen-Bergmann M, Bouillon T, Bruhn J, Hoeft A. Surgical stimulation shifts EEG concentration-response relationship of desflurane. *Anesthesiology* 2001; 94:390-399.
21. Gan TJ, Glass PS, Windsor A, et al. Bispectral index monitoring allows faster emergence and improved recovery from propofol, alfentanil, and nitrous oxide anesthesia. BIS Utility Study Group. *Anesthesiology* 1997; 87:808-815.
22. Wong J, Song D, Blanshard H, Grady D, Chung F. Titration of isoflurane using BIS index improves early recovery of elderly patients undergoing orthopedic surgeries. *Can J Anaesth* 2002; 49:13-18.
23. O'Connor MF, Daves SM, Tung A, Cook RI, Thisted R, Apfelbaum J. BIS monitoring to prevent awareness during general anesthesia. *Anesthesiology* 2001; 94:520-522.
24. Struys MM, Jensen EW, Smith W et al. Performance of the ARX-derived auditory evoked potential index as an indicator of anesthetic depth: a comparison with bispectral index and hemodynamic measures during propofol administration. *Anesthesiology* 2002; 96:803-816.
25. Sebel PS, Lang E, Rampil IJ et al. A multicenter study of bispectral electroencephalogram analysis for monitoring anesthetic effect. *Anesth Analg* 1997; 84:891-899.
26. Guignard B, Menigaux C, Dupont X, Fletcher D, Chauvin M. The effect of remifentanil on the bispectral index change and hemodynamic responses after orotracheal intubation. *Anesth Analg* 2000; 90:161-167.
27. Mantzaridis H, Kenny GN. Auditory evoked potential index: a quantitative measure of changes in auditory evoked potentials during general anaesthesia. *Anaesthesia* 1997; 52:1030-1036.
28. Luginbuhl M, Schnider TW. Detection of awareness with the bispectral index: two case reports. *Anesthesiology* 2002; 96:241-243.

# Depth of Anaesthesia Monitors in Paediatric Anaesthesia

ANDREW DAVIDSON, MB, BS, FANZCA
Royal Children's Hospital, Melbourne

*Andrew Davidson is a staff anaesthetist at the Royal Children's Hospital, Melbourne. His research interests include assessment of anaesthesia depth, brain protection and ethics in research.*

There has been a resurgence of interest in depth of anaesthesia assessment. This has arisen due to a greater fear of awareness, the increased use of TIVA and the advent of new devices and technologies. What role do the new depth monitors have in paediatric anaesthesia? I will first outline the inherent limitations of the current technology, whether current monitors are in theory suitable for children, how much they have been validated in children, and finally how paediatric anaesthesia differs in its suitability for the use of depth of anaesthesia monitors.

**Limitations of current monitors**

Several depth of anaesthesia monitors are now available for clinical use and more are being trialled. These include the Bispectral Index (BIS), the Patient State Index, the entropy monitor, processed auditory evoked responses and ocular micro-tremor. The BIS has spawned hundreds of academic publications and has been used in millions of anaesthetics. In contrast, the entropy monitor has only resulted in a handful of publications.[1] Which monitor is better? It is unlikely one will be significantly better than any other, due to the inherent limitations in what they are trying to achieve and how.

All current monitors measure and analyse aspects of brain electrical activity. Various mathematical derivatives are then generated. These mathematical derivatives are compared to those of previously analysed records and a number allocated to each set of derivatives. The number thus generated is on a scale that equates lower numbers with periods of adequate anaesthesia and higher ones with measures of wakefulness. Often, the monitors go further to equate very low numbers with higher concentrations of anaesthetic.

It is a very important concept that these monitors do not specifically measure anything that represents consciousness or memory processing. They measure electrical activity and this is associated with some crude assessments of consciousness. Unfortunately, consciousness is a slippery thing that we have difficulty defining and even greater difficulty measuring. Thus, the monitors have never had a "gold standard" to be calibrated against. Alas, we have even greater trouble defining consciousness in infants and neonates. Some theories of consciousness would even argue that children less than 18 months of age have no state of consciousness at all![2]

A further problem is that the EEG can be influenced both by degree of arousal and directly by the anaesthetic agent itself. At high concentrations of an anaesthetic, the

derived number will still go up and down with the dose of the anaesthetic, but may not vary significantly with stimulation. The patient is unconscious. Thus, the idea that consciousness is being measured is nonsense. The monitor is mainly responding to the direct effect of the anaesthetic agent on the EEG. The numbers generated have two determinants; the concentration of a particular drug (particularly at higher concentrations) and the degree of arousal (particularly at lower concentrations).

It has also been known for many years that different anaesthetic drugs can have quite different effects on the EEG.[3] Ketamine, nitrous oxide, xenon, high doses of opioids and even halothane have direct EEG effects different to those of isoflurane, propofol and thiopentone. The direct EEG effects of the former anaesthetic agents make interpretation of the monitors problematic, as they have been calibrated using the latter agents. Unfortunately, ketamine and halothane are still frequently used in paediatric anaesthesia.

Even if the monitors do not measure any real entity, they are still potentially useful. The algorithms have been designed to reliably guide an anaesthetist through a typical anaesthetic. However, the algorithms are derived from adult cases and therefore are not intended to fit paediatric anaesthesia. Thus, any outcome benefit demonstrated in adults may not be applicable to paediatrics.

**Could they work in theory?**

The awake EEG changes with age.[4,5] The rate of change slows as age increases, with most changes occurring in the neonatal and infant groups. The most pronounced difference is in the dominant background frequency, which increases with age. At 6 months, the dominant frequency is 5 Hz, from 9-18 months it is 6-7 Hz, at two years it is 7-8 Hz, by 7 years it is 9 Hz and by 15 years of age reaches the adult levels of 10 Hz. If a depth monitor is partly determined by the power frequency relationship, then an age related bias could occur. Young children can also demonstrate peculiar and specific EEG changes when progressing from the asleep to the awake phase. Children aged 6 months to 4 years have short bursts of 4-8 Hz activity lasting 1-5 seconds. Longer periods of 1-3 Hz activity may be seen in 3 month to 5 year olds. This activity is seen maximally at 12 months.

There is very little systematic data about the EEG of children under anaesthesia. What is available suggests that the EEG under anaesthesia does behave differently in children compared to adults.[6-8] Even less is known of the auditory evoked response. One study concludes that the auditory evoked responses are unreliable as a depth of anaesthesia monitor in children younger than 2 years.[9]

**Validating the monitors in children**

Given the theoretical basis for a difference between adults and younger patients, it would be instructive to examine the efforts made to validate these monitors in children, particularly the lower age groups. These are difficult studies to perform well, as adult measures of consciousness may be impossible to use in children. As well as it being unethical to recruit child volunteers, very young children cannot be asked to perform tasks, or be quizzed to assess word retention. Also, small children tend to awaken rapidly making definition of subtle end-points difficult.

Nevertheless, several investigators have tried to validate the BIS in children. Denman et al showed a decrease in BIS with increasing sevoflurane concentration in infants (described as less than 2 years of age) and children during maintenance of

anaesthesia.[10] The BIS at various stages of anaesthesia was recorded and compared to historical adult controls. No difference was detected between age groups in BIS during emergence, induction or maintenance of anaesthesia. The study was limited, however, by the imprecise manner in which emergence was assessed and had insufficient power to compare infants with older children. Degoute et al assessed the correlation between BIS and the OAA/S scale (Observer's Assessment of Awareness/Sedation).[11] They compared children aged 3.5 to 13 years with adults and found no difference between BIS numbers at both loss of consciousness and return of consciousness.

In a more rigorous study, BIS was also assessed prior to awakening after a specific stimulus.[12] This compared children older than 1 year with infants aged 6 months to 1 year. In children, the number prior to awakening was similar to that described in adults. However, in infants the BIS number prior to awakening was lower. Infants also failed to show a rise in BIS as the anaesthetic concentration was lowered. These results could imply that the EEG is sufficiently different in infants to make interpretation of the BIS unreliable in this age group. The BIS should therefore be used with caution in infants and neonates.

Finally, a phenomenon known as paradoxical delta activity occurs rarely in some patients just prior to arousal and is reflected by a sudden and brief plunge in the BIS immediately prior to awakening.[13] Paradoxical delta activity is more common in children emerging from volatile anaesthesia. This phenomenon may be related to the slow frequencies mentioned previously, which are normally seen in small children during the transition from drowsy to the awake state.

It is tempting to follow the line of some authors and insist proper validation be performed in children before clinical use of this type of monitoring.[14] Another approach is to accept that good validation studies may never be performed. We do much in paediatrics using technology that has never been specifically validated in children. If well-designed studies in relevant age groups can demonstrate improvements in important clinical outcomes, then complex validation studies may be superfluous.

There is no information yet about validating the Patient State Index, entropy monitors or ocular micro-tremor in children.

**Would a depth monitor be useful in paediatric anaesthesia**

Having tentatively established that, in older children, at least one monitor (the BIS) corresponds with consciousness as well (or as badly) as it does in adults, what is its role in paediatric anaesthesia?

Firstly, it could be used as it is in adult anaesthesia — to guide anaesthesia delivery for optimising dose. The benefits from this are faster emergence, lower costs in drug usage and post-operative care and, possibly, reduced risks of awareness. In adults, faster emergence and drug cost savings have been demonstrated.[15] However, paediatric anaesthesia is different to adult anaesthesia in both the way it is conducted and in some of the clinical goals. A rapidly "street fit" child is not always our highest priority. Techniques that save recovery time in adults may not save time in children.

Bannister et al have evaluated the use of BIS guided anaesthesia in reducing times to emergence and discharge.[16] They found that, in children aged 3 to 18 years, less anaesthetic drug was consumed, and the children awoke and were ready for discharge sooner. The time to actual discharge was, however, unchanged. In children less than 3 years of age, there was no difference in any of the outcomes. In children less than

6 months they had difficulty titrating the anaesthetic to the desired BIS number. Similarly to Davidson et al,[12] they found a lack of correlation between BIS number and low anaesthetic concentrations in this age group.

Reduction of awareness using the BIS is being evaluated at present in several large trials in the adult patient population. There is no reason to think that using the BIS would not help detect inadequate anaesthesia delivery or equipment malfunction and thus result in a reduction in awareness. It may also help detect that group of patients who are aware due to a higher anaesthesia requirement. There is, however, less evidence that this forms a significant proportion of cases of awareness. Moerman's oft quoted paper suggests anaesthesia records fail to give an indication of awareness because of poor record keeping, not because patients are frequently aware when all seems well.[17]

Very little has been published about awareness in children.[18] In 1973, a study demonstrated an alarming incidence of 5% in 202 children aged 7 to 14.[19] In 1988, two smaller studies from Liverpool of 120 children aged 5-17 and 144 children aged 5-14 reported no cases of awareness.[20, 21] The later studies were limited by a relatively less comprehensive follow-up. Given the low incidence in contemporary adult studies, the small sample sizes of the Liverpool studies are also unlikely to give any meaningful result when comparing paediatric to adult anaesthesia.

The incidence of awareness may well be different in children, as the techniques of paediatric anaesthesia are different and there is less consistency in the pharmacokinetics and pharmacodynamics of anaesthetic agents. Small children also show greater fluctuation in haemodynamic variables, reducing their specificity in indicating inadequate anaesthesia. A child's response to the trauma of awareness may also be very different. A pain-free neonate may not be as upset by hearing the surgeon's conversation compared to the body-conscious adolescent. Children may also lack the coping strategies of adults. Clearly a lot more work needs to be done in the paediatric population before depth monitoring can be advocated as "a must" for reducing awareness in this age group.

Depth of anaesthesia monitors are also useful for total intravenous anaesthesia (TIVA). TIVA does not have the advantage of blood concentration monitoring similar to the end tidal estimates possible with volatile anaesthesia. An end organ effect monitor would therefore be more useful in TIVA. In paediatric anaesthesia, TIVA is less frequently used. One reason for this is the relatively poorly developed algorithms compared to adult TIVA.[22] Use of depth monitors may encourage the development of better algorithms and increased use of TIVA in paediatrics.

**Outside theatre**

The use of depth of anaesthesia monitors is being advocated to monitor sedation in the ICU and the emergency departments. It is plausible that these monitors could be better suited to sedation rather than anaesthesia, as in the algorithm design there were more meaningful end-points at higher planes of sedation/anaesthesia. Numerous studies have already explored the use of the BIS for sedation, sometimes demonstrating remarkable reliability.[23]

Measuring pain and sedation in paediatric ICU patients is particularly challenging. Demonstrating any improved outcome is difficult in the ICU setting. Nevertheless, there may be a role for these monitors in the paediatric ICU.[24-26]

Giving sedation to children for procedures in the emergency department is a controversial issue amongst paediatric anaesthetists. These children deserve effective and safe sedation and any advance in this area would be most welcome. BIS monitors are being used in paediatric emergency departments, but far more careful and extensive outcome analysis needs to be done before their widespread use can be recommended.[27]

**Conclusion**

We are still far from the perfect depth monitor. Despite this, interest in measuring anaesthetic effect is likely to continue and increase. As we learn more of the neurobiology of the effects of anaesthesia and as new monitors and new applications of existing monitors emerge, we must make the effort to specifically investigate the paediatric population. Children are different; their neurophysiology and the concepts of consciousness are distinctly different to adults. New drugs and technologies often neglect children. Children may well benefit from depth monitoring but simply applying adult technology to younger patients could be useless or misleading.

**References**

1. Bruhn J, Ropcke H, Hoeft A. Approximate entropy as an electroephalographic measure of anaesthetic drug effect during desflurane anesthesia. *Anesthesiology* 2000; 92:715-726.
2. Frith C, Frith Y. Interacting minds — a biological basis. *Science* 1999; 286:1692-1695.
3. Faulconer A, Bickford R. Electroencephalography in anesthesiology. Charles C Thomas, Springfield, USA 1960.
4. Scher MS. Pediatric neurophysiologic evaluation. In: Swaiman KF, Ashwal S, eds. Pediatric Neurology. Mosby, St Louis, USA 1999; 142-181.
5. Niedermeyer E. Maturation of the EEG: development of waking and sleep patterns. In: Niedermeyer E, Lopes da Silva F, eds. Electroenephalography: basic principles, clinical applications and related fields. Williams and Wilkins, Baltimore, USA 1999; 189-214.
6. Kitahara Y, Kojima Y, Nozaki F. Aperiodic analysis of the EEG in children during halothane anesthesia. *Marsui* 1990; 39:396-402.
7. James PD, Volgyesi GA, Burrows FA. α-Pattern EEG during pediatric cardiac operations under isoflurane anesthesia. *Anesth Analg* 1986; 65:525-528.
8. Sugiyama K, Joh S, Hirota Y, et al. Relationship between changes in power spectra of electroencephalograms and arterial halothane concentration in infants. *Acta Anaesthesiol Scand* 1989; 33:670-675.
9. O'Kelly SW, Smith DC, Pilkington SN. The auditory evoked potential and paediatric anaesthesia. *Br J Anaesth* 1995; 75:428-430.
10. Denman W, Swanson EL, Rosow D et al. Pediatric evaluation of the Bispectral Index (BIS) monitor and correlation of BIS with end-tidal sevoflurane concentration in infants and children. *Anesth Analg* 2000; 90:872-877.
11. Degoute C, Macabeo C, Dubreuil C et al. EEG bispectral index and hypnotic component of anaesthesia induced by sevoflurane: comparison between children and adults. *Br J Anaesth* 2001; 86:209-212.
12. Davidson A, McCann M, Devavaram P et al. Differences in the bispectral index between infants and children during emergence from anesthesia after circumcision surgery. *Anesth Analg* 2001; 93:326-330.
13. Rampil I, Cai H, Embree P, Greenwald S. "Paradoxical arousal" during bispectral index monitoring. *Anesthesiology* 2000; 93:A107.
14. Watcha M. Investigations of the bispectral monitor in pediatric anesthesia: first things first. *Anesth Analg* 2001; 92:805-807.
15. Gan T, Glass P, Windsor A, et al. Bispectral Index monitoring allows faster emergence and improved recovery from propofol, alfentanil and nitrous oxide anesthesia. *Anesthesiology* 1997; 87:808-815.
16. Bannister CF, Brosius KK, Sigl JC et al. The effect of Bispectral Index monitoring on anesthetic use and recovery in children anesthetized with sevoflurane in nitrous oxide. *Anesth Analg* 2001; 92:877-881.
17. Moerman N, Bonke B, Oosing J. Awareness and recall during general anesthesia. Facts and feelings. *Anesthesiology* 1993; 79:454-464.

18. Davidson A. Awareness and paediatric anaesthesia. *Paed Anaesth* 2002; 12:567-568.
19. McKie BD, Thorpe EA. Awareness and dreaming during anaesthesia in a paediatric hospital. *Anesth Intens Care* 1973; 1:407-414.
20. Hobbs AJ, Bush GH and Downham DY. Peri-operative dreaming and awareness in children. *Anaesthesia* 1988; 43:560-562.
21. O'Sullivan EP, Childs D, Bush GH. Peri-operative dreaming in paediatric patients who receive suxamethonium. *Anaesthesia* 1988; 43:104-106.
22. McFarlan CS, Anderson BJ, Short TG. The use of propofol infusions in paediatric anaesthesia: a practical guide. Paed Anaesth 1999; 9:209-216.
23. Palmer GM, Davidson AJ, Sethna N, Laussen P. Use of the Bispectral Index monitor to aid titration of propofol during a drug-assisted interview. *Paed Anaesth* 2001; 11:245-248.
24. Berkenbosch JW, Fichter CR, Tobias JD. The correlation of the Bispectral Index monitor with clinical sedation scores during mechanical ventilation in the pediatric intensive care unit. *Anesth Analg* 2002; 94:506-511.
25. Crain N, Slonim A, Pollack MM. Assessing sedation in the pediatric intensive care unit by using BIS and the COMFORT scale. *Ped Crit Care Med* 2002; 3:11-14.
26. Fletcher JE, Aneja R, Heard A, Kielma D, Heard C. monitoring the sedation of children receiving neuromuscular blocking drugs in the PICU - the use of the Bispectral Index (BIS). *Anesthesiology* 2001; 95:A402.
27. Overly FL, Fontaine B, Wright RO, Jay GD, Linakis JG. Bispectral analysis in the pediatric emergency department during conscious sedation. *Acad Emerg Med* 2002; 9:431.

# Anaesthesia — "A Modern Concept"

ROGER TRAILL, M.B., B.S.Hons, FANZCA
Department of Anaesthesia, Royal Prince Alfred Hospital

*Roger Traill is a senior staff specialist at Royal Prince Alfred Hospital; he is also Director of Facilities within the Department of Anaesthesia and Director of the Peri-operative Service. His anaesthetic practice is primarily in neuroanaesthesia.*

On October 16, 1846 at the Massachusetts General Hospital, Thomas Morton gave the first successful public demonstration of general anaesthesia. He anaesthetised Edward Gilbert Abbot, so that John Warren could remove a vascular tumour from the left side of his neck. Warren then uttered the famous words, "Gentlemen, this is no humbug." From that moment, the use of ether anaesthesia spread around the world and surgery was revolutionised. Thus, Thomas Morton is commonly honoured as the "discoverer of anaesthesia" and October 16, 1846 as the date of the birth of anaesthesia. Unfortunately, this is not true and simply confirms that those who publish first are most likely to have their ideas adopted. Once a fiction is established, it is difficult to reverse. The real honour of discovery belongs to Crawford Long who, on March 30 1842, gave the first successful ether anaesthetic in Jefferson, Georgia. Unfortunately for him, Long did not publish his results until 1849 and history barely records his efforts, although his statue stands in Washington's Capitol building as one of Georgia's two most famous people.

As an aside, Crawford Long also submitted the first anaesthetic bill in history. He charged $2.00 for the surgery, which he also performed, but only 25 cents for the anaesthetic. So started the historical imbalance between surgical and anaesthetic fees which exists to this very day. Perhaps Long really does deserve little honour!

Oliver Wendell Holmes first coined the term "Anaesthesia" in 1846 as a single word to apply to the newly described state. The Professor of Anatomy and Physiology at Harvard, Holmes is better known as a humorist and poet. Soon after this, in 1847, Plomley[1] described three stages of anaesthesia in a letter to the *Lancet*. He wrote, "I have breathed the ether on several occasions, and think its effects may be divided into three stages or degrees, the first is merely a pleasurable feeling of half intoxication; the second is one of extreme pleasure, being similar to the sensations produced by breathing nitrous oxide, or laughing gas; ... The third stage, the only one, I think, for performing operations in, is one of profound intoxication and insensibility." Later that same year John Snow, the first physician anaesthetist, wrote, "I shall divide the effects of ether into five stages or degrees; premising, however that the division is, in some measures, arbitrary — that these different degrees run gradually into each other, and are not always clearly to be distinguished."[2] The concept of "depth of anaesthesia" was born from Snow's ideas.

Arthur Guedel was a young doctor who served with the American forces in World War I. He was put in charge of the anaesthetic services for the US Army in Vosges,

France. As there were few doctors involved, it was necessary for non-professionals to give ether anaesthesia and Guedel provided supervision by rounding between the hospitals on a bicycle! To ensure the safe administration of ether, he devised a system for checking on the patient's condition. After his return to the US, Guedel published his first article on the "Signs and Stages of Ether Anesthesia" in 1920,[3] followed by a book in 1937.[4] His vertically oriented tables fitted neatly with the concept of "depth", with a progression from awake, through stages of anaesthesia and, finally, to death. The stages were described by a series of observations based on breathing, muscle relaxation, pupils, lacrimation and eyelid reflexes.

Artusio in 1954[5] refined this further, describing three planes of Stage 1 of ether anaesthesia (Table 1). He referred to Plane 3 as "Amnesia Wakefulness". Cardiac surgery could be performed whilst patients were in this state!

**Table 1**
Planes of stage 1 of ether anaesthesia

| | | |
|---|---|---|
| Plane 1 | No Analgesia | No Amnesia |
| Plane 2 | Partial Analgesia | Total Amnesia |
| Plane 3 | Total Analgesia | Total Amnesia |

In 1950, Rees and Gray[6] published their seminal paper in which the famous "triad of anaesthesia" was described. This was said to consist of narcosis, relaxation and analgesia. The concept is still taught today and anaesthetists commonly refer to giving patients "analgesics" during an anaesthetic. Interestingly, the concept was mentioned as an aside in the paper, which was about the anaesthetic effects of methyl-n-propyl ether. The sole reference to the "triad" was a single sentence, "It might, however, give the drug a certain value in a balanced anaesthetic when a relaxant is used to complete the triad of relaxation, narcosis and analgesia". There is no discussion around this concept, and I can find no other reference to it, until another paper by Gray and Rees in 1952.[7]

They start this paper by saying, "There has been widespread support for the concept which views anaesthesia as a triad", although they provide no references to support this assertion. In the paper, they described the "Liverpool Technique" combining the use of a muscle relaxant, nitrous oxide and hyperventilation. The conclusion suggests: "The time has come, in our opinion, to substitute for the triad principle of anaesthesia outlined at the beginning of this paper a tetrad, which can be regarded as a pyramid which has a base apnoea upon which are constructed the sides of the pyramid — narcosis, relaxation and analgesia." Whilst the "triad" and the "Liverpool Technique" became popular, the "tetrad" does not appear to have resonated with the anaesthetic community; this is the only reference to the "tetrad" I could find.

In 1960, Mushin[8] noted, "One of our commonest phrases is 'depth of anaesthesia'. We speak of getting the patient 'deeper', of getting him 'under', of getting him 'down'. All these words imply movement in a vertical direction: towards the grave." The vast majority of anaesthetists still use these terms. Virtually all medical students are taught about "depth of anaesthesia" and, as a consequence, all our surgical colleagues use the term as well. The famous "triad" is usually taught as part of this anaesthesia education as well.

Once this concept of depth had become firmly entrenched in the minds of anaesthetists, the search was on to find a measure of it. For if only we could measure

it, we could then "scientifically" deliver it! A huge range of possible "measures" have been studied, from clinical scores (PRST), EEG (and processed derivatives especially the BIS), evoked potentials (auditory mid-latency potentials being the most popular), EMG (frontalis muscles), skin vasomotor tone and lower oesophageal sphincter tone. The search continues today, with a myriad of articles published each year on "depth of anaesthesia". A Medline search done in preparation for this manuscript, using the words "depth" and "anesthesia", revealed more than 350 articles published in the last ten years. Many of these strived to deliver better and better measurements of "depth". In an erudite paper discussing the ways in which one might statistically derive a measure of depth of anaesthesia, Smith et al[9] talk in an abstract manner of this being "judged against a gold standard indicator of anaesthetic depth", without in any meaningful way discussing what the term might mean.

Interestingly, whilst "depth" is still the dominant paradigm (a collection of beliefs shared by scientists), a more logical way of dealing with the problem of "anaesthesia" has arisen in parallel, although this has not succeeded in supplanting the old ideas. In 1957 Woodridge[10] wrote, "How deep is this patient? That question has become more puzzling as time goes by." He suggested that there was not a single entity of anaesthesia, but that it could be broken into several components, which he called "nervous depression" of sensory, motor, reflex and mental systems. Unfortunately, at the end of this article he decided that "anaesthesia" was not even the best word and proposed that "northria" (a Greek word meaning torpor) was more appropriate and that we should all be called "Northrotists"! Whilst this may have been easier for the average patient to say and spell, it failed to have a lingering effect in our field.

Cecil Gray published the single most important article on the concept of "Depth of Anaesthesia" in 1960.[11] In this article, he debunked the whole concept of "depth", in particular producing a demolition of many of the "signs" put forward by Guedel to indicate stages and planes. He proposed a modified "triad", consisting of "narcosis" (unrousable unconsciousness), "reflex depression" (replacing "analgesia") and "relaxation". Amazingly, for a man so influential in this whole debate, he chose to recant his previous views in the *Irish Medical Journal*. It was almost as if he were somehow ashamed of his previous views and, whilst wishing to come clean on the issue, did not want to make it a very public retraction. Not surprisingly, I have never seen this article quoted in any article dealing with this topic! Others have written since then questioning the validity of the term "Depth of Anaesthesia".[12-17] Amongst them; Pinsker[12] suggested, "... paralysis, unconsciousness, and attenuation of stress response. This is complete anesthesia." This view is remarkably similar to that of Gray's, which he does not quote.

Developing this theme, Prys-Roberts[13] stated, "There cannot be degrees of anaesthesia, nor for that matter can there be variable depths of anaesthesia. The continuing search for some method to measure anaesthetic depth resembles that for the Philosophers Stone." Kissin[14] agreed regarding the search for a single measure of anaesthetic depth. He suggested that anaesthetic action consisted of "different actions used to achieve variable goals of anesthesia". In a paper titled "Monitoring Depth of Anaesthesia", Schneider and Sebel[15] wrote, "In the modern practice of anaesthesia, the term 'depth of anaesthesia' and the definition of stages are irrelevant. Anaesthesia is not 'deep' or 'light': it may or may not be adequate." Why they then gave the article the title they did is a little harder to understand. With similar thinking to Kissin, Eger[16] felt that there were two components to anaesthesia, "immobility" and "amnesia",

suggesting that each resulted from actions at separate anatomical sites. Despite these ideas, the search continues for the measure of "depth of anaesthesia". Kissen[17] recently stated, "the term depth of anesthesia becomes irrelevant for major components taken together … (but) it could still be relevant for each of the components measured separately".

It is clear that the paradigm started in 1847 is hard to break. How are we to deal with this issue in a way that will allow us to deliver anaesthesia in a logical manner?

**In order to communicate we need to define some terms**

The most fundamental of these is the question of consciousness. Many textbooks have been written on this topic, despite the simple statement of Rene Descartes (1596-1650), "I think therefore I am", as a self-evident truth. William James (1842-1910), the American philosopher and leader in the philosophical movement of "pragmatism", stated in 1892,[18] "Everyone knows what consciousness is until he tries to define it". Stanley Cobb used the definition "Awareness of environment and of self" in 1948.[19]

From a contrary viewpoint, Feldberg[20] suggested, "There is no need to define unconsciousness. We all know what it means". I think that, for the most part, that this statement is true. We all see our family asleep each night and can usually judge if they are unconscious. Sleep is a physiological (as apposed to pharmacological) state of unconsciousness from which we can be roused. However, it must be conceded that, at its most fundamental level, the interpretation of whether or not someone is conscious is an internal one; i.e. we assume someone is unconscious when they fail to respond to some stimuli by interacting with us in a purposeful manner, for example, by opening their eyes to command. But what if a person chooses not to respond? They would then be clearly conscious, but we would not be able to determine this. We could also imagine a circumstance in which a person could be aware of themselves but not of their environment, as in a sensory deprivation tank. Also, if a person is unable to respond when paralysed, then it becomes very difficult to determine if indeed they are conscious.

The definition that I prefer to use is:

**Consciousness** **Awareness of one's self**

therefore:

**Unconsciousness** **Loss of Consciousness**

Now whilst a person may be unconscious, it does not mean that he or she cannot be roused (made conscious) by some stimuli. Rousability is a separate, although related, issue.

**Rousable** **Someone unconscious can be made conscious by stimulation, e.g. talking, shaking.**

"Pain" is, in some ways, a prototypical conscious experience. Certain stimuli (noxious) will produce "pain" in the conscious person in addition to reflex responses. We can use the following definition:

**Pain:** **The (usually) unpleasant sensation associated with actual, potential or perceived tissue damage.**

I say "usually unpleasant" because there is that small subset of the community who finds pain pleasurable! As well, "pain" does not have to be associated with tissue damage. Indeed, as a protective mechanism, it should occur in situations before tissue damage occurs, so this can be prevented by the person's response to the "pain". There

are also situations where the person perceives "pain" when no actual or potential tissue damage exists, e.g. phantom limb pain or during application of an electrical current as a torture mechanism. (Provided the current is not excessive, this will produce no damage at all, e.g. a nerve stimulator's tetanic setting). "Pain" is simply the conscious interpretation (or sensation) of these neural signals.

**Analgesia:** **The relief or prevention of "pain".**

**Nociceptive:** **The nervous impulses associated with tissue damage. These produce "pain" (in the conscious patient) and reflex responses.**

**Aims of General Anaesthesia**

There are really only two aims for general anaesthesia:

**1. Narcosis**, or unrousable unconsciousness

Unconsciousness must surely be the first and most important aim for a general anaesthetic. However, that alone is insufficient as the patient must also be unrousable. Sleep, from which one can roused, should be clearly seen to be different. There may be times when it may not be possible or safe to achieve unconsciousness, eg during surgery for ruptured aortic aneurysm.

"Amnesia" (the failure to remember events that occurred when the patient was conscious) and "Analgesia" (as defined above) are often stated to be aims. However, these clearly imply that the patient is conscious; if we accept the first aim, then these terms have no meaning in the setting of general anaesthesia and so should not be used. We may give substances during an anaesthetic so that the patient will awake pain free, but we are not giving them as analgesics during the general anaesthetic. Opioids given during "general anaesthesia" are primarily used as reflex depressants (see later).

Similarly, there may be times in which we would wish a patient to be amnesic (during the performance of a difficult block). However, if we aim to have an unconscious patient, then amnesia is not an aim as such during "general anaesthesia". (Of course, all drugs that produce unconsciousness will also produce amnesia in lower doses.)

Consciousness is a quantal response, one is either conscious or not. One should not confuse the difficulty in determining whether or not a paralysed patient is conscious with its quantal nature. If unconscious, then one may be either "rousable" or "unrousable", depending on the particular stimulus. Scott and White[21] have pointed out that many Sedation Scores are, in fact, measures of rousability. Table 2 shows a typical sedation score:

Note that the patient is actually conscious only with a score of 4 or 5. With scores of 3, 2, or 1, the patient is unconscious but "rousable". With a score of 0, they are "unrousably unconscious" (the primary aim of general anaesthesia). What physicians and neurosurgeons mean when they speak of a "deeply unconscious" patient is that the

**Table 2**
Sedation score

| | |
|---|---|
| 5 | Responds readily to spoken word in normal tones |
| 4 | Lethargic response to name spoken in normal tones |
| 3 | Responds only after name is called loudly and/or repeatedly |
| 2 | Responds only after mild prodding or shaking |
| 1 | Responds only after painful trapezius squeeze |
| 0 | Does not respond to painful trapezius squeeze |

patient is "unrousably unconscious" and has a variable degree of reflex depression, i.e. flexor or extensor responses only. The conscious patient's level of mentation, (e.g. oriented, alert, confused) can be assessed, but these are not measures of consciousness; rather, they are measures of higher cortical function.

Induction of anaesthesia is first and foremost about rendering the patient "unrousably unconscious". One determines this by the same means we assess whether someone is conscious in any other non-medical setting, by talking to the patient. The common habit of trying to determine if the eyelash reflex has gone is clearly not appropriate, as not only is it a test of a reflex (and therefore unrelated to consciousness) but is also unpleasant should be patient actually be conscious.

**2. Reflex Depression**

Whilst a patient may be "unrousably unconscious", this is not sufficient for general anaesthesia. The same nociceptive impulses that would have produced pain had the patient been conscious may also produce other non-conscious (or reflex) responses. Reflexes are, by their very definition, non-voluntary responses and so have nothing whatsoever to do with consciousness. Again, we should not get confused about the separate nature of the aims just because we commonly give drugs that may produce both "unrousable unconsciousness" and "reflex depression".

Reflexes can be classified into:

**Motor Reflexes**, e.g. movement, coughing

**Autonomic Reflexes**

**Cardiovascula**r, e.g. BP and HR changes

**Neuro-endocrine**, e.g. cortisol, vasopressin increases

These can be modified at any point in the reflex arc, consisting of a receptor organ, an afferent limb, central processing area (spinal, brainstem, cortical), an efferent limb and an effector organ, (e.g. skeletal or vascular muscle, heart, endocrine gland).

There is no pre hoc way of describing reflex responses; one can only say whether a response has occurred and, if so, to what degree. The reflex arc can be interfered with at any one (or more) of the points in the arc. We may very specifically block particular reflexes (e.g. neuromuscular junction (NMJ) blockers for motor reflexes or beta-blockers for HR responses) whilst leaving others unaffected. We might also use agents, e.g. opioids, which have their major effect by blocking central processing (spinal cord, brainstem).

The responses can clearly be graded, but each reflex type has different responses to different drugs and differs between patients, as demonstrated by MAC and MAC BAR.[22] Sometimes, one reflex will be blocked whilst another is present, as with a patient whose BP increases with incision, but who does not move because paralysed. Some patients will cough on incision as the sole motor response. There is a multitude of reflex responses, each with its own pathway and intermediate relays and neurotransmitters. The neuro-endocrine responses, in particular, are especially varied, although not able to be measured easily at the time of anaesthesia and surgery.

Clearly, as reflex depression has nothing to do with consciousness, or lack thereof, and as we can break reflex responses into component parts, each of which can be blocked separately, it is impossible to have a single measure of all these things. Hence, there can be no such thing as "depth of anaesthesia" and, philosophically, if it does not exist it cannot be measured.

It is also now clear that the effects of anaesthetic agents in producing

unconsciousness and interfering with reflex responses, occur via differing mechanisms, with the movement responses being, to a large extent, mediated at a spinal cord level.[23] Thus, there are pharmacological as well as physiological reasons to consider reflex depression as a separate aim to unrousable unconsciousness, even if both are achieved with the one drug.

**Decreased Muscle Tone**

The literature has got itself confused in this area. Authors often use the term "muscle relaxation" as a requirement of anaesthesia. However, what is really meant is a lack of motor response to surgical and anaesthetic stimuli (motor reflex depression). Provided a patient does not move during a procedure, then a decrease in muscle tone is not always needed, e.g. during breast biopsy. There are many occasions when a decrease in muscle tone is needed, but it is not a universal requirement. When needed it can be achieved in a number of ways. Muscle tone is a function of:

1. efferent (motor) nerve activity,
2. neuromuscular junction activity, and
3. muscle function/ mass/ resting fibre length.

We can decrease tone by interfering with one or more of these. However, we are only able to objectively measure the second mechanism, with nerve stimulators and either the force or EMG response to the stimuli, as with the Train of Four response. There are many clinical situations where the muscle tone is low without any pharmacological intervention, e.g. spinal cord injury (decreased efferent activity), myasthenia gravis (inadequate neuromuscular junction function), myopathies (poor muscle function), the cachectic patient (decreased muscle mass) and post-caesarean section (abdominal muscles lax due to loss of stretch). In these situations, no pharmacological intervention may be needed to provide the low muscle tone that surgery may require.

Clearly, we need to consider decreased muscle tone as a separate, but not always necessary, part of the anaesthetic care of the patient. It is not a part of "general anaesthesia" per se. A decrease in muscle tone is highly dependant on the mechanisms by which a particular drug achieves this and different patients will respond quite differently. One needs to think clearly as to how one might achieve a decrease in muscle tone (if indicated) and how one might monitor it. Whilst many drugs that interfere with motor reflex responses at the spinal cord level or higher will also tend to decrease efferent motor activity and, hence, decrease muscle tone (volatile agents), some (e.g. opioids), may actually increase muscle tone and cause rigidity in high doses. Depolarising and non-depolarising NMJ blockers are commonly used to achieve a decrease in muscle tone and their effects should always be monitored with a nerve stimulator.

It must be remembered that the muscle tone achieved will be a function of all the factors listed above, not just the function of the neuromuscular junction. A patient with a high $P_aCO_2$ may show diaphragmatic movement, even thought they have barely one twitch visible in response to a TOF stimulus. The most appropriate response for the anaesthetist could be lowering the $P_aCO_2$ or giving opioids or other respiratory depressants, rather than just giving more NMJ blockers.

**Conclusions**

Consciousness is a quantal phenomenon, rousability can be graded, reflexes are quite independent of consciousness and reflexes themselves, whilst gradable, come in

a multitude of forms many of which can be individually depressed. As a consequence, there can be no single measure of "depth of anaesthesia"; this is a philosophically meaningless term and cannot possibly be measured.

As Prys-Roberts[13] wrote, "the search for a measure of Depth of Anaesthesia is the modern day equivalent of the search for the Philosopher's Stone". Those who keep up with the latest literature will know that the Philosopher's Stone was destroyed by Professor Dumbledore to prevent the evil Lord Voldemort from acquiring it (Harry Potter and the Philosophers Stone, 1997). With the stone gone, it is time to for us to abandon the long lived and futile search for it and the related depth of anaesthesia.

1. As Gray wrote[11] in 1960 (and was ignored):
   "I suggest that these concepts 'stages of anaesthesia' and 'depth of anaesthesia'— no longer serve any useful purpose and should no longer be taught. I believe that they should be relegated to the museum, already crowded, for outworn concepts ..."
2. One should therefore describe the patient in terms of the aims described above.
3. General Anaesthesia can thus be defined as: **A reversible iatrogenic state characterised by unrousable unconsciousness and reflex depression.**

By logically dealing with every general anaesthetic in this way, one can then deal appropriately with the management of each patient. By ensuring that the patient remains unrousably unconscious and their reflex responses are depressed (usually so they don't move and their BP and HR are within acceptable limits) and by considering these as separate issues (even if managed with a single agent), one can optimally manage each general anaesthetic. If a decrease in muscle tone is needed it also should be thought of in a logical way, but not be considered part of "general anaesthesia", rather just another part of the intraoperative management of the patient, just as is keeping them warm or cold and well hydrated. It is time to abandon the old paradigm and embrace the new.

**References**


1. Plomey F. Operations on the Eye. *Lancet* 1847, I:134-135.
2. Snow J. On the Inhalation of the vapour of Ether. 1847, Churchill London.
3. Guedel AEA. Systemization of the Signs of Inhalational Anesthesia. Current Research in Anesthesia 1920, Macmillan New York.
4. Guedel AE. Inhalational Anaesthesia — A Fundamental Guide 1937, Macmillan New York.
5. Artusio JF. Di-ethyl ether Analgesia: a detailed description of the first stage of ether analgesia in man. *J Pharmacol* 1954; 111:343.
6. Rees GJ, Gray TC. Methyl-n-propyl ether. *Brit J Anaesth* 1950; 22:83.
7. Gray TC, Rees GJ. The role of apnoea in anaesthesia for major surgery. *Brit Med J* 1952; 2:891.
8. Mushin WW. Signs of Anaesthesia. In: General Anaesthesia. FT Evans (ed). London, Butterworths 1960.
9. Smith W, Dutton R, Smith N. Measuring the performance of anaesthetic depth indicators. *Anesthesiology* 1996; 84:38-51.
10. Woodridge P. Changing concepts concerning depth of anaesthesia. *Anesthesiology* 1957; 18:536-550.
11. Gray CA. Reassessment of the signs and levels of anaesthesia. *Irish J Med Science* 1960; 499-508.
12. Pinsker MC. Anesthesia: a pragmatic construct (Letter). *Anesth Analg* 1986; 65:819-820.
13. Prys-Roberts C. Anaesthesia: a practical or impractical construct? *Brit J Anaesth* 1987; 59:1341-1345.
14. Kissin I. General anesthetic action: an obsolete notion? (Edit). *Anesth Analg* 1993; 76; 215-218.
15. Schneider G, Sebel PS. Monitoring depth of anaesthesia. *Europ J Anaesthes* 1997; 15(Suppl):21-28.
16. Eger EI, Koblin D. A step towards consensus on general anesthesia. *Anesth Analg* 1998; 86:446.
17. Kissin I. Depth of anesthesia and bispectral index monitoring. Anesth Analg 2000; 90:1114-1117.
18. James W. *Textbook of Psychology* 1892. Macmillan, London
19. Cobb S. Foundations of Neuropsychiatry, 4th Ed 1948. Williams and Wilkins, Baltimore.
20. Feldberg WA. Physiological approach to the problem of general anaesthesia and of loss of consciousness. *Brit Med J* 1959; 771-781.

21. Scott RW White PF. Sedation scales: Measures of calmness or somnolence. *Anesth Analg* 1993; 76:219-221.
22. Stanski DR. Monitoring Depth of Anaesthesia. In Miller RD, ed. Anesthesia 5th Ed. Churchill Livingston, Philidelphia 2000; 1095-1100.
23. Antogninini JF. Exagerated anesthetic requirements in the preferentially anesthetized brain. *Anesthesiology* 1993; 79:1244-1249.

# Fitness to Drive After Intravenous Sedation and General Anaesthesia: A Literature Review

P. F. TUCKER, MB, BS(Hons), BMedSci
C. R. CHILVERS, BMed, DA(UK), BMed, FANZCA
Department of Anaesthesia, Launceston General Hospital, Launceston, Tas.

*Dr Pat Tucker is a second year accredited anaesthetics registrar currently working at the Launceston General Hospital. She graduated from the University of Tasmania in 1996 and has worked in both NSW and Tasmania. Dr Chilvers is a consultant anaesthetist. He has previously worked as a Clinical Research Fellow in Ambulatory Anaesthesia in Vancouver, Canada. Currently, he is the anaesthetist responsible for the Day Procedure Unit at the Launceston General Hospital.*

**Introduction**

Current recommendations by the Australian and New Zealand College of Anaesthetists are that after undergoing general anaesthesia or intravenous sedation, a patient should not drive a car until the following day.[1] In the United Kingdom patients are advised not to drive for 24 hours. In this age of modern day case anaesthesia is this based on sound evidence or are we maintaining the status quo? To date there has not been a systematic review or meta-analysis of this topic, i.e. level I evidence as defined by the National Health and Medical Research Council.[2] In this paper the available evidence is examined. The anaesthetic agents discussed are those most commonly used in current day case anaesthesia. Fitness to drive is assessed by psychomotor function tests, multiple sleep latency tests, and simulated driving tests. Recovery after undergoing sedation or anaesthesia, is also compared with functioning at the legally safe blood alcohol levels (BALs) for driving as recommended by the World Health Organisation.[3]

**Clinical Trials**

Investigators have been studying the effects of general anaesthesia on driving ability for decades. In 1972, Ogg found that 73% of car owners drove within 24 hours after undergoing a general anaesthetic, and 9% actually drove home![4] Recently the percentage of patients driving after an anaesthetic has fallen, however Correa et al found that 4% of patients still drove within 24 hours, despite advice not to.[5] Most agree that patients should never drive immediately after anaesthesia but for what period remains debatable. In the 1970s Havard wrote: "With most general anaesthetics, it is safer to advise against driving for 48 hours afterwards."[6] This was based on the finding that tissue concentrations of anaesthetics and/or their metabolites are evident for 48 hours.

Articles used for this literature review were initially found by a search of the National Library of Medicine. Key words used were: general anaesthesia, driving, midazolam, diazepam, propofol, fentanyl, alfentanil, nitrous oxide, desflurane, sevoflurane, and isoflurane. References from studies found by the above means were then used to find further studies. Trials were excluded if: they involved anaesthetic agents not commonly

in current use or not available in Australia; aimed at studying short term recovery and recovery room behaviour only; studying recovery after operations involving overnight hospital stays (except reference 19); not written in English. The trials reviewed and listed in the tables are all randomised controlled trials or equivalent, i.e. level II evidence. (2)

**Relationship between Anaesthetic Agents, Driving skills, Psychomotor Tests, and Multiple Sleep Latency Tests**

Driving is a complex task. It involves attention, information processing, judgement, sensorimotor skills, and perception.[7] All these functions are affected by the anaesthetic agents used in day case anaesthesia. Research has attempted to evaluate these anaesthetic effects in the hope of gaining sufficient evidence to confidently identify a prudent non driving period.

The clinical trials quoted in this review utilised testing of a variety of psychomotor skills, multiple sleep latency time periods, and driving skills. Brief descriptions of individual psychomotor tests are available in the Appendix. Practice (or learning) effect, whereby results improve with repeated testing, is mentioned where applicable to a study. The tables indicate if training has been used to minimise this effect.

Psychomotor studies (upon which most of this review is based) generally define "recovery time" as the time taken for the mean test results of the experimental group to return to baseline or control levels. Unfortunately this methodology could fail to identify an individual subject who is much slower to recover than the rest of the group. For safety in real life we would prefer to know the recovery time the very slowest person could take. In addition there is conflicting evidence on the sensitivity of the different psychomotor tests. Tracking tasks, the peg board test, the maddox wing test, and perceptive accuracy tests appear to be the most sensitive psychomotor tests. Choice reaction tests, critical flicker fusion threshold, and free recall appear to be the most controversial. The recovery of free recall was inconsistent with other tests in most papers and has therefore been ignored in the summaries made in this review. Marshall et al (1992) found that control subjects also had impaired free recall with subsequent testing.[8]

Multiple sleep latency tests are an alternative but valuable method of assessing if a person is fit to drive. Sleep latency is measured as the time taken to reach the first epoch of non-wakefulness.[9] Improved driving performance has been shown to be associated with increased sleep latency times.[10] Lichtor et al (2002) found these tests to be more sensitive than psychomotor tests.[11]

Actual car driving ability is the gold standard in assessing fitness to drive after an anaesthetic. This is especially true when ability is compared with subjects at the legal BAL for driving. However, actual driving is often not a practical option. Driving simulators can be used instead. The main features recorded from the simulators are brake reaction times and performance errors (neglected instructions, driving off the road, and collisions).[12]

The ultimate test of fitness to drive, the motor vehicle accident mortality and morbidity figures from recently anaesthetised patients driving, cannot of course be studied. Interestingly, a review of motor accidents in Tasmania over the past 15 years has shown no deaths have occurred which could be attributed to either general anaesthesia or sedation. (Personal communication, Kathryn Campbell, Government Analyst.) No information was available on motor accidents not resulting in death.

The electroencephalogram (EEG) was also performed in one study. (13) The relationship between EEG abnormalities and driving ability has not yet been determined.

**Midazolam and Diazepam Trials**

Many trials have studied recovery after midazolam and diazepam used for sedation. These are listed in Table 1.

In summary, the effects of both midazolam and diazepam on psychomotor tests and car driving ability are essentially gone by 10 hours.

**Propofol Trials**

Propofol is a commonly used both as an anaesthetic and a sedation agent. Details of relevant research on propofol are given in Table 2.

To summarise, the effects of propofol on psychomotor function appear to be largely dissipated by 2 hours post anaesthetic. At a blood propofol concentration of 0.2 mcg/ml psychomotor impairment is equivalent to a BAL of 20 mg/100 ml.

**Fentanyl and Alfentanil Trials**

Few clinical trials have studied the recovery of psychomotor function and driving ability post fentanyl or alfentanil when administered alone. Details of the available trials are summarised in Table 3.

In the very limited studies performed, recovery (apart from EEG changes) after a single dose of fentanyl or alfentanil appears to occur by 6 hours. The recovery time of fentanyl is effected by the dose administered.

**Inhalational Anaesthetic Trials**

Details of the available trials of inhalational agents are given in Table 4.

It can be seen that testing of the volatile agents in isolation in regard to post-anaesthetic fitness to drive is limited. Studies conducted have been relatively small and time-limited. Objective recovery after $N_20$ appears to be consistently less than 30 minutes, but subjective recovery has been reported as being prolonged for up to 8 hours. Recovery occurs after desflurane in 3 hours. Sevoflurane recovery has not been adequately studied. However, with its pharmacokinetic profile, it could be assumed to lie somewhere between the 2-3 hour recovery period for desflurane and the 5-7 hour recovery period for enflurane and halothane.

**Combined Anaesthetic Agent Trials**

In practice most sedative and anaesthetic regimens utilise a combination of agents. Table 5 summarises these studies.

Trials studying a combination of midazolam, fentanyl, and propofol in sedative doses have found a recovery time of within 8 hours. Studies using propofol as the maintenance agent have revealed a recovery time of 1 hour. Trials utilising propofol and $N_20$ as the maintenance agents have shown a recovery time of 3 hours or less. Studies using a desflurane/$N_20$ anaesthetic have found a recovery time of within 3 hours. Trials studying recovery after an isoflurane/$N_20$ maintenance anaesthetic have usually found a recovery time of 1.5 hours or less. One exception is the Marshall et al trial in which recovery only approached baseline by 3 hours.[8] In the same study, another group who had received an alfentanil infusion added to this anaesthetic did not achieve baseline levels by 5 hours. Trials studying recovery after a sevoflurane/$N_20$ maintenance combination anaesthetic are lacking. Only one study met this reviews criteria, hence, no comment can be made on recovery time.

Interestingly, recovery from a balanced anaesthetic using a combination of agents

**Table 1**
Summary of recovery after midazolam and diazepam

| Anaesthetic agent | Number subjects | Procedure | Tests used | Psychomotor recovery time period | Comments | Reference |
|---|---|---|---|---|---|---|
| Midazolam 15 mg (oral) | 7 | None | 2 psychomotor tests & car driving ability | 10 hours | Practice effect evident. Looked at recovery the next morning. No testing between 1 & 10 hours. | (14) |
| Midazolam 0.05/ 0.1/0.15 mg/kg | 11 | None | 8 psychomotor tests | 7 hours | Minimal practice effect | (15) |
| Midazolam 5/10 15/20 mg (oral) | 7 | None | 2 psychomotor tests | 7 hours | Minimal practice effect | (16) |
| Midazolam 0.1 mg/kg | 16 | None | 6 psychomotor tests | 3 hours (except maddox wing and subjective fatigue) | Minimal practice effect | (17) |
| Midazolam 0.05 mg/kg | 25 | Bronchoscopy | 3 psychomotor tests, Romberg's test & ability to walk in a straight line. | 2 hours undertaken. | No training | (18) |
| Midazolam 0.1 mg/kg | 24 | Bronchoscopy | d | >2 hours (significant impairment of ability to walk in a straight line) | No further testing after 2 hours. | (18) |
| Midazolam 0.075 mg/kg (mean) (+ spinal block) | 13 | TURP | 8 psychomotor tests | 2 hours — 3/9 tests remained impaired. 24 hours — free recall only - impaired | No testing between 2 & 24 hours. No training. Included because of comparison with propofol (see later section) | (19) |
| Diazepam 10/20 mg | 10 | None | 8 psychomotor tests & EEG monitoring | 6 hours (EEG changes & memory recall >8 hours) | Minimal practice effect | (13) |
| Diazepam 0.15/ 0.3 mg/kg | 11 | None | 8 psychomotor tests | 7 hours | Minimal practice effect | (15) |
| Diazepam 0.2 mg/kg | 27 | Bronchoscopy | 3 psychomotor tests, Romberg's test, and ability to walk in a straight line. | 2 hours | Practice effect evident | (18) |
| Diazepam 0.3 mg/kg | 11 | None | 5 psychomotor tests | 8 hours | Practice effect evident | (20) |

Note: Recovery times are based on the time taken for the mean test result to return to baseline or control levels.

**Table 2**
Summary of recovery after propofol

| Anaesthetic agent | Number subjects | Procedure undertaken | Tests used | Psychomotor recovery time period or alternative measure | Comments | Reference |
|---|---|---|---|---|---|---|
| Propofol 0.54 mg/kg + 5/4/3 mg/kg/hr Mean anaesthetic duration 40 minutes | 13 | TURP | 8 psychomotor tests | 1 hour — except free recall (normal by 24 hours) | This study compared propofol with midazolam (see earlier comments) Propofol offers a more rapid recovery. | (19) |
| Propofol (TCI) 0.8/0.4/0.2 mcg/ml | 10 | None | 3 psychomotor tests | 0.2 mcg/ml equivalent to BAL 20 mg/ 100 ml (0.02%) although compared with baseline | Good study design in that functioning at a set propofol concentration was compared with functioning at a set BAL. Training undertaken. | (21) |
| Propofol Plasma concentration 3.9±1.0 mcg/ml (mean±SD) Anaesthetic duration 1 hour | 7 | None | Subjective sleepiness, fatigue, and sleep latency | Subjective tests normalized within 1 hour 20 minutes | | (22) |
| Propofol 4.4 mg/kg | 18 | Minor gynaeco-logical | 6 psychomotor tests | 2 hours | Training undertaken however practice effect evident. | (23) |

Note: Recovery times are based on the time taken for the mean test result to return to baseline or control levels.
TCI=target controlled infusion

appears to be no longer than that using an individual agent. In fact it is usually shorter. This is probably a result of lower individual drug doses being used.

**Discussion**

In setting guidelines as to when a person is fit to drive after sedation or anaesthesia many factors need to be taken into account. Most studies test fitness to drive using psychomotor tests, simulated driving tests, or multiple sleep latency tests. Although only an indirect way of assessing the risk to the patient and the public they can be performed scientifically and have been validated against the medical and legal benchmark of BALs. Lack of impairment on one particular test may not of course equate to being fit to drive a motor vehicle. A whole battery of tests need to be investigated and practice effects need to be taken into account. The practice effect noted in approximately half the studies may bias the results towards a shorter recovery. Considerable inter-individual variation in psychomotor testing after anaesthetics has also been found.

Most of the studies were been done on fit, healthy volunteers who did not undergo any operative procedure. Patients were excluded from the studies if they were taking other medications. Many studies have also excluded subjects with organ dysfunction and/or other illnesses, for example obstructive sleep apnoea. Obstructive sleep apnoea alone may

**Table 3**
Summary of recovery after fentanyl and alfentanil

| Anaesthetic agent | Number subjects | Procedure undertaken | Tests used | Psychomotor recovery time period | Comments | Reference |
|---|---|---|---|---|---|---|
| Fentanyl a) 100 mcg | 10 | None | 8 psychomotor tests & EEG monitoring | 2 hours | Minimal practice effect. | (13) |
| b) 200 mcg | d | d | d | 6 hours EEG effects seen for up to 8 hours (Increase in frontal fast activity) | Highlights the dose dependent effect of fentanyl on recovery. | (13) |
| Fentanyl 2.5 mcg/kg | 7 | None | 2 psychomotor tests | 2 hours | Minimal practice effect | (24) |
| Fentanyl 100 mcg | 9 | None | 1 psychomotor test only — tracometer | >2 hours | Testing only undertaken for 2 hours. Practice effect evident. | (25) |
| Fentanyl 167 mcg + Etomidate 20 mg + maintenance doses | 22 | Cystoscopy | 2 psychomotor tests | >1 hour | No training undertaken. | (26) |
| Alfentanil 500 mcg + Etomidate 20 mg + Maintenance doses | 25 | Cystoscopy | d | 1 hour | No training undertaken — practice effect evident. | (26) |

Note: Recovery times are based on the time taken for the mean test result to return to baseline or control levels.

affect one's ability to drive, via increased fatigue[10], and if combined with an anesthetic these patients might have prolonged effects.

In addition, anxiety experienced by the patients prior to procedures and its effect on sleep deprivation and increased sedation/anaesthetic requirements has not been evaluated objectively in any study. The effect of post-operative pain on driving ability has not been evaluated either.

As driving is an individual action which may inadvertently affect others we need to take into account the time needed for the slowest patients to recover from the anaesthetic. It is not safe to give a time when the "average" patient has recovered. It must be the time when all patients are fit to drive. Most time periods mentioned in this review are unfortunately only "mean" times. Documenting of the range of times of recovery from anaesthetics may be a more appropriate method. There is some variation between studies and we need to weight decisions on "evidence" from those with the most conservative rate of recovery.

Conversely, it also needs to be accepted that to drive a car, total recovery after sedation or anaesthesia need not occur. Recovery of function only needs to be equivalent to that with a BAL of 50 mg/100 ml.

**Table 4**
Summary of recovery after inhalational agents

| Inhalational agent | Number subjects | Procedure undertaken | Tests used | Psychomotor recovery time period or alternative measure | Comments | Reference |
|---|---|---|---|---|---|---|
| $N_20/0_2$ 50/50 | 12 | Colonoscopy | 4 psychomotor tests | No post-procedural impairment. | Practice effect evident on one test despite training. Testing undertaken as soon as patient able to walk into testing area (within 30 minutes). | (27) |
| N20/02 50:50 | 80 | Flexible Sigmoid - oscopy | 1 psychomotor test only — adaptive tracking task. | No post-procedural impairment. Better than at a BAL of 80 mg/ 100 ml. | Compared function-ing with that at a BAL of 80 mg/100 ml & with controls. | (28) |
| $N_20/0_2$ 30:70 Anaesthetic duration 40 minutes | 11 | None | 4 psychomotor tests | 22 minutes | Compared with the above 2 studies impairment was seen. Training undertaken. | (29) |
| N20/02 Up to 50:50 | 5 | None | 5 psychomotor tests | 20 minutes Except free recall. | Subjects felt subjectively unwell for up to 8 hours. Training undertaken. | (30) |
| Desflurane ET 8.3% (mean) Mean anaesthetic duration 53 minutes | 16 | Elective day case surgery | 2 psychomotor tests | 2 hours | Training under-taken. Subjects received only local anaesthetics or oral ibuprofen for analgesia. | (31) |
| Desflurane ET 7.4% (mean) Anaesthetic duration 1 hour | 20 | None | 9 psychomotor tests | 3 hours | Training undertaken | (32) |
| Sevoflurane Fi 1-2% + $N_2O/0_2$ 67:33 Mean anaesthetic duration 17.6 minutes | 34 | Colonoscopy | 9 psychomotor tests | 30 minutes except free recall impaired for >120 minutes | Testing only undertaken for 2 hours. Training occurred. | (33) |
| Halothane/ $N_20/0_2$ Induced at high conc, then reduced to Fi 0.7% for 3.5 minutes | 11 | None | 4 psychomotor tests & simulated driving test. | >5 hours (psycho-motor testing); driving ability recovered by 7 hours. | Psychomotor testing only under-taken for 5 hours. Study included as a comparison with the newer agents. | (12) |
| Enflurane/ $N_20/0_2$ Induced at high conc, then reduced to Fi 1.5% for 3.5 minutes | 11 | None | ‹ | >5 hours (psycho-motor testing) although driving ability recovered by 4.5 hours. | ‹ | (12) |

Note: Recovery times are based on the time taken for the mean test result to return to baseline or control levels. ET=end-tidal. Fi=fraction inspired.

**Table 5**
Summary of recovery after combined anaesthetics

| Anaesthetic agents | Number subjects | Procedure undertaken | Tests Used | Psychomotor recovery time period or alternative measure | Comments | Reference |
|---|---|---|---|---|---|---|
| a) Propofol 2.5 mg/kg | 12 | None | 5 psychomotor tests & multiple sleep latency test. | 4 hours including sleep latency period. | Sleep latency period is more sensitive than psychomotor testing. Combinations including propofol but excluding midazolam decreased sleep latency the least. | (11) |
| b) Propofol 2 mg/kg + Fentanyl 2 mcg/kg | " | " | " | " | " | (11) |
| c) Propofol 2 mg/kg + Midazolam 2 mg/70 kg | " | " | " | 4 hours 6 hours until recovery of sleep latency period. Range: up to 8 hours. | " | (11) |
| d) Midazolam 0.07 mg/kg + Fentanyl 2 mcg/kg | " | " | " | " | " | (11) |
| Fentanyl 1 mcg/kg + Midazolam 0.05 mg/kg + 0.025 mg/kg PRN + Propofol 10-20 mg PRN | 35 | Colonoscopy | 9 psychomotor tests | 2 hours except free recall | Training undertaken Testing occurred for for only 2 hours. | (33) |
| a) Fentanyl 50 mcg/70 kg + Propofol 35 mg/70 kg | 12 | None | 4 psychomotor tests | 30 minutes until = BAL 100 mg/100 ml | To be more relevant to the Australian population a BAL | (34) |
| b) Fentanyl 50 mcg/70 kg + Midazolam 2 mg/70 kg | " | " | " | 60 minutes until = BAL 100 mg/100 ml | of 50 mg/100 ml would be more appropriate. | (34) |
| c) Fentanyl 50 mcg/70 kg + Midazolam 2 mg/70 kg + Propofol 35 mg/70 kg | " | " | " | 75 minutes until = BAL 100 mg/100 ml | Relatively low drug doses used. | (34) |
| Midazolam 0.1 mg/kg + Fentanyl 2 mcg/kg | 12 | None | 8 psychomotor tests | 3 hours Approaching baseline. | Training undertaken. | (35) |
| a) Propofol Induction 203.7 mg (mean) + 12/8 mg/kg/hr + Alfentanil 500 mcg + 250 mcg 15 minutely (operation time 15-79 minutes) | 15 | Knee Arthroscopy | 2 psychomotor tests | 1 hour | Practice effect evident with choice reaction time though not with the perceptive accuracy test. | (36) |
| b) Propofol Induction 212 mg (mean) + Fi Isoflurane 0.5-2% + Alfentanil 500 mcg + 250 mcg 15 minutely (operation time 18-73 minutes) | " | " | " | 1 hour | " | (36) |
| Propofol 10/8/6 mg/kg/hr +/- Alfentanil 10 mcg/kg Mean anaesthetic duration 13 minutes. | 46 | Oesophago-scopy & Bronchoscopy | 1 psychomotor test only - Critical flicker fusion threshold. | 1 hour | Minimal practice effect | (37) |

**Table 5**
Continued

| Anaesthetic agents | Number subjects | Procedure undertaken | Tests Used | Psychomotor recovery time period or alternative measure | Comments | Reference |
|---|---|---|---|---|---|---|
| a) Propofol 2.5 mg/kg + 10 mg PRN+ $N_2O/O_2$ 66:34 Mean anaesthetic duration 9.7 minutes | 30 | Outpatient gynaecological | 2 psychomotor tests | 1 hour | Compared their study subjects with controls hence practice effect not an issue. | (38) |
| b) Propofol 2.5 mg/kg $N_2O/O_2$ 66:34 + Fi Isoflurane 1% Mean anaesthetic duration 12.9 minutes | " | " | " | 1 hour | " | (38) |
| a) Propofol 2 mg/kg + 10 mg/kg/hr+ $N_2O/O_2$ 67:33 Mean anaesthetic duration 51.4 minutes | 20 | Knee Arthroscopy | 2 psychomotor tests | 1.5 hours | Did not take the practice effect which occurred into account. | (39) |
| b) Propofol 2 mg/kg+ $N_2O/O_2$ 67:37+ Fi Isoflurane 0.9% Mean anaesthetic duration 49.9 minutes | " | " | " | 1.5 hours | " | (39) |
| a) Propofol 2.5 mg/kg + Propofol 9/6 mg/kg/hr + $N_2O/O_2$ 67:33 Mean anaesthetic Duration 33.5 minutes | 32 | Dental surgery | 4 psychomotor tests | 3 hours (Return to baseline levels only except for free recall.) | Matched with a control group who achieved consistently better results than baseline at 3 hours (practice effect) except with free recall. | (8) |
| b) Propofol 2.5 mg/kg + Propofol 9/6 mg/kg/hr + $N_2O/O_2$ 67:33 + Alfentanil 10 mcg/kg/hr Mean anaesthetic duration 34.5 minutes | 25 | Gynaecological laparoscopy | " | 5 hours (Just approaching baseline) | 8% of subjects unable to even take the tests. Tissue stores must have been saturated with alfentanil to account for this prolonged sedation. | (8) |
| c) Propofol 2.5 mg/kg + Fi Isoflurane 1% + $N_2O/O_2$ 67:33 Mean anesthetic time 32.3 minutes | 32 | Dental surgery | " | 3 hours (Return to baseline levels only.) | See above comments for (a) | (8) |
| d) Propofol 2.5 mg/kg + Fi Isoflurane 1% + $N_2O/O_2$ 67:33 + Alfentanil 10 mcg/kg +10 mcg/kg/hr Mean anaesthetic duration 40.2 minutes | 25 | Gynaecological laparoscopy | " | 5 hours (Just approaching baseline) | 2% unable to take tests. | (8) |
| a) Propofol 230 mg (mean) + 10/6 mg/kg/hr + $N_2O/O_2$ 67:33 + Alfentanil 920 mcg (mean) Mean anaesthetic duration 33 minutes | 24 | Knee Arthroscopy | 2 psychomotor tests | >1 hour | This group received a higher dose of alfentanil compared with the group below. No testing occurred after 1 hour. Large inter-individual variation seen. | (40) |

**Table 5**
Continued

| Anaesthetic agents | Number subjects | Procedure undertaken | Tests Used | Psychomotor recovery time period or alternative measure | Comments | Reference |
|---|---|---|---|---|---|---|
| b) Propofol 237 mg (mean)+ Fi Isoflurane 0.5-2% + $N_20/0_2$ 67:33 + Alfentanil 740 mcg (mean) Mean anaesthetic duration 24 minutes | 26 | " | " | <1 hour | | (40) |
| Propofol 2.5 mg/kg + Desflurane 1.25 MAC + $N_20/0_2$ 60:40 Anaesthetic duration 1 hour | 20 | None | 9 psychomotor tests | 3 hours | Training undertaken | (32) |
| Propofol 2.5 mg/kg + Fentanyl 1mcg/kg + $N_20/0_2$ 50:50 + Desflurane 1 MAC for 30 minutes | 4 | None | Simulated driving | 3 hours At this time simulated driving was comparable with a BAL 50 mg/ 100ml. | Good study design. Compared functioning with that at the legal BAL for driving. Unfortunately only used small numbers. No training undertaken. | (41) |
| a) ET desflurane 4.7% Mean anaesthetic duration 146 minutes | 16 | Elective limb orthopaedic surgery | 2 psychomotor tests | 1 hour | No training undertaken Possible practice effect. | (42) |
| b) ET Isoflurane 0.8% Mean anaesthetic duration 149 minutes<br><br>Both groups were combined with: Midazolam 1-2 mg + Fentanyl 50 mcg+ Thiopentone 7 mg/kg + $N_20/0_2$ 60:40 | 9 | " | " | 1 hour | " | (42) |
| Propofol 2-3 mg/kg + Fi Isoflurane 0.5-2% + Alfentanil 980 mcg (mean) Mean anaesthetic duration 39.5 minutes | 15 | Knee Arthroscopy | 4 psychomotor tests | 1 hour | Practice effect seen with 3 of the 4 tests. Not seen with the perceptive accuracy test. | (43) |
| Sevoflurane ET Sevo 0.23% + Fentanyl 1.5 mcg/kg + $N_20/0_2$ 66:34 Mean anaesthetic duration 57 minutes | 13 | Non - neurological elective surgery | 2 psychomotor tests & visual evoked potentials | 1 hour except visual evoked potentials > 90 minutes | No further testing after 90 minutes. Practice effect evident but not accounted for. | (44) |

Note: Recovery times are based on the time taken for the mean test result to return to baseline or control levels unless otherwise specified.
MAC = minimal alveolar concentration
ET = end tidal
Fi = fraction inspired

Table 6 gives a summary of the longest recovery times for all agents. Even acknowledging the cautionary points in previous paragraphs the evidence in this review has shown remarkably rapid psychomotor recovery with the modern anaesthetic agents. The studies to date have shown that midazolam, fentanyl, and propofol anaesthetics appear to be associated with virtually total recovery within 10 hours. Some studies have shown the occasional psychomotor test, for example free recall, to remain impaired for up to 24 hours but whether this test is very discriminatory or even has a significant effect on driving skills is debatable. EEG changes have also been shown to persist for >8 hours but the significance of this for driving has not been determined. When compared with legal driving limits for alcohol, these anaesthetics compare very favourably, skills being the same as at a 50 mg/100 ml alcohol level within 3 hours. Nitrous oxide has been shown to have a very rapid objective recovery period when used as a single agent, but in one study was associated with a significant subjective feeling of being "unwell" for up to 8 hours. The implications of this for driving are difficult to ascertain. Is it like driving when we have a cold or flu? The anaesthetics conducted with volatile agents are not as clear cut. Studies involving desflurane show rapid recovery times of within 3 hours. However, evidence on sevoflurane is insufficient to make conclusions, and evidence on isoflurane is only available from combination anaesthetics. Most of the combination or balanced general anaesthetic techniques have shown a recovery within 8 hours.

It appears that if a patient needs to be as alert as possible postoperatively the best anaesthetic agents to use for sedation are propofol or $N_20$, with or without fentanyl (in doses <100 mcg). For general anaesthetics the best maintenance agents are propofol with

**Table 6**
Summary of longest recovery times of all agents (from 2 or more level II evidence trials)

| Anaesthetic agent | Longest psychomotor recovery time period or alternative measure | Number of studies |
|---|---|---|
| Midazolam (Maximum dose 0.15 mg/kg) | 10 hours | 6 (14, 15, 16, 17, 18, 19) |
| Diazepam (Maximum dose 0.3 mg/kg) | 7 hours | 4 (13, 15, 18, 20) |
| Propofol | 2 hours | 4 (19, 21, 22, 23) |
| Fentanyl (Maximum dose 2.5 mcg/kg) | 6 hours | 4 (13, 24, 25, 26) |
| $N_20/0_2$ (Maximum concentration 50:50) | 0.5 hours | 4 (27, 28, 29, 30) |
| Desflurane | 3 hours | 2 (31, 32) |
| Propofol/midazolam/fentanyl/ alfentanil | 8 hours | 6 (11, 33, 34, 35, 36, 37) |
| Propofol/$N_20$ maintenance (+ alfentanil infusion >5 hours (8)) | 3 hours | 4 (8, 38, 39, 40) |
| Desflurane/$N_20$ maintenance (±midazolam/fentanyl/propofol/ thiopentone) | 3 hours | 3 (32, 41, 42) |
| Isoflurane/$N_20$ maintenance (±midazolam/fentanyl/alfentanil boluses/propofol/thiopentone) (+ alfentanil infusion >5 hours (8)) | 3 hours | 7 (8, 36, 38, 39, 40, 42, 43) |

or without $N_20$, desflurane/$N_20$, or isoflurane/$N_20$. If analgesia is required after a procedure it would seem logical to use non-sedative drugs such as paracetamol, non-steroidal antiinflammatory agents or local anaesthetics.

**Conclusion**

After conducting this review one could conclude that after a minimum period of 12 hours following day case sedation or anaesthesia it is probably safe to drive a motor vehicle. The caveat being that this time may have to be extended for individuals in poorer health, taking other medications, or having larger total doses of agents than quoted here. In addition, as a patient who drives 12 hours after a procedure is almost certainly driving at night, night driving ability then becomes important. In practice it would be appropriate to advise patients that it would be safe to drive the next morning. As with all medical decisions requiring justification with scientific evidence there is room for further large controlled trials. A multi-centre study involving a real-life cross section of patients having actual day surgical procedures, and looking specifically at their functioning after 12 hours, would be ideal.

**Acknowledgements**

Special thanks to Dr John Paull and Dr Simon Morphett for proof reading this paper and to Dr Phil Ogden for the original idea.

**References**

1. Australian and New Zealand College of Anaesthetists. Endoscopy of the Airways. Review PS17; 1997.
2. National Health and Medical Research Council. A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines. Canberra; 1999.
3. Willette RE, Walsh JM. Drugs, Driving, and Traffic Safety. Geneva: World Health Organisation, 1983, Publication No. 78.
4. Ogg TW. An assessment of postoperative outpatient cases. *British Medical Journal* 1972; 4:573-766.
5. Correa R, Menezes RB, Wong J, Yogendran S, Jenkins K, Chung F. Compliance with postoperative instructions: a telephone survey of 750 day surgery patients. *Anaesthesia* 2001; 56:447-484.
6. Havard JDJ, in Medical Aspects of Fitness to Drive, (ed.) A Raffle, 3rd edn, p. 43. London, Medical Commission on Accident Prevention, 1976.
7. Mitchell MC. Alcohol-Induced Impairment of Central Nervous System Function: Behavioral Skills Involved in Driving. *Journal of Studies on Alcohol* 1985; Supp 10:109-116.
8. Marshall CA, Jones RM, Bajorek PK, Cashman JN. Recovery characteristics using isoflurane or propofol for maintenance of anaesthesia: a double-blind controlled trial. *Anaesthesia* 1992; 47:461-466.
9. Rechtschaffen A, Kales A (eds): A manual of Standardized Terminology Techniques and Scoring System for Sleep Stages of Human Subjects. Washington DC, National Institutes of Health 1968, p. 204 .
10. George CFP, Boudreau AC, Smiley A. Effects of nasal CPAP on simulated driving performance in patients with obstructive sleep apnoea. *Thorax* 1997; 52:648-653.
11. Lichtor JL, Alessi R, Lane BS. Sleep Tendency as a Measure of Recovery after Drugs Used for Ambulatory Surgery. *Anesthesiology* 2002; 96; 878-883.
12. Korttila K, Tammisto T, Ertama P, Pfaffli P, Blomgren E, Hakkinen S. Recovery, Psychomotor Skills, and Simulated Driving after Brief Inhalational Anaesthesia with Halothane or Enflurane Combined with Nitrous Oxide and Oxygen. *Anesthesiology* 1977; 46:20-27.
13. Ghoneim MM, Mewaldt SP, Thatcher JW. The Effect of Diazepam and Fentanyl on Mental, Psychomotor and Electroencephalographic Functions and Their Rate of Recovery. *Psychopharmacologia (Berl.)* 1975; 44:61-66.
14. Hindmarch I, Subhan Z. The Effects of Midazolam in Conjunction with Alcohol on Sleep, Psychomotor Performance and Car Driving Ability. *International Journal of Clinical Pharmacologic Res* 1983; 3(5):323-329.
15. Nuotto EJ, Korttila KT, Lichtor JL, Ostman PL, Rupani G. Sedation and Recovery of Psychomotor Function After Intravenous Administration of Various Doses of Midazolam and Diazepam. *Anesthesia and Analgesia 199*2; 74:265-271.

16. Gudgeon AC, Hindmarch I. Midazolam: Effects on Psychomotor Performance and Subjective Aspects of Sleep and Sedation in Normal Volunteers. *British Journal of Clinical Pharmacology* 1983; 16 Supp 1:121S-126S.
17. Lichtor JL, Zacny J, Korttila K et al. Alcohol After Midazolam Sedation: Does It Really Matter? *Anesthesia and Analgesia* 1991; 72:661-666.
18. Korttila K, Tarkkanen J. Comparison of Diazepam and Midazolam for Sedation during Local Anaesthesia for Bronchoscopy. *British Journal of Anaesthesia* 1985; 57:581-586.
19. Kestin IG, Harvey PB, Nixon C. Psychomotor Recovery after Three Methods of Sedation during Spinal Anaesthesia. *British Journal of Anaesthesia* 1990; 64:675-681.
20. Korttila K, Linnoila M. Skills Related to Driving After Intravenous Diazepam, Flunitrazepam or Droperidol. *British Journal of Anaesthesia* 1974; 46:961-969.
21. Grant SA, Murdoch J, Millar K, Kenny GNC. Blood propofol concentration and psychomotor effects on driving skills. *British Journal of Anaesthesia* 2000; 85(3):396-400.
22. Ozone M, Itoh H, Yamadera W, et al. Changes in subjective sleepiness, subjective fatigue and nocturnal sleep after anaesthesia with propofol. *Psychiatry and Clinical Neurosciences* 2000; 54:317-318.
23. Sanders LD, Clyburn PA, Rosen M, Robinson JO. Propofol in short gynaecological procedures. *Anaesthesia* 1991; 46:451-455.
24. Manner T, Kanto J, Salonen M. Simple devices in differentiating the effects of buprenorphine and fentanyl in healthy volunteers. *European Journal of Clinical Pharmacology* 1987; 31(6):673-676.
25. Stevenson GW, Pathria MN, Lamping DL, Buck L, Rosenbloom D. Driving Ability After Intravenous Fentanyl or Diazepam. A Controlled Double-Blind Study. *Investigative Radiology* 1986; 21(9):717-719.
26. Kay B, Venkataraman P. Recovery After Fentanyl and Alfentanil in Anaesthesia for Minor Surgery. *British Journal of Anaesthesia* 1983; 55:169S-171S.
27. Trojan J, Saunders BP, Woloshynowych M, Debinsky HS, Williams CB. Immediate Recovery of Psychomotor Function After Patient-Administered Nitrous Oxide/Oxygen Inhalation for Colonoscopy. *Endoscopy* 1997; 29:17-22.
28. Martin JP, Sexton BF, Saunders BP, Atkin WS. Inhaled patient-administered nitrous oxide/oxygen mixture does not impair driving ability when used as analgesia during screening flexible sigmoidoscopy. *Gastrointestinal Endoscopy* 2000; 51(6):701-703.
29. Korttila K, Ghoneim MM, Jacobs L, Mewaldt SP, Petersen RC. Time Course of Mental and Psychomotor Effects of 30 Per Cent Nitrous Oxide during Inhalation and Recovery. *Anesthesiology* 1981; 54:220-226.
30. Cheam EWS, Dob DP, Skelly AM, Lockwood GG. The effect of nitrous oxide on the performance of psychomotor tests. *Anaesthesia* 1995; 50:764-768.
31. Fletcher JE, Sebel PS, Murphy MR, Smith CA, Mick SA, Flister MP. Psychomotor Performance After Desflurane Anaesthesia: A Comparison With Isoflurane. *Anesthesia and Analgesia* 1991; 73:260-265.
32. Apfelbaum JL, Lichtor JL, Lane BS, Coalson DW, Korttila KT. Awakening, Clinical Recovery, and Psychomotor Effects After Desflurane and Propofol Anaesthesia. *Anesthesia and Analgesia* 1996; 83:721-725.
33. Theodorou T, Hales P, Gillespie P, Robertson B. Total Intravenous Versus Inhalational Anaesthesia for Colonoscopy: A Prospective Study of Clinical Recovery and Psychomotor Function. *Anaesthesia and Intensive Care* 2001; 29:124-136.
34. Thapar P, Zacny JP, Choi M, Apfelbaum JL. Objective and Subjective Impairment from Often-Used Sedative/ Analgesic Combinations in Ambulatory Surgery, Using Alcohol as a Benchmark. *Anesthesia and Analgesia* 1995; 80:1092-1098.
35. Lichtor JL, Zacny J, Apfelbaum JL et al. Alcohol after Sedation with I.V. Midazolam-Fentanyl: Effects on Psychomotor Functioning. *British Journal of Anaesthesia* 1991; 67:579-584.
36. Larson LE, Gupta A, Ledin T, Doolan M, Linder P, Lennmarken C. Psychomotor recovery following propofol or isoflurane anaesthesia for day-care surgery. *Acta Anaesthesiol Scand* 1992; 36(3):276-282.
37. Kestin IG, Chapman JM, Coates MB. Alfentanil used to supplement propofol infusions for oesophagoscopy and bronchoscopy. *Anaesthesia* 1989; 44:994-996.
38. Milligan KR, O'Toole DP, Howe JP, Cooper JC, Dundee JW. Recovery from Outpatient Anaesthesia: A Comparison of Incremental Propofol and Propofol-Isoflurane. *British Journal of Anaesthesia* 1987; 59:1111-1114.
39. Zuurmond WWA, Van Leeuwen L, Helmers JH. Recovery from propofol infusion as the main agent for outpatient arthroscopy. *Anaesthesia* 1987; 42:356-359.
40. Gupta A, Kullander M, Ekberg K, Lennmarken C. Assessment of recovery following day-case arthroscopy. A comparison between propofol and isoflurane-based anaesthesia. *Anaesthesia* 1995; 50:937-942.

41. Sinclair D, Chung F, Smiley A. Recovery and Simulated Driving after Outpatient Anaesthesia. *Canadian Journal of Anaesthesia Part II* 1999; 46:A28.
42. Tsai SK, Lee C, Kwan WF, Chen BJ. Recovery of Cognitive Functions after Anaesthesia with Desflurane of Isoflurane and Nitrous Oxide. *British Journal of Anaesthesia* 1992; 69:255-258.
43. Gupta A, Larson LE, Sjoberg F, Lindh ML, Lennmarken C. Thiopentone or propofol for induction of isoflurane-based anaesthesia for ambulatory surgery? *Acta Anaesthesiol Scand* 1992; 36:670-674.
44. Iohom G, Collins I, Murphy D, et al. Postoperative changes in visual evoked potentials and cognitive function tests following sevoflurane anaesthesia. *British Journal of Anaesthesia* 2001; 87(6):855-859.

# Appendix

**Description of Psychomotor Tests**

| Psychomotor Test | Description |
|---|---|
| Choice reaction time | Time required to react to one specific stimulus, with a distinct response, when presented with multiple stimuli. |
| Critical flicker fusion | Requires the subject to discriminate flickering in a set of "four light emitting threshold diodes in foveal fixation at one metre." |
| Free recall | Ability to recall 9 objects off a picture card 15 minutes later. |
| Maddox wing test | Measures the divergence of the eyes (extraocular muscle balance) and is representative of general muscle tone. |
| Peg board test | Subject puts tight-fitting pegs through holes as quickly as possible. |
| Perceptive accuracy test | Involves reacting to a 2 digit number flashed transiently up on a screen by pressing the same numbers on a keyboard. |
| Tracking task | Subject maintains a marker icon in contact with a target circle moving across the screen. |
| Tracometer | A steering task. |
| Visual Evoked Potentials | Measurement of the integrity of the visual neural pathway. |

# Do the Right Thing: A New Way to Manage Mistakes

MERRILYN WALTON

*Merrilyn Walton is an A/Professor of Ethical Practice, Faculty of Medicine, University of Sydney. Prior to this appointment Merrilyn was the founding Commissioner for the Health Care Complaints Commission in NSW. This paper is based on her presentation at the NSW Anaesthetic Continuing Education Meeting, November 2002.*

Do you tell your patient when you make a mistake in treatment? Even though honest communication is something patients could reasonably expect, many good doctors do not disclose mistakes. Fear of litigation, the medical culture and possible damage to reputations inhibit truth telling and promote secrecy. Open disclosure is the process of the clinician informing the patient about an adverse event, its cause and what steps will be taken to fix the problem. This article argues for open disclosure and describes the national "Open Disclosure" project.

**What is the evidence for patient harm?**

Harm to patients as a result of their health care is not new. Nearly 50 years ago the link between errors and harm to patients was shown in a study by Beecher and Todd who examined all deaths between 1948 and 1952 occurring on surgical services in 10 university hospitals.[1] They found that 7.6 per cent of the deaths resulted from gross errors in anaesthesia management. More recent studies in Australia, the United States of America, the United Kingdom and New Zealand confirm the harm to patients as a result of their health care. These studies using retrospective medical record review have been pivotal in identifying the types of errors and the extent of preventable injury to patients.[2-4]

Even though the extent of iatrogenic injury in the health system has long been recognised, the degree to which they are acknowledged and managed varies greatly across the system and across health professions. Ignorance of the extent of injury, and the fact that most errors do not cause harm, may explain the historical lack of concern. Outcome data still is not routinely published in the media and not routinely collected by health providers. Medical mistakes mainly occur in isolation (one patient at a time) and this too makes it harder for health professionals to appreciate the extent of errors in the system. Most iatrogenic injuries are due to preventable mistakes involving human beings.[5]

In a survey of 831 physicians published in the *New England Journal of Medicine* in December 2002, 35 per cent said that either they or members of their family had experienced medical errors in the course of their treatment. Many said the errors led to death, serious long term disability or severe pain. Three out of 10 doctors reported an error causing serious harm to patients outside their families in the last 12 months.[6]

Despite experience of error and the literature confirming the extent of the problem,

many clinicians are reluctant to engage openly with patients about medical mistakes. During the 15 years I was responsible for the management of complaints involving doctors in New South Wales, it became evident that most complaints involving bad outcomes for patients would not have been made had the treating doctors been honest and open about the causes of the injuries at the time they occurred. However, the medico-legal dominated health environment places patients and doctors in adversarial positions; patients seeking information and answers through complaint mechanisms and clinicians (often on legal advice) remaining mute. This leaves no room for engagement and reconciliation.

There are both benefits and potential harms associated with disclosure of medical mistakes to patients.[7] For patients, the greatest benefits relate to access to treatment to fix their medical problem, prevention of needless worry about the source of the problem, and provision of information to allow informed decisions and promote trust. The potential harms relate to the increased anxiety as a result of the disclosure, loss of confidence in the treating doctor and increased mistrust in the medical profession.

Clinicians who disclose may feel relief at telling the patient and a strengthened doctor-patient relationship. Candour may decrease the risk of litigation because patients appreciate having timely information. They may improve practice and learn from others' mistakes. Possible harms include stress caused by an angry patient, exposure to a medical negligence suit, increased insurance premiums, loss of referrals, and damage to reputation and career prospects.

The health system is now implementing systems for managing adverse events. These practices include incident monitoring, root cause analysis and medical record review. Some hospitals have guidelines designed to encourage health professions to capture and study errors. Error management, now centre stage for health departments and hospitals, will soon include standards on the process of disclosing adverse events to patients.

**The impact of medical errors on clinicians and patients**

Clinicians who want to "do the right thing" and tell their patients about adverse outcomes come up against significant barriers. The fear of litigation is easily recognised but the literature shows other issues such as the medical culture and the fear of being judged by peers and patients are equally significant. In a study designed to explore the emotional impact of the most memorable mistake made by family physicians, it was found that doctors experienced a range of emotions: self-doubt (96%) disappointment (93%) self blame (86%) shame (54%) and fear (50%).[8] Participants disclosed the lack of support from their peers and their own reluctance to unconditionally help a colleague with identical needs for support. Beliefs about perfection are also heavily influential on clinicians and how they manage mistakes.

*Litigation*

Doctors do not set out to hurt their patients and are distraught when patients suffer as a result of their treatment. In theory, doctors agree with a patient's right to know when a mistake has been made but, in practice, patients are not routinely told of mistakes. Litigation alone makes good doctors deceptive. Fear of being sued for medical negligence and the psychological stress and shame from a pending court case are major factors preventing open disclosure. This was not always the case. In the late nineteenth and early twentieth centuries, mistakes were commonly reported in the

journals as an educational tool.[9] By 1934, this had changed because of new scientific requirements for reporting, but also because of the threat of malpractice suits.

The medical indemnity providers, for economic reasons, have supported the adversarial position taken by clinicians. The nett effect of an adversarial position is that clinicians often think they are prohibited from talking about a mistake with anyone, except their lawyers. Without reporting of mistakes within hospitals and peer groups, there can be no investigations to determine contributing factors, a necessary step before improvements can be made. Equally, without system improvements and practice reviews the chances of mistakes recurring are high.

Many clinicians want to talk to patients about mistakes but feel constrained by legal advice. Hospital lawyers and medical defence organisations fear that full and frank discussions with patients about mistakes will expose doctors and hospitals to litigation because it will be seen as an admission of liability. In the consultation phase of the National Council on Quality and Safety Open Disclosure Project that commenced in 2002, the medical defence organisations did agree that there were no legal impediments to open disclosure. But, in practice, medical defence organisations send mixed messages to their members, leaving many clinicians in a state of uncertainty.

Of the thousands of mistakes made in the health system, few are due to negligence. Leape et al found that more than two-thirds of adverse events are preventable and that while 28% were due to negligence, 42% were caused by non negligent errors.[10] Yet many health professionals treat all mistakes as if they were "negligent". Poor understanding by health professionals of clinical risk is extensive. Many clinicians are experiencing unnecessary anxieties as a result of their misconceptions of the law of negligence and the legal system. Training and education of doctors in the past has not included clinical risk management.

Imagine you are a locum filling in for a colleague who works in a private hospital that you have not worked at previously. Your patient is a young woman with juvenile rheumatoid arthritis in both wrists. The agreed treatment is surgery on both wrists to be done separately, a different procedure for each wrist. The patient wanted the left wrist done first and arrangements were made, including advising you, the anaesthetist. However, unbeknown to you, the patient changed her mind about the order of the wrists the night before the planned operation and the surgeon notified theatres about the changed procedure. You arrive early and make the necessary preparations for the left wrist but you fail to confirm that this is the correct side. You proceed to perform a regional anaesthetic including sedating the patient prior to commencement of the block. When the surgeon arrives the error is detected.*

The causes of this simple slip, failing to confirm the site with the patient, or anyone else, are multifactorial. They involve environmental, organisational, patient and clinician factors. What do you tell the patient? Avoiding her, making excuses, or accepting total responsibility may not be appropriate, nor will any of these approaches lead to an analysis of why the error happened in the first place. The best approach is to advise the patient as soon as possible about the error and undertake an evaluation of the factors that led to it. When you have ascertained the main causes of error, you are then in position to provide detailed information to the patient as well as put in place some processes to help avoid repeating the same mistake.

*This circumstances of this hypothetical case were provided by Dr Leone Watterson, who developed it from a number of actual cases.

There is strong evidence that openness about mistakes is an antidote to litigation.[11] When doctors tell the truth about an adverse event to the patient they are less, not more, likely to be sued. A 1994 study published in Lancet found that the deciding factors for initiating legal action were the nature of the original injury, insensitive handling and poor communication by the doctors after the original event.[12] Another study found that patients who suffered from moderate to severe mistakes were significantly more likely to consider litigation if the doctor did not disclose the error.[13]

*The ethical position*

There is no mention of the necessity to disclose to patients after an adverse event in the Australian Medical Association Code of Ethics. The American Medical Association Code of Ethics expressly obliges doctors to disclose mistakes to patients, even though they may be concerned about legal liability which might result following truthful disclosure. In the UK, the General Medical Council advises doctors to act immediately to put matters right, if that is possible. The patient should have a full explanation by the doctor of what happened and the likely long- and short-term effects. They suggest an apology, if appropriate. The Royal Australasian College of Surgeons requires honesty with patients and stresses that the relationship should at all times be characterised by trust, which allows an honest exchange of information and facilitation of patient autonomy. But it falls short on explicit instructions about acknowledging mistakes.

Honesty and integrity remain hallmarks of the profession. They are the touchstone for the community's trust. If non-disclosing continues, the community's trust in the profession erodes because two primary ethical principles underpinning ethical medicine (respect for patient autonomy and putting the interests of the patient first) are breached.

How do we make ethics the driver for open disclosure? Over the last 25 years, medicolegal uncertainties have usurped ethics as the main instrument for ethical and professional practice. This incursion is responsible for driving doctors and patients apart by instilling fear in doctors and mistrust in patients. One way to promote ethical practice is by providing clear explanations as to what doctors are expected to do and how they can do it. In 2002, The Quebec College of Physicians required doctors to tell their patients about all "incidents", "accidents" or "complications" occurring during treatment which could have a substantial impact on their health.[14] The code does not use the term "medical error". A failure to advise patients accordingly will be treated as a disciplinary matter and could result in a fine or suspension. The College avoided the term "error" because of its association with negligence.

If your patient suffers an adverse drug reaction from an overdose of an anaesthetic drug you prescribe, what do you tell him? Do you tell the truth or do you tell your patient that he had a reaction to the drug? How would you document this in the medical record? If you fail to document the truth in the record, this compounds the lie to the patient. There is no justification for writing a false note in the medical record, because the false information may deprive your patient of a drug that may be beneficial to his future health. Confessing the error promptly and acting upon and attending to your patient with care and open communication are your best strategies from many perspectives. Should litigation follow, then you have done everything possible for your patient; fulfilling your fiduciary duties of good faith, trust, confidence and candour.

Your ethical obligation to put your patient's interests first is met. Sharing the mistake with your patient may promote trust and sustain your relationship.

Clinicians often use methods other than open disclosure to manage mistakes. Patients are told their problems are part of the progress of their or disease/condition, instead of the after effects of a mistake. The doctor may deny the error or blame others for the problem. Others provide incomplete information about the patient's problem by saying complications have arisen. Some may resort to falsehood, telling the patient that a complication occurred and that they are lucky to survive. Patients are often so grateful that they may not question the original nature of the problem. I have also seen doctors avoid patients who have suffered an adverse event, or delegate the responsibility for informing the patient or their family to a junior member of staff who may not have all the facts. Because of their junior position, they may be reluctant to refuse the task. Finally, some doctors may use rationalisation to manage the mistake, "the patient was going to die in any event", so the error in treatment is not seen as a significant factor.

*The uncertainty of medicine*

Nearly fifty years ago, sociologist Renee Fox observed that medical training prepares clinicians for uncertainty by teaching them to accept the limitations and gaps in medical knowledge and to appreciate that available knowledge can not always be applied perfectly. Today, even with the development of specialisation, there are areas of medicine beyond reach despite the best evidence and expert advice.

Uncertainty is part and parcel of clinical practice. McIntyre and Karl Popper pointed out in the *BMJ* that professional ethics for most of the twentieth century were mistakenly based on the view that scientific knowledge was certain knowledge, that knowledge grows normally by accumulation and can be acquired and stored in a person's mind.[15] This perception, they believed, permitted a culture of "authority" to rise and dominate. Many doctors still subscribe to this view and argue that their authority is undermined when patients question their knowledge or seek additional information.[16] The corollary of this premise is that mistakes by clinicians are caused by insufficient knowledge or skill. Some still maintain this position, even though they recognise that it is impossible for doctors to know everything.

*The medical culture*

Other factors, such as notions of perfection and infallibility within medicine, militate against open disclosure. These make individuals reluctant to own their mistakes for fear of censure. The "blame" culture in medicine is renowned for discouraging candour about mistakes. Whether doctors disclose medical mistakes involves a number of considerations and conflicts, such as personal morality, professional obligation to prevent recurrence of the error, concern that disclosure will jeopardize the professional relationship with another colleague, the patient's right to know the truth, concern that information will cause more suffering for the patient, and damage to patient's confidence in the physician.[17]

Identifying an adverse event as a mistake rather than a complication will also depend on other factors such as the hospital culture. The willingness of senior clinicians to discuss their own errors with colleagues will significantly impact on how clinicians generally identify and manage errors. Many discussions about error take place in a blame environment with punitive overtones such as Morbidity and Mortality Meetings.

A fear of being blamed and worries about litigation are significant factors in the under-reporting of mistakes by junior doctors in hospitals. The first three years in a young doctor's professional life are crucial for their development and consolidation of professional values and attitudes. An unsupportive environment can lead to them concealing mistakes rather than discussing them openly. This sends a message to junior doctors, "don't be open about your mistakes". Consequently, many fear their careers will be ruined if they self-report or that their supervisors will lose confidence in them. Twenty years ago Bosk, a sociologist, categorised errors for young surgical trainees into technical, judgemental and normative errors.[18] He defined errors as technical errors when they were speedily acknowledged, reported and treated. If these conditions were met the technical error was forgiven. Popper and McIntyre dispute that doctors were encouraged to be open about their mistakes even 20 years ago.[15]

The complexity of the health care environment today complicates training generally and in particular error management. The proliferation of drugs and technology, diagnostic tests, and invasive procedures collectively increase the opportunities for errors. With sicker patients and shorter hospital stays, there is also decreased opportunity for detailed discussions about patients and sharing of information among clinicians. Fear of litigation, whether perceived or real, has had a major impact on mistake management in hospitals. A medical negligence court hearing can seriously damage a doctor's career even if the doctor was not negligent or the mistake minor.

*Impact on patients*

Ten years ago, Charles Vincent and others surveyed 154 surgical patients who were injured by their treatment.[19] They found that these patients were more distressed than people who had suffered serious accidents or bereavement. They suffered more pain (>1 year later) than if they had untreated postoperative pain, their psychosocial adjustment was considerably worse than patients with serious illnesses and the poor explanations they received were associated with disturbing memories and poorer adjustment. We have under-appreciated the impact of medical errors on patients. The little research available indicates that patients want to be told about errors, even minor ones. If the error was severe, patients had higher expectations for the provision of detailed information and explanation.

**Does open disclosure work?**

A literature review I undertook for the National Open Disclosure Project (2001) uncovered many studies supporting the benefits to clinicians and patients of a frank and honest exchange of information when errors are made. But despite this, many still doubt its value even though it has been successfully introduced by various organisations. In 1995, the US Department of Veteran Affairs adopted a risk management framework that depended on full disclosure and compensation. The guidelines on disclosure[20] require that:

> "The medical center will inform the patient and/or the family as appropriate of the event, assure them that medical measures have been implemented and that additional steps are being taken to minimize disability, death, inconvenience or financial loss to the patient and/or the family. ...District Counsel will advise the medical Center Director about informing the patient and/or family of their right to file an Application for Compensation and pension...or to file an administrative tort claim ..."

The policy includes the following components: early injury review, steadfast maintenance of the relationship between the hospital and the patient, full disclosure to patients that have been injured because of an accident or medical negligence and fair compensation for injuries. While there is no evidence that the open disclosure policy saves money, there is no escalation of costs. It also improves relationships with patients, improves staff morale and helps doctors maintain honest relationships with patients.

**What is the Open Disclosure project?**

As part of a broader agenda to improve the health system and make it safer, the Australian Council for Safety and Quality in Health Care in 2001 sponsored the development of national standards for open disclosure of adverse events to patients and their families. The aim of the Open Disclosure Project is to improve the practice of open disclosure by providing guidance on how to communicate an adverse event to a patient and how to ensure that mistakes are examined with a view to prevention.

The standards are designed to be flexible, to allow local adaptation in both the public and private sectors. Many clinicians already practice open disclosure but often feel unsupported by hospital management or insurers. The guidelines will also assist organisations to better support both patients and clinicians.

Arguments about whether a bad outcome is a result of a complication or an error are avoided, because the guidelines cover all incidents deemed to have caused harm to the patient during the course of their care and treatment. An adverse event is defined as an incident in which an unintended harm resulted to a person receiving health care.

Communication lies at the heart of the standards and is reflected in all the main themes underpinning them. The principles upon which the standards have developed are:

1. Openness and timeliness of communication
2. Acknowledgment
3. Expression of regret
4. Recognition of the reasonable expectations of patients and their nominees
5. Staff support
6. Integrated risk management and systems improvement
7. Good governance
8. Confidentiality

The guidelines have been developed through a process of literature and legal reviews, consultation with all the stakeholders, dissemination of a discussion document, examination of submissions, education and implementation of guidelines, and dissemination of open disclosure packages. Underpinning the guidelines is the recognition that we need to move away from blaming individuals when an adverse event occurs and move towards organisation and clinician support for disclosing medical mistakes. Only when adverse events are reported and analysed can improvements be made.

Karl Popper and Neil McIntyre wrote in the British Medical Journal a decade ago that doctors are expected to learn from their mistakes and that this requires a willingness to admit errors and discuss them.[15] They lamented that, unfortunately, medical students and doctors see little evidence of this. Gorowitz and MacIntyre think that a misunderstanding of fallibility in medicine is a major problem.[21] They wrote:

"No species of fallibility is more important or less understood than fallibility in medicine. The physician's propensity for damaging error is widely denied perhaps because it is so intensely feared. ... Physicians and surgeons often flinch from even identifying error in clinical practice, let alone recording it, presumably because themselves hold that error ... arises from their or their colleagues' ignorance or ineptitude."

While guidelines by themselves will not change practice or a clinician's attitude towards error, they do provide a reference point for those clinicians wishing to be more constructive about mistakes and more inclusive of patients.

Communication is the key to successful open disclosure. Clinicians wishing to practise open disclosure will be able to follow the guide which sets out how, when and where patients, clinicians and the organisation are to interact. Included in the standard is information about documentation, grading the event, investigation steps, and reporting for improvement.

**Legal issues**

Clarifying the legal position for open disclosure has been a major preoccupation with stakeholders. Although the courts accept that an explanation of the facts is different from an admission of liability in negligence, the perception that honesty with patients will expose clinicians and hospitals to litigation is real and is one of the major hurdles before open disclosure is embraced by clinicians. A legal review for the Open Disclosure project concludes that better communication will decrease legal risk, but adds a caveat that open disclosure poorly handled may inflame emotions and have the opposite effect. Staff training, support and guidance are essential for effective communications with patients suffering an adverse event.

After adverse events, many clinicians and hospital managers rely on legal advice to determine if patients are informed. The quality of that advice varies significantly. How willing doctors will be to communicate with patients will often depend on how knowledgeable and aware the lawyers are about open disclosure. Many lawyers are unfamiliar with medical ethics and the research on adverse events. Because open disclosure requires both communication with patients and an exploration of the event, there must be consideration on a case by case basis of privacy, confidentiality, and privilege (legal and qualified) issues.

Anaesthetists reviewing adverse events via the Deaths under Anaesthesia Committee use qualified privilege to protect their discussions and deliberations. Qualified privilege prevents anyone from discussing and divulging communications to unauthorised people. But open disclosure is about the opposite, providing detailed explanations about what happened and why. Some stakeholders want to use qualified privilege for the investigation process, even though most hospital investigations are currently done without legal protection.

**Summary**

The core value in open disclosure is communication. To deny or limit the information you provide to your patients may not meet their expectations and, in circumstances where trust may have already been damaged, your patients may think you or the hospital are covering up and have something to hide. This is a major reason for suing doctors. Many clinicians already practice open disclosure and know the benefits. But for a significant majority open disclosure will be new.

## References


1. Beecher and Todd. A study of the deaths associated with anesthesia and surgery based on a study of 599,548 anesthesias in ten institutions. Reported in: Kohn LT Corrigan JM Donaldson MS (eds) To Err is Human: Building a safer health system. 2000 National Academy Press Washington DC.
2. Brennan TA, Leape LL, Laird N et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study. *New England Journal of Medicine* 1991; 324:370-376.
3. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L.,Hamilton JD. The Quality in Australian Health Care Study. *Medical Journal of Australia* 1995; 163:458-471.
4. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal* 2001; 322:517-519.
5. Dubois RW, Brook RH. Preventable deaths: who, how often, and why? *Annals of Internal Medicine* 1998; 109:582-589.
6. Blendon RJ, DesRoches CM, Brodie M et al. Views of practicing physicians and the public on medical errors. *New England Journal of Medicine* 2002; 347:1933-1940.
7. Wu AW, Cavanaugh T, McPhee S, Lo B, Micco G. To tell the truth: ethical and practical issues in disclosing medical mistakes to patients. *Journal of General Internal Medicine* 1997; 12:770-775.
8. Newman MC. The emotional impact of mistakes on family physicians. *Archives of Family Medicine* 1996; 5:71-75.
9. Pincus RL. Mistakes as a social construct: an historical approach. *Kennedy Institute of Ethics Journal* 2001; 11:117-133.
10. Leape LL, Lawthers AG, Brennan TA, Johnson WG. Preventing medical injury. *Quality Review Bulletin* 1993; 8:145.
11. Walton M. Open Disclosure to patients or families after an adverse event: A literature review. 2001 http://www.nsh.nsw.gov.au/teachresearch/cpiu/CPIUwebdocs/WebversionLiteratureReview.pdf.
12. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. *Lancet* 1994; 343:1609-1613.
13. Witman AB, Park DM, Hardin SB. How do patients want physicians to handle mistakes? A survey of internal medicine patients in an academic setting. *Archives of Internal Medicine* 1996; 156:2565-2569.
14. Spurgeon D. Quebec doctors must tell patients about medical accidents. *British Medical Journal* 2002; 325:1192.
15. McIntyre N Popper K. The critical attitude in medicine: the need for a new ethics. *British Medical Journal* 1983; 287:24-31.
16. Walton M. Patient autonomy: Does doctor know best? *Medicine Today* 2002; 3:107-111.
17. Sweet MP, Bernat JL. A study of the ethical duty of physicians to disclose errors. *The Journal of Clinical Ethics* 1997; 8940;341-348.
18. Bosk CL. Forgive and Remember. University of Chicago Press 1979, p. 38.
19. Vincent C et al. Patients' experience of surgical accidents: 154 surgical patients injured by their treatment. *Qual Health Care* 1993; 2:77-82.
20. Kramann SS, Hamm JD. Risk Management: Extreme honesty may be the best policy. *Annals of Internal Medicine* 1999; 131:963-967.
21. Gorowitz S, MacIntyre N. Toward a theory of medical fallibility. *Journal of Medical Philosophy* 1976; 1:51-71.

# Root Cause Analysis: The NSW Health Incident Management System

SARAH MICHAEL, RN, GradDipQHCM

PAUL DOUGLAS, MB, BS, DRACOG, MHA, FRACMA

*With a background in intensive care, Sarah is a Principal Analyst in the Quality and Clinical Policy Branch in NSW Health. Paul is Deputy Director of Clinical Governance for Hunter Health (NSW), and the Director of the Patient Safety Programme for the Institute of Clinical Excellence in New South Wales. This paper was presented at the NSW Anaesthetic Continuing Education Meeting on decision-making and risk management, in November 2002.*

Since the Harvard Medical Practice Study in 1991[1] and the 1995 Quality in Australia Health Care Study,[2] the extent of the problem of error in clinical care has been highlighted. Health services and governments have sought to put in place constructive solutions to address this. At the national level, the Australian Council for Safety and Quality in Health Care was established in January 2001. At the state level, NSW Health has developed two key documents, the Framework for Managing the Quality of Health Services in 1999 and, more recently, the Clinician's Toolkit for Improving Patient Care. In 2001, NSW Health also established the Institute for Clinical Excellence (ICE). Together with the Quality and Clinical Policy Branch, ICE has worked collaboratively to provide practical solutions for health services and clinicians on how to respond to the demands for help in addressing patient safety matters.

The ICE/NSW Health Patient Safety Program, based to a large extent on the Veteran Affairs (VHA), National Centre for Patient Safety (NCPS) program in the USA, is now introducing a state wide program of "Root Cause Analysis". The VHA was able to roll out a similar program to more than 180 facilities within 10 months, with a resultant 30-fold increase in incident reports and a 900-fold increase in reports of near misses or less serious adverse events. The NSW program is designed to support improved practice in both the clinical and non-clinical environment, where staff is able to undertake systematic root cause analysis to find the real cause(s) of problems.

This paper outlines the components of the NSW education and training program and highlights how incidents can be prioritised and then investigated by appropriately selected teams. Through systematic analysis and utilisation of practical tools, causation statements can be developed which assist in bringing about changes to prevent recurrence of similar events.

## INTRODUCTION

Patient safety and continuous quality improvement have been firmly on the health policy agenda for over a decade. The Harvard Medical Practice Study,[1] the Quality in

Australian Health Care Study,[2] the Institute of Medicine Report[3] and, more recently, the Douglas Report on events in a Perth hospital[4] outlined high levels of adverse events in health care institutions. The Australian Council on Safety and Quality in Health Care, the NSW Quality and Clinical Policy Branch and the NSW Institute for Clinical Excellence believe the high incidence of preventable adverse events in health care is a major issue for the population. To address these concerns, organisations need to focus on approaches that:

- Have a positive impact in improving patient and staff safety;
- Aid understanding of the causes underlying adverse events;
- Facilitate changes in the systems and processes to reduce the probability of such events in the future;
- Increase the knowledge about events, their causes, and strategies for prevention; and,
- Improve the safety of health care for the consumer and maintain the confidence of the public in the care and services provided.

To date it has been the norm for incidents arising in healthcare to be blamed on an individual. However, it is now clearly recognised that it is a combination of specific circumstances and the work environment that commonly combine to result in unwanted outcomes. It is acknowledged that people have not come to work to do a bad job or make a mistake. Today, it is known that the "root causes" of problems leading to incidents are usually found in the design of the system that permitted the event in the first place, and are rarely attributable to an individual.

Organisations with the right processes for the investigation and analysis of adverse incidents can create an environment that focuses on systems and not on individuals. To minimise error, we must have a healthcare system that makes it easy to do the right thing and difficult to do the wrong thing.[3]

**CULTURE CHANGE**

To be successful, a change in culture across the health system is imperative for the implementation of this initiative. An approach that emphasises prevention, not punishment is required. Clinicians and managers need to be provided with the skills to identify and understand the deeper underlying factors in adverse events and translate them into corrective actions. This cultural change requires the following criteria:

- A commitment from the leaders of the organisation to quality improvement;
- The needs of patients/consumers, including open disclosure, be considered;
- A "quality culture" that emphasises empowerment, flexibility and multidisciplinary teamwork must be developed;
- Every clinician must take responsibility for patient safety;
- System improvements need to be the focus of corrective actions;
- Reliable, valid, timely and objective information necessary for decision making needs to be available to all who require it;
- Effective feedback processes must be in place;
- Monitoring and evaluation of performance must happen on a continuous basis; and,
- Management must commit to provide training and education to all staff, particularly on the use of "quality" tools.

Following a comprehensive literature review and consultation process, NSW Health has developed a model for a safety improvement program based to a large extent on that of the VHA in the United States. The VHA program was unique for health care

because of its focus on prevention, not punishment. The integration of human factor analysis and quality improvement methodologies was integral to their identification and elimination of system vulnerabilities.

"Root cause analysis" is the tool that is used by VHA and NSW Health to identify prevention strategies. As well, it defines processes to develop solutions, to test and implement them and then measure outcomes in order to improve patient safety. RCA has been a major factor in building a culture of safety within the VHA and in moving from a "blame" to a "just" culture. The NSW Health model, whilst based substantially on the VHA methodology, is also aligned to the Standards Australia Guideline for Managing Risk in the Healthcare Sector.[5] It relies strongly on the following principles:

**"Oversight" — Governance of the Process**
- Incorporates individual and team accountabilities;
- Defines committee structures and reporting mechanisms; and,
- Formulates and standardises policy.

**"Doing and learning" — The quality improvement and risk management process —**
- Prevention — such as credentialling, clinical guidelines and review of current and new procedures.
- Management of incidents — ensuring that particular types of incidents (including near misses and close calls) are identified, reviewed and analysed at an appropriate level and that recommendations are implemented.
- Audit and review of performance — including peer review, clinical indicators and variance analysis.

**"Knowing how" — Training and education —**
- Identification of training and education requirements.
- Communication of sentinel alerts, system learnings, the organisational safety process and skills training.

While the primary focus is on improving patient care, encouraging clinician participation and improving the work environment, RCA assists in the development of systems that:
- Identify and report incidents that occur, in a manner that encourages self-learning from the analysis;
- Lead to the investigation of serious adverse events and critical incidents in order to promote the redesign of systems as the main method for improving safety;
- Ensure action upon recommendations from these investigations;
- Ultimately improve patient safety and the individual's health care experience;
- Support a culture where every clinician takes responsibility for patient safety and where reporting of events and problems are rewarded, not punished.

**ROOT CAUSE ANALYSIS — THE METHOD**

RCA reviews events to find the most basic reason(s) that can be readily identified and that are in management's control to fix. RCA has the following characteristics:[6]
- It focuses primarily on systems and processes, not individual performance;
- It repeatedly digs deeper by asking "why";
- It identifies changes that could be made in systems and processes -either through redesign or the development of new systems or processes;

- Its focus is non punitive;
- Its focus is on how to improve systems in order to prevent the occurrence of sentinel events;
- It digs into existing systems to find new ways to do things.

Rare situations arise where patient harm has resulted and it is found that individuals have acted with intention to harm, under the influence of drugs or in the full knowledge that they are not adhering to policy. This should not be handled through a quality improvement process, but managed via normal human resource investigation procedures.

The following process has been designed to support improved practice in both the clinical and non-clinical work environments and to help clinicians and managers to perform systematic root cause analysis to find the real causes of problems. It is recognised that RCA is only one of many processes that work together to improve patient care and that it is used principally when things have gone wrong! The steps involved in the RCA method include:

1. Reporting;
2. Prioritisation;
3. Investigation — including team selection, flow charting;
4. Cause and effect diagramming and determination of the root causes;
5. Developing causation statements, listing contributing factors, actions and recommendations; and
6. Implementation of recommendations and evaluation of effectiveness

**1. Reporting**

Each Area Health Service in NSW will have its own mechanism for incident reporting. It is anticipated that, by late 2003, a state wide information system will standardise this mechanism.

**2. Prioritising Incidents**

The Severity Assessment Code (SAC) below is a simple method that allows Area Health Services to quantify the actual and potential risk associated with an incident. By using the SAC score, all incidents/events are rated from 1 to 4, with 1 being the most severe. A rating of 1 will always require an investigation and notification to Area Executive and the Department of Health. A rating of 2 will require notification to the Area Executive and local assessment as to the level of investigation that may be required. Incidents that are rated 3 or 4 will be managed locally (Figures 1 and 2).

**3. Investigation**

There are two fundamental challenges in the investigation process:

- To understand how and why an event occurred; and,
- To prevent the same or similar event from occurring in the future.

An identified problem is often the result of multiple causes, at different levels. This means that some causes affect other causes, which in turn create the visible problem. The process of RCA is somewhat analogous to clinical investigations, where symptoms and signs are further explored to identify underlying factors contributing to or causing the disease state.

Causes can be classified as one of the following:

| Serious | Major | Moderate | Minor | Minimum |
|---|---|---|---|---|
| Patients with death unrelated to the natural course of the illness and differing from the immediate expected outcome of the patient management or any of the following:<br>Sentinel events reportable to Australian Council for Safety and Quality in Health Care<br>• Procedures involving the wrong patient or body part<br>• Suicide<br>• Retained instruments<br>• Unintended material requiring surgical removal<br>• Intravascular gas embolism resulting in death or neurological damage<br>• Haemolytic blood transfusion<br>• Medication error leading to death<br>• Maternal death or serious morbidity associated with labour or delivery<br>• Infant abduction or discharge to wrong family<br>Requires notification under existing legislative reporting requirements | Patients with Major permanent loss of function (sensory, motor, physiologic or intellectual) unrelated to the natural course of the illness and differing from the expected outcome of patient management or any of the following:<br>Disfigurement as a result of the incident<br>Absconded involuntary mental health patient<br>Patient or staff assault requiring external involvement | Patients with permanent lessening of bodily functioning (sensory, motor, physiologic, or intellectual) unrelated to the natural course of the illness and differing from the expected outcome of patient management or any of the following:<br>• Increased length of stay as a result of the incident<br>• Surgical intervention required as a result of the incident | Patients requiring increased level of care including:<br>• Review and evaluation<br>• Additional investigations<br>• Referral to another clinician | Patients with no injury or increased level of care or length of stay |
| **Staff:** Death of staff member related to work incident or suicide, or hospitalisation of 3 or more staff | **Staff:** Permanent injury to staff member, hospitalisation of 2 staff, or 3 or more staff experiencing lost time or restricted duty or illness | **Staff:** Medical expenses, lost time or restricted duties or injury / illness for 2 or more staff | **Staff:** First aid treatment only with no lost time or restricted duties. | **Staff:** No injury or review required |
| **Visitors:** Death of visitor or hospitalisation of 3 or more visitors | **Visitors:** Hospitalisation of up to 2 visitors | **Visitors:** Medical expenses incurred or treatment for up to 2 visitors but not requiring hospitalisation | **Visitors:** Evaluation and treatment with no expenses | **Visitors:** No treatment required or refused treatment |
| **Services:** Complete loss of service or output | **Services:** Major loss of agency / service to users, including cancellation of booked surgery, more than twice | **Services:** Disruption to users due to agency problems | **Services:** Reduced efficiency or disruption to agency working | **Services:** No loss of service |
| **Financial:** Critical financial loss >$1,000,000 | **Financial:** Major financial loss $100,000 - $1,000,000 | **Financial:** Moderate financial loss $10,000 - $100,000 | **Financial:** Minor financial loss <$10,000 | **Financial:** No financial loss |
| **Environmental:** Toxic release off-site with detrimental effect. Fire requiring evacuation | **Environmental:** Off-site release with no detrimental effects or fire that grows larger than an incipient stage | **Environmental:** Off-site release contained with outside assistance or fire incipient stage or less | **Environmental:** Off-site release contained without outside assistance | **Environmental:** Nuisance releases |
| **IF ANY OF THE ABOVE INCIDENTS ARE LIKELY TO EVOKE EXTERNAL INTEREST, CONSIDERATION MUST BE GIVEN TO REAPPRAISING THE SAC RATING FOR THAT INCIDENT.** | | | | |

**Figure 1.** Severity Assessment Code: all incidents are analysed against actual and potential outcomes.

| CONSEQUENCE / LIKELIHOOD | Extreme | Major | Moderate | Minor | Insignificant |
|---|---|---|---|---|---|
| Frequent (almost certain) | 1 | 1 | 2 | 3 | 3 |
| Probable (likely) | 1 | 1 | 2 | 3 | 3 |
| Occasional (possible) | 1 | 2 | 2 | 3 | 4 |
| Uncommon (unlikely) | 1 | 2 | 3 | 4 | 4 |
| Remote (rare) | 2 | 3 | 3 | 4 | 4 |

**Figure 2.** Severity Assessment Code, scoring.

- Symptoms — these are not regarded as actual causes but rather flags of existing problems or the final outcome or event that occurred.
- First level causes — causes that directly lead to a problem.
- Lower level causes — causes that lead to the first level causes. While they do not directly cause the problem, lower level causes form links in the cause and effect relationships.

The lowest-level cause of a problem is called the root cause. It is "the evil at the bottom" that sets in motion the cause and effect chain that creates the problem(s).[7]

The NSW Health model has broken the investigation, cause and effect diagramming and actions into three separate meetings in its training program. This allows participants to more clearly focus on their objectives and complete the investigation in a timely manner.

***The sequence of events: Meeting 1***

The first key step taken by the root cause analysis team is to define the incident and determine the sequence of events that led to it. This involves the team mapping out the flow of what happened and when it happened, based on the initial understanding of the sequence of events and the known facts. An initial flow diagram is drawn up. This is an outline of the story of the key events, progressing chronologically from the first known fact, through the actual event being reviewed, and concluding with the final known event. Figure 3 is an example of the initial diagram that describes what has lead up to a disaster in an obstetric patient. In this story, the patient has undergone a failed trial of labour. A tired anaesthetist, who rarely does obstetric anaesthesia, reluctantly gives a general anaesthetic for the patient, who suffers anaphylaxis and a subsequent cardiac arrest.

Following the initial flow diagram, it is necessary to address each of these key events within the chain, asking "how and why" each condition existed, until there are either no more questions or no more answers. If the answer results in blaming an individual or group of individuals, another "why" question needs answering.

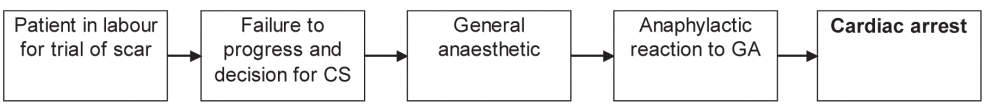


**Figure 3.** Initial flow diagram of events.

Based on the NCPS Triage Cards, NSW Health and ICE have developed a "Checklist Flipchart" to help identify system and process issues. At this stage in the process, the team familiarises themselves with the checklist flipchart questions that will help identify those who may need to be interviewed and assist in outlining other background information that needs to be collected. These may include records, policies and procedures, and equipment contributing to the event. The flipchart includes questions like:

- Were there issues related to patient assessment in this event?
- Were issues related to staff training or staff competency a factor in this event?
- Was equipment (or its use or lack thereof) involved in this event in any way?
- Was a lack or misinterpretation of information a factor in this event?
- Was communication a factor in this event?
- Were appropriate Policies/Procedures or guidelines, or lack thereof, a factor in this event?
- Was the failure of a safety mechanism or barrier designed to protect the patient, staff, equipment, or environment a factor in this event?
- Were specific patient issues a factor in this event.

Eventually, an expanded flow chart such as that shown in Figure 4 will be developed.

Following this step, unanswered questions may be resolved in interviews with those involved and identify information that still needs to be collected.

The team should not jump to conclusions, thinking they know the cause of an adverse event without doing any investigation. The natural tendency is to think in terms of a straight line from an adverse event to the preceding action and is influenced by "hindsight bias". In reality, multiple decision points are encountered and must be

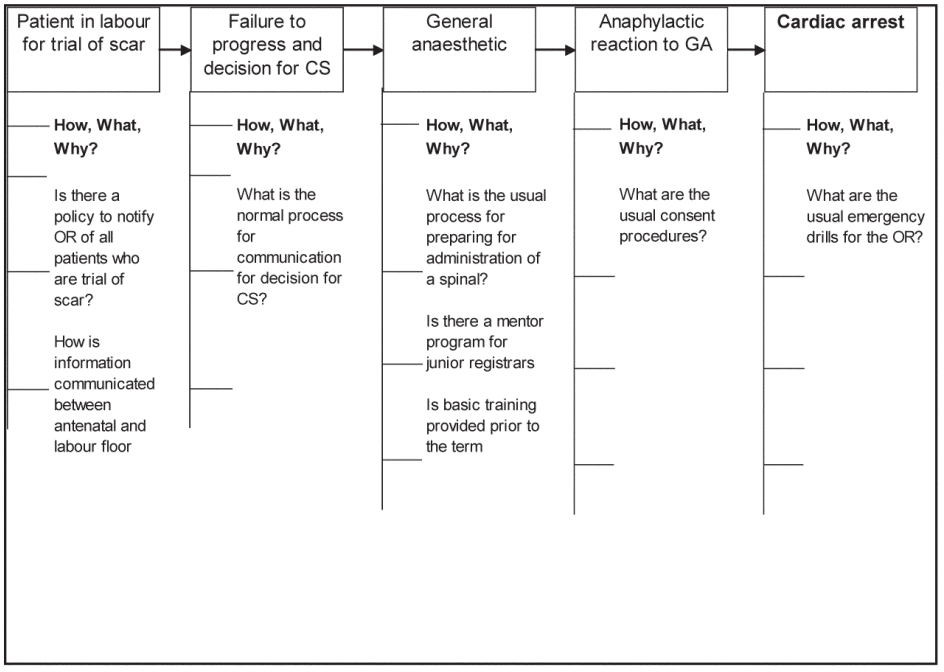


**Figure 4.** Expanded flow diagram.

dealt with. These environmental factors and decision points must be understood in order to identify the root cause or contributing factors leading to the adverse event. Thus, the importance of this stage of the process cannot be overstated. It identifies gaps in the story and sequence of events and assists in highlighting contributive factors and behaviours and provides the initial substance for the RCA, without which a team cannot proceed.

To finalise the first meeting, the team should agree on the following tasks:

- The questions that need to be asked and to whom;
- The additional information that needs to be obtained; and,
- Who is going to do each task.

## 4. Determining the root causes

### *Detailed flow diagram and Cause and Effect Diagram: Meeting 2*

After the interviews with staff identified at Meeting 1 have been completed and the additional information has been collected, the team needs to collate this information and ask "so what", or "what is the relevance", of each piece of information collected. This will help identify the symptoms or primary causes. This will lead to the root causes and contributing factors and what might be seen as the crucial "Barrier Points" where the event may have been prevented and can be set out in a detailed flow diagram, as in Figure 5.

The next step is the development of a cause and effect diagram, such as that in Figure 6. This assists the team in analysing relationships between a problem or symptom and its causes. It is a systematic way to combine the previous brainstorming, interviews and flowcharting tasks. The cause and effect diagram assists in:

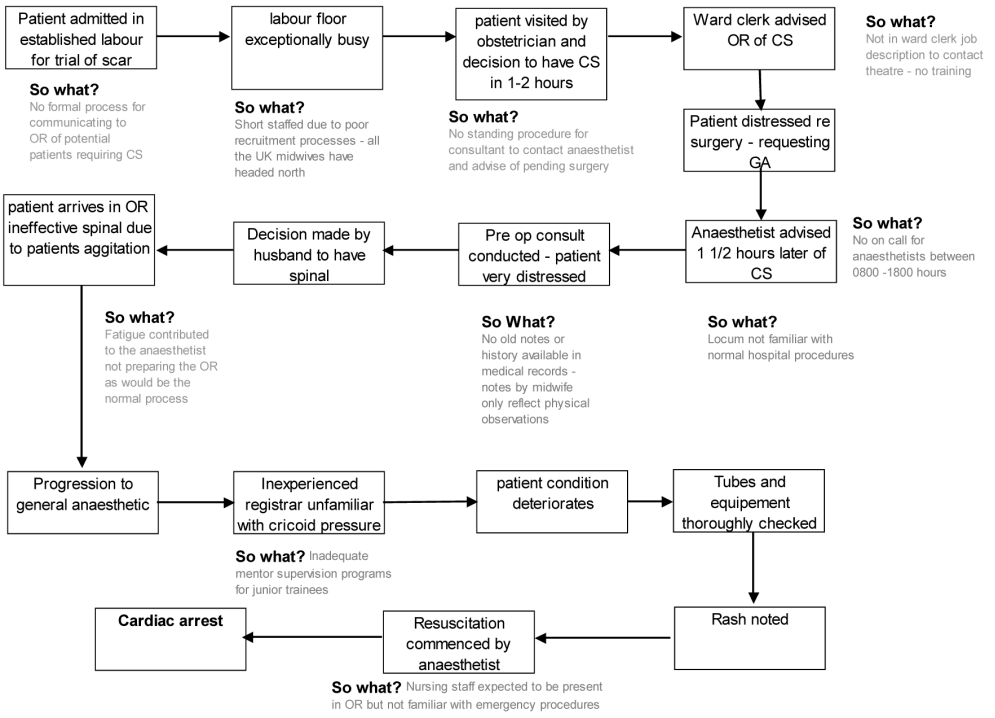


**Figure 5.** Detailed flow diagram.

- Clearly describing the problem for which the causes are sought — "What is the real problem we want to prevent?"
- Identifying the main categories of causes of the problem.
- Encouraging the team to dig deeper into each primary cause by repeatedly asking "What was this was caused by or what did it result in?", until the contributing factor root cause is identified.
- For each new answer to the question, ask the question again until no new answer results. (As a rule of thumb this often takes three to five rounds).

The cause and effect diagram is a four-step process:

1. Reviewing the event flow diagram and clarifying the problem statement — what is it you want to prevent occurring again?
2. Brainstorming a list of primary causes and choosing the most important.
3. Completing the causal chain for each of the primary causes or symptoms.
4. Concluding the investigation by developing root cause/contributing factor statements, actions and recommendations.

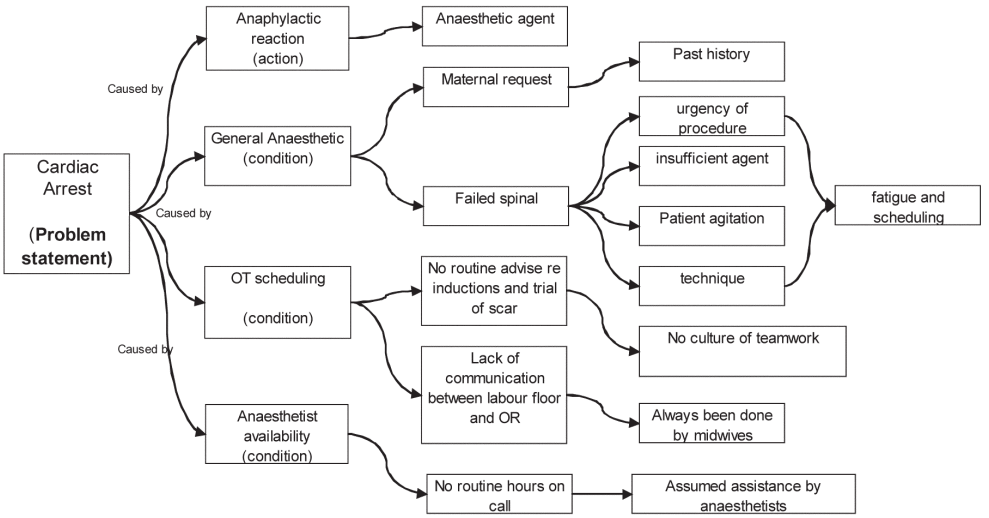


**Figure 6.** Cause and effect diagram.

***5. Causal statements, actions and recommendations: Meeting 3***

Once the cause and effect diagram has been completed, the team needs to develop root cause and contributing factor statements. These address why something occurred, not by whom it was caused. They should focus on what are the system vulnerabilities, not on individuals. The wording of causal statements has been described as cumbersome. We believe that this wording is crucial in that it prevents ambiguity and drives management to specific action that will prevent this incident in the future.

Root causes may include:

- Errors;
- Omissions;
- System deficiencies;
- Inadequate competencies;
- Poor communication or documentation;

- Inadequate facilities or equipment;
- Inadequate skill mix or availability of the health care team; and,
- Managerial inaction.

It is usual to identify more than one root cause; consequently, the team will need to prioritise the solutions to each cause.

***Actions and Recommendations***

The key step to the RCA process is a decision on "actions and recommendations" to address the root causes that either directly or indirectly contributed to the event. Each root cause should be described and have the corresponding category from the checklist flipchart noted. The actions and recommendations will need to answer the question, "Will this prevent this incident in the future?" These actions may:

1. E*liminate* the factor — remove, fix or replace a piece of equipment or put a measure in place to ensure the problem cannot recur.
2. *Control* the factor — checklists, cognitive aids, enhanced documentation, reduced disturbances. Or,
3. *Accept* the factor — place a warning notice or have reminders at team meetings and orientation etc, in effect acknowledging that there is an associated risk and accept it.

In addition, the team is expected to make recommendations about:

- What is the most effective remedial action?
- Who should be responsible for implementing the action?
- Outline a reasonable outcome measure for the action; and
- Set realistic target dates for measurement.

In our example, recommendations may include review of guidelines for "trial of scar", protocols for communication between labour ward and operating theatres, when an urgent Caesarean section is possible, and the rostering of an anaesthetist specifically for obstetric patients during the day.

**6. Monitoring and Evaluation**

While not specifically addressed in the training program, it is essential that organisations have a process in place that reviews the implementation of recommendations and their effectiveness. This should be undertaken by the management of organisations and reported regularly through Area Committees.

**Conclusion**

It is recognised that there are many ways to improve the health care environment to make it safer for patients. The program that we have outlined on root cause analysis is one of these and concentrates efforts on improving patient safety through a focus on all types of errors. The causes of injury are rooted at the deepest organisational levels of clinical care and are difficult to observe and measure. Root cause analysis methodology provides an opportunity to reduce future injuries by identifying and addressing system related issues in health care. Through addressing clinical incidents in an environment that encourages reporting, thorough analysis of system vulnerabilities and making management accountable to ensure that recommendations are acted upon, Area Health Services not only stand to develop much more efficient services, but ones that will provide much greater customer satisfaction.

The training program provided by the Institute for Clinical Excellence has already

witnessed cultural change in places where RCA has been implemented. The end of 2003 will see over 2500 staff trained in the methodology. This should provide an immense impetus to cultural change and, even more importantly, lead to thousands of system improvements throughout NSW to provide a safer environment for patients. This process finally is using available information to put actions in place.

## REFERENCES

1. Brennan T et al. Incidence of adverse events in hospitalised patients. *New Engl J Med* 1991; 384:370-372.
2. Wilson R et al. The Quality in Australian Health Care Study. *Med J Aust* 1995; 163:458-471.
3. Institute of Medicine Report: To Err is Human 1999.
4. Lessons from the inquiry into obstetrics and gynaecological services at King Edward Memorial Hospital 1999-2000. The Safety and Quality Council, July 2002.
5. Guidelines for managing risk in the healthcare sector (AS/NZS HB 228:2001).
6. Sentinel events: evaluating cause and planning improvement. The Joint Commission on Accreditation of Healthcare Organizations, 2nd Edition, 2002.
7. Anderson B, Fagerhaug T. Root Cause Analysis. Simplified Tools and Techniques. ASQ Quality Press, Wisconsin 2000.

# In-flight Medical Emergencies — A Difficult Medical and Legal Environment

MARK LOVELL, MB, BS, FANZCA
The Children's Hospital at Westmead, Sydney, NSW

*Mark Lovell is a Staff Specialist at Westmead Hospital and The Children's Hospital at Westmead, Sydney. His interests include paediatric and obstetric anaesthesia.*

## INTRODUCTION

There is an old joke about in-flight emergencies that goes something like this.

An announcement comes over an aircraft intercom, "Is there an anaesthetist on the plane?" After many requests, an anaesthetist reluctantly volunteers to a grateful flight attendant who thanks him and says breathlessly that there is a surgeon in first class who urgently needs his light adjusted! Despite widespread misconceptions about the role of the anaesthetist, we are actually well placed to provide help during in-flight emergencies. Anaesthetists are specifically trained in the care of critically-ill patients, with airway and resuscitation skills in advance of many other specialties. This article discusses the environmental and physiological factors peculiar to airline travel, the assessment and treatment of in-flight emergencies and the medico-legal issues relevant to your decision to volunteer assistance, especially when such emergencies are attracting the attention of the media, travellers and the medical profession.

Every year, commercial airlines transport hundreds of millions of passengers. Air transport is extremely safe with a very low death rate of about 0.31 deaths per 1 million passengers.[1] However, passengers with pre-existing disease are now travelling in increasing numbers and there is an expectation that commercial airlines will assist them with any in-flight emergencies. Medical practitioners from all specialities are frequently asked to volunteer during these events. How should they respond to in-flight emergencies, when they are passengers?

## EPIDEMIOLOGY OF IN-FLIGHT EMERGENCIES

The incidence of in-flight medical events appears to be increasing.[2] There have been a number of studies of this, but currently there is no mandatory reporting of events aboard commercial aircraft and so accurately assessing the incidence of such emergencies has been difficult. In 1999, a British report found an incidence of one in-flight emergency for each 11,000 passengers,[3] while another study quoted an incidence of 1 in 39,000 in-bound passengers.[4] The death rate amongst passengers from a 1997 study was 1 in 6.4 million paying passengers.[5] Only about one in five deaths are amongst passengers with a documented severe illness.[1]

Most in-flight events on commercial flights are not serious. By far the most common are fainting episodes, or "syncope". After syncope, angina, cardiac events, gastro-intestinal conditions, asthma, anxiety and panic attacks were the next most common.[2]

Anxiety, alcohol abuse, sleeplessness, immobility and barotrauma contributed to the difficulties faced by passengers.[3]

## DOCTOR PARTICIPATION

Despite the relatively low frequency with which in-flight emergencies occur, a recent survey showed that 62% of physicians had been called on at least once to care for a sick or injured passenger on a commercial airline.[6] The same study showed that medical personnel were generally happy to give assistance during in-flight emergencies, but fear of litigation made physicians reluctant to offer their help.[6] A study by the Federal Aviation Authority (FAA) in the US found that 69% of in-flight emergencies aboard US airlines were attended by some form of health care professional — medical doctors 40%, nurses 25% and paramedics 4%.[7]

## AIRCRAFT DIVERSIONS

Diversions for in-flight emergencies are a time-consuming and costly exercise for all commercial airlines. Syncopal episodes, suspected heart attacks, neurological events and dyspnoea make up the majority of cases requiring unscheduled landings.[2] Other causes include seizure activity, abdominal pain and haemorrhage. A recent report of 201 in-flight medical emergencies on one commercial airline over a 4-year period reported eleven diversions and three in-flight deaths. The most common reasons for diversion in this study were cardiac disease (36.3%) and bleeding problems (27.2%).[8] Commercial airlines are increasingly looking for ways to decrease the cost imposed by these emergency diversions, which can range from $3000 to $100,000 because of the need to dump fuel prior to landing and provide accommodation for stranded passengers.[9]

## ENVIRONMENTAL & PHYSIOLOGICAL FACTORS

*Cabin Pressure*

For reasons of passenger comfort and safety, it would be desirable in passenger aircraft to maintain a cabin pressure close to the atmospheric pressure at sea-level. However, this is not possible for structural design and operational efficiency reasons. Modern commercial aircraft fly at altitudes of between 9000 m and 12000 m, primarily to decrease fuel costs but also to avoid unfavourable weather systems.[5] At 12000 m, the atmospheric pressure outside the aircraft drops from 760 mmHg to 176 mmHg. Within the cabin, pressure is maintained by drawing in external air and limiting its outflow. Thus, a differential pressure is established between the outside and the inside of the plane. Most aircraft are able to increase the pressure within the cabin, but only to a pressure differential of about 430 mmHg. Thus, cabin pressure changes in parallel with altitude, but to a much lesser degree than the drop in external atmospheric pressure. For example, at 12000 m where atmospheric pressures outside the aircraft is 176 mmHg, the cabin pressure will be 176 mmHg+430 mmHg, or 606 mmHg. This is equivalent to the atmospheric pressure found at an altitude of 1600 m, therefore giving a “cabin altitude” of 1600 m.

Aboard the supersonic Concorde, the situation is somewhat different as it operates up to a cruising altitude of 18300 m (60,000 ft). By necessity, a greater pressure differential must be reached to provide a near normal cabin environment. This is achieved by drawing even more compressed air from its significantly more powerful engines into the cabin. A differential pressure of 530 mmHg can be maintained,

ensuring reasonable cabin pressures at this extreme altitude, where the external atmospheric pressure is only 51 mmHg.[10]

The drop in cabin pressures generally has minimal effect for healthy passengers, but will result in a drop of arterial partial pressure of oxygen from about 95 mmHg to about 56 mmHg. This represents only a 4% reduction in the amount of oxygen carried by the blood as, even at a $PaO_2$, of 56 mmHg the passenger still lies on the flat part of the oxygen-haemoglobin dissociation curve. Difficulties occur in those patients with significant cardio-respiratory disease, who have lower $PaO_2$ at sea level. Any reduction in atmospheric pressure may see them move onto the steep portion of the dissociation curve, where oxygen saturations may fall quickly. These passengers are at risk of hypobaric hypoxia. Their vulnerability to hypoxaemia at altitude depends on their $PaO_2$ at sea level and their physiological ability to compensate for the decrease in $PiO_2$ as the plane ascends.[5]

Fortunately, depressurisation of commercial aircraft through equipment failure or accident is an extremely rare event. When there is a slow loss of cabin pressure, then an aircraft is able to descend to a level at which the outside atmospheric pressure is adequate for passenger oxygenation. For healthy passengers who are exposed to external atmospheric pressures, 100% oxygen given by face mask should still provide adequate protection against loss of consciousness, up to an altitude of about 12000 m (40,000 ft). However, Concorde operates at altitudes at which even 100% oxygen would be insufficient, should the passengers be exposed to external atmospheric pressures. Its design therefore includes features such as small windows, which aim to minimise the rate of accidental decompression and is also engineered to have a reserve capacity for cabin pressurisation. Special pressurised breathing apparatus ensures that the pilots and crew can function normally in the event of a rapid depressurisation.[10] This equipment is normally only used for pilots operating military aircraft at high altitudes.

*Gas Expansion*

In accordance with Boyle's Law, air and gas trapped in body cavities expands in direct proportion to the decrease in atmospheric pressure. A cabin pressure equivalent to the pressure at an altitude of 1600 m will result in expansion of gas volume by about 30%.[11] Symptoms such as abdominal cramping, ear and tooth ache occur in healthy passengers as a result of this gas expansion.[3, 5] The case involving Dr Angus Wallace, a trauma surgeon from the UK, dramatically illustrates the potential for gas expansion. He was faced with a passenger who developed a tension pneumothorax aboard a commercial aircraft; he managed to relieve this with the aid of a coat hanger, urinary catheter and a brandy bottle![12]

More particularly, this effect must be considered in the circumstances of patient transport. For instance, patients who have had recent surgery are at increased risk of wound dehiscence. Care must be taken with medical equipment such as pneumatic splints, urinary catheters, cuffed endotracheal tubes and tracheostomy tubes in which gas containing closed spaces may expand during flight.[5] Instillation of water into the endotracheal cuff has been recommended to avoid this problem.

*Cabin Air Quality*

Most studies of the composition and quality of commercial aircraft cabin air have found that it is remarkably clean, even more so than in the home or workplace.[13] When

compressed air has passed through the jet engines, its temperature is initially very hot (250°C); it is then cooled at high pressure. These harsh physical properties of aircraft intake systems provide essentially sterile cabin air.[13]

When passengers exhale potential respiratory pathogens, the infectious agents enter laminar sheets of air flowing vertically from top to bottom of the cabin and are then diluted by frequent air exchanges. (Commercial aircraft exchange cabin air every 3-4 minutes.) Studies have shown that the concentration of bacteria within cabins is less than that found on city buses, shopping malls or even in airline terminals[14] although there have been sporadic reports of the transmission of tuberculosis and other infectious diseases aboard commercial flights.[15, 16] However, the risk of cross-infection aboard commercial airlines seems to be primarily determined by the length of the flight (>8 hours) and the proximity to the index passenger (i.e. sitting within 2 rows indicating an increased risk).[17]

Because the air drawn in from the engines at altitude has very low water content and engineering constraints do not allow for cabin humidification systems, humidity is invariably low. Most aircraft cabins have a relative humidity of between 15-25%.[10] Passengers will often complain of dry skin and itchy eyes. Additionally, the low humidity will tend to exacerbate reactive airways disease in susceptible passengers.[13]

*Air Rage and Other Disturbances*

Aggressive or disruptive behaviour aboard aircraft is an increasing problem and presents great difficulties for the air crew, because of the disruption to other passengers, the potential for passenger and crew injury and the very real risk to the safe working of the aircraft. Sporadic reports of "Air Rage" appear in both medical journals and the lay media. The intoxicating effects of alcohol are enhanced at altitude and may contribute to the problem, along with frustration associated with the smoking ban for nicotine addicts. There have been instances requiring physical restraint of passengers, and some occasions requiring physicians to assist with sedating agitated or violent passengers. In one incident, a passenger became abusive and violent on a flight to Turkey and had to be restrained. When his agitation worsened, a doctor aboard the aircraft was asked to assist by sedating the passenger, which he did with 10 mg of intravenous diazepam. Approximately five minutes before landing, the patient stopped breathing and soon died. The aircraft made an emergency landing in Istanbul, where both the crew and the treating doctor were arrested. They were questioned overnight, but allowed to continue their travel the next day.[18] Such incidents highlight the difficulties associated with treating aggressive patients who may be under the influence of other drugs, who may not be fasted and who need to be treated within the close confines of a commercial aircraft.

**FITNESS FOR AIR TRAVEL**

Providing medical advice on flying to passengers with pre-existing disease is very difficult. It might be expected that those with prior significant illness would have greater morbidity and mortality during air travel. One approach to prevention of in-flight emergencies might be to increase travel restrictions for people who have a documented severe illness.[1, 5] Implementing these types of restrictions would be complex and possibly discriminatory. The Air Carrier Access Act of 1986 prohibits airlines from discriminating against passengers with disabilities; however, an airline would still be able to refuse passengers considered at risk for travel. None the less, a

policy that excluded "at risk" passengers may only eliminate up to 20% of in-flight deaths; as noted earlier, only about 1 in 5 deaths aboard commercial aircraft occur in passengers with documented severe illness.[1]

Many guidelines on assessing fitness for air travel have been published but, in general, travel on commercial airline is contraindicated if a medical condition would be adversely affected by hypoxic or pressure changes.[1, 5, 13]. If a patient has a pre-flight arterial oxygen tension of less than 70 mmHg, supplementary oxygen should be recommended.[19] For individuals in whom excessive oxygen can worsen hypoventilation, there are centres where a simulation of altitude, by inhalation of hypoxic gas mixtures, can be done pre-flight. Supplemental oxygen for in-flight medical use can be arranged, but requires 48 hours notice and a prescription for the gas. Passengers cannot use their own equipment during flights, as oxygen is considered a hazardous material.

Specific medical contraindications to air travel include recent myocardial infarction, recent stroke, and decompensated cardiac disease. Others are pneumothorax, recent eye surgery, uncontrolled seizures, recent skull fractures, severe anaemia, sickle cell disease (especially above 7000 m) and pregnancy beyond 35 weeks gestation.[5, 20]

## IN-FLIGHT MEDICAL KITS

The standard of in-flight medical resources varies greatly worldwide. The issue has recently been raised in both the US Congress and in a special report by the Chicago Tribune entitled "Code Blue-Survival in the sky".[21] As public expectation increases, a balance has had to be found between the competing needs of cost effectiveness, storage space, staff training, security concerns and the short shelf life of drugs. Currently, the FAA requires a compulsory medical kit which includes basic equipment such as stethoscope and sphygmomanometer, along with a range of emergency drugs including adrenaline, nitroglycerin, lignocaine, antihistamines and bronchodilator inhalers.[9] Of interest to anaesthetists is that the equipment available includes self-inflating bag, masks and oropharyngeal airways. However, there is no requirement for equipment such as laryngoscopes and endotracheal tubes.

In Australia, Qantas carries a more extensive medical kit that includes laryngoscopes, endotracheal tubes and suction equipment. Medications such as haloperidol, morphine, benztropine and naloxone are also included.

### *Automated External Defibrillators*

Virgin Atlantic Airways was the first to carry automatic external defibrillators in 1990. Provision of defibrillators was controversial until findings showed that cardiac arrest aboard aircraft was most commonly due to ventricular fibrillation.[22] The previous policy of cardiopulmonary resuscitation and emergency diversion was found to be generally unsuccessful for the patient and expensive for airlines. Recent litigation has increased the availability of defibrillators, as well as the training of cabin staff in their use. A further benefit has been the use of defibrillators as portable ECG monitors for patients with cardiac ischaemia, allowing better decision making.[22] From April 2004, all US flights with at least one flight attendant must carry an automated external defibrillator.

There are also trials of advanced telemetry systems capable of transmitting 12 lead ECG, pulse oximetry, end-tidal capnography, pulse rate, blood pressure, together with a video and communications link direct to ground-based medical assistance teams. Virgin Atlantic is currently the only airline to use this system.[11] Recent studies have

shown that long-term survival following the use of automated external defibrillators aboard commercial airlines is still only about 27%.[22]

*Ground-based Medical Assistance*

Recently airlines have been using the services of ground-based medical assistance companies to cover those instances when medical assistance is not available aboard a commercial aircraft. For instance, Medlink is a ground-based medical assistance company based in the emergency department of the Good Samaritan Regional Medical Centre in Phoenix, Arizona. It provides direct communication between flight crew, volunteering medical staff and ground-based doctors. Information can be provided about treatment of patients as well as lists of airports suitable for diversions and details of their medical facilities. Reassuringly, once the company has been contacted by an air crew, the airline and any assisting physician are relieved of liability in the treatment of a sick passenger.[11] It is hoped ground-based medical assistance will be cost effective, given the very high costs associated with emergency diversions.

## MANAGEMENT OF IN-FLIGHT EMERGENCIES

A doctor travelling on board a commercial aircraft may encounter any type of medical emergency. The cramped conditions, difficult access to the patient, lack of privacy, cultural and language difficulties, noise and vibration make providing assistance very difficult. Simple monitoring procedures such as auscultation and blood pressure measurement are affected, even by low levels of ambient noise and vibration.

It is not within the scope of this discussion to recommend treatments for all the various medical emergencies that may be encountered. However, there are some overriding principles that can be employed. As a general rule, the flight crew is responsible for responding to a passenger who is ill. The role of any medical personnel who volunteer is to assist the crew with the management of this patient, but not to take over. Overall, the goal is to stabilise the patient until the aircraft has landed and more definitive treatment can be instituted. As anaesthetists, we have skills for effective treatment of the critically-ill but usually simple first-aid techniques will prove the most useful.

Firstly, the provision of oxygen may be life-saving. If cardiac or respiratory disease is the concern, it may be of benefit to ask for the plane to descend to below 7000 m (22,000 ft). In less serious circumstances, vaso-vagal episodes may be simply be treated by laying the patient down and raising the legs. Hypoglycaemic episodes are relatively common, as patients may inject insulin prior to the flight and then be delayed in eating their meals. Oral glucose preparations or intravenous 50% dextrose are present in most kits. Allergic reactions can often be improved with antihistamines, corticosteroids or adrenaline. Some airlines will carry adrenaline pens in their kits.

If available, consultation with ground-based medical staff will assist with management as well as providing some protection from medical litigation. When simple measures are inadequate, diversion of the aircraft should be suggested. The following conditions are examples:

- Cardiac arrest.
- Chest pain (unrelieved by nitrates, aspirin and oxygen).

- Shortness of breath (unrelieved with oxygen, lowering of altitude or bronchodilator therapy).
- Cerebrovascular accident.
- Uncontrolled seizure activity.
- Unresponsive patient (not improved with 50% Dextrose administration).

**MEDICOLEGAL ISSUES**

Despite recent advances in ground-based medical assistance, there is still the need for cabin staff to call on travelling doctors for medical assistance. A recent survey of 850 physicians showed that 62% had, been asked to provide medical assistance aboard a commercial airliner at one time or another.[6] The legal situation regarding liability and obligation to treat aboard commercial airlines varies worldwide. In the United States, the 1998 Aviation Medical Assistance Act contains wording designed to protect airlines and individual doctors from medical liability. It states that individuals "will not be liable for providing assistance unless the individual is guilty of gross negligence or wilful misconduct".[24] In addition to being medically qualified, the assisting passenger must be a volunteer, render care in good faith and receive no monetary compensation. Gifts in the form of travel voucher, alcohol or seat upgrades are generally not considered monetary compensation. However, there is a report of an Australian orthopaedic surgeon, who was given a travel voucher for providing medical assistance aboard a US airline. After redeeming the voucher he read the conditions of acceptance; these included absolving the airline of any liability in the treatment of the passenger.[25]

There are cases where volunteering medical personnel have attempted to obtain payment for providing medical assistance aboard an aircraft. In one case, the volunteering doctor was a specialist psychiatrist who billed the airline for the five hours he treated a passenger, who was suspected of having a pulmonary embolus.[26] The psychiatrist, who was initially reluctant to volunteer, only billed the company after feeling aggrieved that he had only been given a bottle of champagne and a travel voucher worth $50. The airline refused to pay the bill and the doctor took them to court!

At this time there is no record of a passenger bringing suit against a physician in a US court related to an in-flight medical event.[24] There have, however, been instances where airlines have successfully been sued for medical incidents. Lufthansa Airlines was forced to pay $2.7 million to a passenger who claimed to be having a heart attack during a 9-hour flight. The captain had refused to land, having been advised by a doctor aboard the aircraft that there was no necessity to do so.[27]

*When does a doctor travelling as a passenger have an obligation to treat?*

In the United States, Canada and the United Kingdom doctors do not have a legal duty to render assistance, unless there is a pre-existing physician-patient relationship. In fact, in the UK, the law seemingly discourages the "Good Samaritan". If a doctor comes to the aid of the sick person, he or she undertakes a duty of care and will be liable if skills fail. On the other hand, European Law does impose an obligation to treat. For example, under French law there is an obligation to render assistance where a person is in need of help; backed up by threat of fines or imprisonment.[27]

In Australia, there is case law to suggest that there is an obligation to treat. One case involved a general practitioner who declined to attend a patient having a grand mal

seizure, a short distance from her practice. The patient subsequently developed cerebral ischaemia and severe neurological deficit. The court found a duty of care did in fact exist and found against the doctor, despite there being no previous patient-doctor relationship.[25]

Developing recommendations from all the various case law, international law and conflicting jurisdictions is extremely difficult. However, there are some general principles which may help to guide a doctor who is called on to assist with an in-flight emergency and is concerned about liability.

1. Always wait for the airline to ask for your assistance. In this way, the airline and the doctor are at least equally responsible for incidents related to the ill passenger.
2. Clearly identify yourself and state your qualifications.
3. Obtain as comprehensive medical history as is possible, with the aid of interpreters, if available. Tell the patient and their family of your clinical impression and obtain consent prior to any examination or treatment.
4. If you are unsure of diagnosis or management, ask if there are other medically qualified persons aboard the aircraft.
5. Quickly assess the equipment and drug kit.
6. Consider moving the passenger to a less populated or more spacious area within the aircraft, if this is possible.
7. Ask if the airline has the services of ground-based medical assistance and use them. When this is done, the service then takes overall liability and is responsible for any flight diversion necessary.
8. Ask the flight crew for an estimated time of arrival or, alternatively, the time to the nearest appropriate airport. If the patient's condition is serious, inform the captain of the need for urgent treatment and discuss diversion options.
9. Always make written notes during treatment or shortly after, including clinical impression, treatment and details of communication with flight crew and ground-based medical services.
10. Do not use any treatment you are uncomfortable with; basic first aid principles will often be sufficient until definitive treatment can be instituted on the ground.
11. Do not accept monetary compensation. (Upgrades or small gifts are probably acceptable.)

These points should be taken as a general guide to the difficult decisions that have to be made when volunteering assistance during in-flight medical emergencies.

We should remember that when we offer to assist someone in distress we perform an act of basic humanity. Let us hope that despite all the environmental and legal difficulties associated with in-flight emergencies we can still provide timely and skilful assistance.

## REFERENCES

1. Cummins RO, Chapman PJC, Chamberlain DA et al. In-flight deaths during commercial air travel. *JAMA* 1988; 259:1983-1988.
2. Donaldson D, Pearn J. First aid in the air. *Aust NZJ Surg* 1996; 66:431-434.
3. Dowdall N. "Is there a doctor on the aircraft?" Top 10 in-flight medical emergencies. *Brit Med J* 2000; 321:1136-1137.
4. Cummins RO, Schubach JA. Frequency and types of medical emergencies among Commercial air travellers. *JAMA* 1989; 261:1295-1299.
5. AMA Commission on Emergency Medical Services. Medical aspects of transportation aboard commercial aircraft. *JAMA* 1982; 247:1007-1011.
6. Raymen RB. In-flight medical kits. *Aviat Space Environ Med* 1998; 69:1007-1010.

7. Dejohn CA, Veronneau SJ, Wolbrink AM et al. The evaluation of in-flight medical care aboard selected US carriers 1996-1997. FAA Administration, Office of Aviation Medicine. (Technical report no. DOT/FAA/AM-0013).
8. Han HM, Chae DH, Kim JH. In-flight medical emergencies in civil airline. *Aviat Space Environ Med* 2000; 71:330.
9. Gendreau MA, DeJohn C. Responding to medical events aboard commercial airline flights. *N Engl J Med* 2002; 346:1067-1073.
10. Nunn. Nunn's Applied Respiratory Physiology.
11. Goodwin T. In-flight medical emergencies: an overview. *Brit Med J* 2000; 321:1338-1341.
12. Wallace WA. Managing in-flight emergencies. *Brit Med J* 1995; 311:1508.
13. Rayman RB. Passenger safety, health and comfort: A review. *Aviat Space Environ Med* 1997; 86:432-440.
14. Wick RL, Irvine LA. The microbiological composition of airline cabin air. *Aviat Space Environ Med* 1995; 66:220-224.
15. Wenzel RP. Editorial. Airline travel and infection. *New Engl J Med* 1996; 334:981-982.
16. Mcfarlane JW, Hickmann C, Osteholm M, McDonald KL. Exposure to mycobacterium tuberculosis during air travel. *Lancet* 1993; 342:112-113.
17. Kenyon TA, Valway SE, Ihle WW, Onorato IM, Castro KG. Transmission of multi-drug resistant mycobacterium tuberculosis during a long airplane flight. *New Engl J Med* 1996; 334:933-938.
18. Kovac C. Airline passenger dies after being sedated by doctor. *Brit Med J* 1999; 318:12.
19. Cottrell JJ. Altitude exposure during aircraft flight. *Chest* 1988; 92:81-84.
20. Barry M, Bia F. Pregnancy and travel. *JAMA* 1989; 261:728-731.
21. Code Blue: Survival in the sky. *Chicago Tribune* June 30, 1996.
22. O'Rourke MF, Donaldson E, Geddes JS. An airline cardiac arrest program. *Circulation* 1997; 96:2849-2853.
23. Page RL, Joglar JA, Kowal RC et al. Use of automated external defibrillators by a US airline. *New Engl J Med* 2000; 343:1210-1216.
24. Lyznicki JM, Dietchman SD, Howe JP. In-flight medical emergencies. *Aviat Space Environ Med* 2000; 71:832-838.
25. The Law Report, Part 2. Health law: radio national transcripts. Sydney Australia: Australian Broadcasting Corporation, August 3, 1999.
26. Dyer C. Doctor demands payment for helping airline passenger. *Brit Med J* 1998; 317:701.
27. Newson-Smith MS. Passenger doctor in civil airliners — Obligations, duties and standards of care. *Aviat Space Environ Med* 1997; 68:1134-1138.

# A Risk-Benefit Analysis of Thoracic Epidural Anaesthesia and Analgesia

STEPHAN A. SCHUG, MD(Cgn), FANZCA, FFPMANZCA
Professor of Anaesthesia, School of Medicine and Pharmacology, University of Western Australia & Director of Pain Medicine, Royal Perth Hospital, Perth, Western Australia

*Stephan Schug was previously the Professor and Chair of Anaesthesiology at the University of Auckland, New Zealand, where he lived and worked from 1989 to 2001. He moved then to Perth, where he took up his present positions with the University of Western Australia and Royal Perth Hospital in March 2001. He is a graduate of the University of Cologne, Germany, where he also obtained his MD in clinical pharmacology and his specialist qualification in anaesthesia and intensive care. His areas of research and clinical interest are in the management of acute, cancer and chronic pain, regional anaesthesia, pharmacology of anaesthetics and analgesics and quality of health care.*

This paper will focus on the use of perioperative thoracic epidural anaesthesia and analgesia for major abdominal and thoracic surgery. It will briefly present the intraoperative benefits, but discuss in more detail postoperative epidural analgesia for this kind of surgery. It will also provide suggestions to maximise the benefits and reduce the risks of this technique, based on scientific evidence and practical experience.

There is now widely discussed, level 1 evidence in the form of a meta-analysis which shows significant advantages of the intraoperative use of epidural anaesthesia in comparison to general anaesthesia, with regard to perioperative morbidity and mortality.[1] This meta-analysis of 141 trials which randomised 9559 patients between neuraxial and general anaesthesia, showed not only significant reduction of peri-operative morbidity (deep vein thrombosis, pulmonary embolism, transfusion requirements, pneumonia, respiratory depression, myocardial infarction and renal failure), but also about a third reduction in overall mortality using neuraxial techniques. While these results alone are a strong argument in favour of the use of epidural anaesthesia, there is now widespread evidence that continuation of epidural anaesthesia as epidural analgesia into the postoperative period can provide further benefits. However, there is also widespread concern about the risks of epidural anaesthesia and analgesia, which some regard as unacceptable.

**BENEFITS**

There is no doubt that appropriate usage of epidural analgesia can provide better postoperative analgesia after major abdominal surgery than any other technique. This was shown in a double-blind placebo controlled trial, where pain scores at rest and on

movement were in the range of 40-50/100 with systemic administration of opioids, but in the range of 0-10 with appropriate thoracic epidural analgesia.[2] Superior pain relief was also the major advantage of epidural analgesia in the MASTER trial, recently published in *Lancet*.[3]

This improved analgesia translates obviously into reduced pulmonary morbidity, by potentially improving respiratory function after surgery and permitting better compliance with physiotherapy. Again, reduced respiratory failure was the only morbid endpoint improved by epidural anaesthesia and analgesia in the MASTER trial.[3] Such reduced pulmonary morbidity has also been shown by a meta-analysis of randomised controlled trials comparing epidural local anaesthetics with systemic opioids.[4]

Further advantages result from blockade of the sympathetic nervous system by epidural administration of local anaesthetics. These include improved gastrointestinal recovery and reduction of duration of peri-operative ileus. For example, Liu and his group showed a halving of the time until first flatus after colonic surgery by the use of thoracic epidural analgesia with a morphine-bupivacaine mixture. Such improved gastrointestinal recovery has further outcome benefits. In the same study, patients with thoracic epidural analgesia fulfilled discharge criteria at 67 hours while patients with PCA did this after 96 hours.[5] Again, a meta-analysis for the Cochrane collaboration confirmed these result.[6] Steinbrook has written an excellent review of this issue.[7]

Obviously, sympathetic blockade also has benefits with regard to attenuated stress response after surgery. This has been shown in a number of randomised controlled trials in high-risk patients undergoing coronary artery bypass grafting. Here, attenuated plasma epinephrine increase, attenuated troponin-T release, reduced incidence of ST changes and reduced incidence of new arrhythmias requiring treatment could be shown.[8] Similar trends have also been shown for upper abdominal surgery in patients with high risk factors for coronary artery disease. De Leon-Casasola's group showed that, in such patients, the incidence of tachycardic episodes and subsequent ST changes were significantly reduced by thoracic epidural analgesia.[9] These results were confirmed by a recent meta-analysis, which showed a significantly reduced perioperative myocardial infarction rate with the use of thoracic epidural techniques in comparison to general anaesthesia. [10]

In conclusion, there is now level-1 evidence, based on meta-analysis or multiple large randomised controlled trials, that thoracic epidural anaesthesia and analgesia reduces postoperative pain severity, respiratory failure and pulmonary complications, time to bowel recovery, perioperative myocardial infarction rate and many other types of perioperative morbidity, as well as perioperative mortality.[1, 3, 4, 6, 10]

**Maximising benefits**

In accordance with the literature and from our own practical experience in more than 8000 patients receiving epidural analgesia postoperatively, the achievable benefits can be maximised by following a number of rules.[11]

1. Thoracic epidural analgesia should always rely on the continuous administration of low concentrations of local anaesthetics. Ropivacaine might be the local anaesthetic of choice here, in view of reduced systemic toxicity and less interference with motor function.
2. Patient-controlled epidural analgesia (PCEA) might offer some slight practical advantages, in particular with regard to the number of interventions required to achieve optimal analgesia.

3. Administration of epidural analgesia for abdominal surgery by lumbar catheters results in insufficient analgesia and potentially significant problems with motor blockade; it should be avoided in favour of thoracic catheter localisation.
4. Combining local anaesthetics with very small amounts of opioids results in a reduction of the incidence of patchy or unilateral blocks and improves the overall quality of analgesia. It is therefore the recommended technique.
5. The discussion on other additives continues. While clonidine improves analgesia and reduces top-up requirements, it results in a decrease in blood pressure and heart rate and increases vasopressor requirements. In my opinion, it is not an adjuvant for routine use. All other adjuvants are in a more experimental state with the exception of adrenaline. There is now some good data which suggests that adding low concentrations of adrenaline (in the range of 2 $\mu$g/ml) to local anaesthetic opioid mixtures results in better pain relief, more widespread block and reduced systemic side effects.
6. Last and possibly most importantly, thoracic epidural analgesia for abdominal surgery needs to be integrated into an overall multi-disciplinary rehabilitation approach to the postoperative period, as has been suggested by Kehlet for many years.[12] Utilisation of excellent analgesia and improved bowel recovery provided by epidural analgesia, permits early extubation, early aggressive mobilisation and early enteral feeding. This can result in significantly improved postoperative outcome with decreased discharge times and even significant cost reduction, as shown for example by van Aken.[13]

## ADVERSE OUTCOMES

While this brief summary illustrates the immense benefits thoracic epidural anaesthesia and analgesia can provide to patients after abdominal surgery, debate about this "dangerous technique" continues and polarises researchers and clinicians. The risks primarily discussed relate to damage to neural structures by needle and catheter insertion, epidural haematoma and epidural infection (with the potential to create the catastrophe of paraplegia) and hypotension as an acute complication.

### Neural damage

Data from large surveys suggest a very low incidence of nerve damage. The incidence of damage to peripheral neural structures is in the range of 0.02 to 0.24%. The larger number originates from a survey of 4185 thoracic epidurals, which found 10 such incidents.[14] All these neurological deficits were of a temporary nature and the authors conclude in a statistical analysis that the predicted maximum risk for permanent neurologic complications (upper bound of the 95% confidence interval) in the study is 0.07%. However, there is no doubt that extremely rare cases of spinal cord damage by epidural needle insertion in the thoracic epidural space have been reported. While this risk seems to be minimal, and possibly related to experience of the operator, it is a real risk with potentially catastrophic consequences in these rare cases.

### Epidural haematoma

Fortunately, epidural haematomas as a result of epidural catheter insertion are so rare that it is impossible to perform randomised controlled trials to establish cause and effect. All we have is anecdotal evidence in the form of case reports and case series. Most of the reported cases are related to use of anticoagulants or pre-existing

coagulopathy, to the use of a catheter, or to difficult traumatic or bloody tap procedures. They can also occur after catheter removal.[15]

On the basis of such case series, two well known reviews calculate the risk as 1:100,000 to 1:150,000 cases, at the upper confidence interval of 95%.[15,16] In view of the rarity of this event, it is not surprising that overall there are far more published cases of spontaneous epidural haematoma than of epidural haematoma caused by epidural anaesthesia and analgesia. This is also in line with my personal experience in running a large chronic pain clinic.

However, this positive impression has been significantly damaged by the horrifying experience of combining epidural anaesthesia with low molecular weight heparin (LMWH) in the United States. By May 1998, there were at least 50 case reports with regard to this complication known to the FDA. It is now obvious, that these complications related not so much to the respective drug or the epidural technique as to the inappropriate combination of both. In contrast to Europe, where such complications were extremely rare, the Americans routinely used low molecular weight heparin twice daily, commonly in too high a dose. This situation has now been brought under control by very well developed guidelines of the American Society of Regional Anesthesia. A simplified version of these guidelines has been used successfully at Auckland Hospital for the last six years.

> "LMWH requires an interval of at least 12 hours between last injection of a standard prophylactic dose and insertion or removal of epidural catheters; the next dose should be given at least two hours later."
>
> "To permit this approach, low molecular weight heparin should be prescribed only once daily in the evening in a standard dose in patients on epidural infusions."

While epidural haematomas will hopefully remain a rare complication and most of us will never see such a case, once one has developed extreme vigilance and urgency are the main factors in preventing permanent damage. Available data suggests emergency laminectomy within eight hours has a very good prognosis with regard to recovery of neurological function, while laminectomy more than 24 hours after the initial event has an extremely poor prognosis and permanent paraplegia commonly ensues. Symptoms to look for are sharp back pain, developing motor weakness, urinary retention and sensory loss, as complete paraplegia develops usually only after 12 hours and more. The best diagnostic technique is immediate MRI scan, as computer tomography is not as sensitive.

**Epidural infection**

In the past, epidural infection was regarded as extremely rare. Reports of an incidence in the range of 1:100,000 and more were common. However, a very careful study of all epidural catheters inserted in Denmark in the year 1998, showed an incidence of 1:2000 leading to neurological deficit in 1:4000.[17] It is obvious that quality of care must have played a role here as the incidence of epidural infections was more than seven times lower in university hospitals than in community hospitals. It is also of note that most of the patients who developed an epidural abscess in this study were immune compromised; some having epidural catheters not for postoperative but for cancer pain relief. Finally, no case occurred in patients with less than two days of catheterisation; the mean duration of catheterisation in patients with epidural infections was eleven days.

Reducing the risk of epidural infection requires a number of procedural and organisational guidelines for the pain service providing epidural analgesia. Again, literature and experience suggest that aseptic insertion technique, standardised dressing technique, regular review of dressing and entry site and preparation of infusion solutions under sterile conditions in a pharmacy or pharmaceutical production facility reduce the risks significantly. Furthermore, in patients with large infected cutaneous wounds and/or generalised septicaemia, the risk of epidural infection seems to be increased. However, these patients very often benefit most from epidural analgesia and one needs to make a risk/benefit assessment in each individual patient.

**Acute hypotension**

Hypotension as an acute problem is often discussed in the literature. Because of inconsistent criteria for its definition, the incidence is reported to be in the range of 2.6-8%. In our experience, hypotension with epidural analgesia is commonly caused by unmasking of hypovolemia, more than by the epidural technique itself. It is of interest that, in a comparison of cholecystectomy patients with thoracic epidural and systemic opioid analgesia, one study could not show significant differences in haemodynamic responses during rest, orthostatic stress or after walking. As well, there were no significant differences in the number of episodes of dizziness, nausea or vomiting during mobilisation.[18]

In our own experience, thoracic epidural analgesia can be provided safely by an anaesthesiology-based acute pain service. Over the last eleven years, the acute pain services under my supervision looked after more than 8000 patients receiving continuous epidural analgesia on normal surgical wards. Over this period, no mortality and no case of severe morbidity could be attributed to the use of this technique. No patient developed an epidural infection and no patient had significant permanent neurological deficit because of epidural analgesia. Potentially severe complications without consequences, collected by a continuous safety audit, occurred in 0.2% of the patients on epidural analgesia — an incidence similar to that we observed with systemic opioid analgesia.[19]

In conclusion, the risks of epidural analgesia, provided by a well organised anaesthesiology-based acute pain service on normal hospital wards, are minimal and are by far outweighed by the benefits this technique offers patients, in particular those at high risk.

**References**

1. Rogers A, Walker N, Schug SA et al. Reduction of post-operative mortality and morbidity with epidural or spinal anaesthesia: results from overview of randomised trials. *BMJ* 2000; 321:1493-1497.
2. Schug SA, Scott DA, Payne J et al. Postoperative analgesia by continuous extradural infusion of ropivacaine after upper abdominal surgery. *Br J Anaesth* 1996; 76:487-491.
3. Rigg JRA, Jamrozik K, Myles PS et al. Epidural anaesthesia and analgesia and outcome of major surgery: a randomised trial. *Lancet* 2002; 359:1276-1282.
4. Ballantyne JC, Carr DB, DeFerranti S et al. The comparative effects of postoperative analgesic therapies on pulmonary outcome: cumulative meta-analysis of randomised controlled trials. *Anesth Analg* 1998; 86(3):598-612.
5. Liu S, Carpenter RL, Mackey DC et al. Effects of peri-operative analgesic technique on rate of recovery after colon surgery. *Anesthesiology* 1995; 83:757-765.
6. Jorgensen H, Wetterslev J, Moiniche S, Dahl JB. Epidural local anaesthetics versus opioid-based analgesic regimens on postoperative gastrointestinal paralysis, PONV and pain after abdominal surgery. *Cochrane Database of Systematic Reviews* 2000; 4:CD001893.

7. Steinbrook RA. Epidural anesthesia and gastrointestinal motility. *Anesth Analg* 1998; 86:830-836.
8. Loick HM, Schmidt C, van Aken H et al. High thoracic epidural anesthesia, but not clonidine, attenuates the perioperative stress response via sympatholysis and reduces the release of troponin T in patients undergoing coronary artery bypass grafting. *Anesth Analg* 1999; 88:701-709.
9. De Leon-Casasola OA, Lema MG, Karbella D, Harrison P. Post-operative myocardial ischaemia: epidural versus intravenous patient-controlled analgesia. *Reg Anesth* 1995; 20:105-112.
10. Beattie WS, Badner NH, Choi P. Epidural analgesia reduces postoperative myocardial infarction: a meta-analysis. *Anesth Analg* 2001; 93(4):853-858.
11. Wheatley RG, Schug SA, Watson D. Safety and efficacy of postoperative epidural analgesia. *Br J Anaesth* 2001; 87:47-61.
12. Wilmore DW, Kehlet H. Management of patients in fast track surgery. *BMJ* 2001; 322:473-476.
13. Brodner G, Pogatzki E, van Aken H et al. A multi-modal approach to control post-operative pathophysiology and rehabilitation in patients undergoing abdominothoracic esophagectomy. *Anesth Analg* 1998; 86(2):228-234.
14. Giebler RM, Scherer RU, Peters J. Incidence of neurologic complications related to thoracic epidural catheterisation. *Anesthesiology* 1997; 86(1):55-63.
15. Wulf H. Epidural anaesthesia and spinal haematoma. *Can J Anaesth* 1996; 43(12):1260-1271.
16. Tryba M. Epidural regional anaesthesia and low molecular heparin. *Anesthesiologie Intrensivmedizin Notfallmedizin Schmerztherapie* 1993; 28(3):177-181.
17. Wang LP, Hauerberg J, Schmidt JF. Incidence of spinal epidural abscess after epidural analgesia. *Anesthesiology* 1999; 91:1928-3196.
18. Moiniche S, Hjortson C, Blemmer T et al. Blood pressure and heart rate during orthostatic stress and walking with continuous post-operative thoracic epidural bupivacaine/morphine. *Acta Anaesth Scand* 1993; 37(1):65-69.
19. Baird M, Schug SA. Safety aspects of postoperative pain relief. *Pain Digest* 1996; 6:219-225.

# A Reappraisal of Metoclopramide

E. J. MURPHY, MB, BS, FANZCA
Consultant Anaesthetist, Department of Anaesthesia, Royal Adelaide Hospital, Adelaide, South Australia

*Dr Murphy has specific interests in cardiovascular physiology and pharmacology, anaesthesia equipment, and monitoring.*

**SUMMARY**

Metoclopramide is still the most widely prescribed anti-emetic in Australia, despite the lack of evidence supporting its use at the recommended doses. It has not been convincingly shown to decrease postoperative ileus after gut surgery, and it may cause dystonia in susceptible patients by virtue of its antidopaminergic properties. Has the time come to discard this drug from our formulary, or is it merely misunderstood?

If we consider the available data on metoclopramide and its effects on dopaminergic and serotonergic neurotransmission, then the latter is probably correct. Much of the early clinical data on the drug was attributed to incorrect mechanisms, yet has been retained in textbooks and prescribing recommendations, and we probably should reconsider our use of metoclopramide in light of this newer data.

The following discourse will review the known mechanisms of action of the drug, discount some early theories, highlight appropriate uses and suggest novel benefits, including myocardial protection.

**BACKGROUND**

Metoclopramide (MCP) was developed experimentally in the early 1960s as a dopamine antagonist drug, with clinical use first described by Justin-Besancon and associates in 1964. The drug is produced by the addition of a 5-chloro and 2-methoxy aryl group to the benzene ring in procainamide, which itself stems from the local anaesthetic, procaine.[1] More widespread use of metoclopramide commenced around 1968 in Europe, where it was used to treat nausea and vomiting in pregnancy. However, it was not studied in America until around 1975, when interest arose regarding its role as a dopaminergic antagonist in the gastrointestinal tract, and in the treatment of smooth muscle disorders.

By 1983, four distinct mechanisms had been postulated to explain the anti-emetic and prokinetic effects of metoclopramide:

1. Direct activation of intramural cholinergic neurons was surmised in the earliest studies, as the gut effects of 10 mg MCP persisted despite vagotomy, but could be abolished by 10 mcg/kg atropine. These effects also depended upon intrinsic stores of acetylcholine, unlike the case with conventional cholinergic compounds.[1]
2. Central nervous system mediated antiemesis was attributed to central nervous system dopamine blockade, as therapeutic success did not always correlate with gastric emptying, and vomiting produced by apomorphine acting on the chemoreceptor trigger zone could be reversed with MCP, as well as the phenothiazines.[1]

3. Promotion of gastrokinesis by dopaminergic blockade had been assumed after dopamine (DA) was presented in 1975 as being a significant inhibitory neurotransmitter in the autonomic nervous system, with specific receptors identified in the GIT, renal, and coronary circulations.[2] The oesophageal and gastric effects of MCP were able to be reversed to some extent by the administration of levodopa, a dopamine precursor, lending support to this mechanism as the cause of increased gastrokinesis.[1]
4. Direct effects on smooth muscle were suggested by the persistence of MCP induced lower oesophageal sphincter contraction in the opossum, despite treatment with either tetrodotoxin (a sodium channel blocker) or atropine. Large doses of MCP were used in this study however, and a separate experiment using isolated gastric guinea pig smooth muscle did show MCP effect reversal with tetrodotoxin and hexamethonium, implicating postganglionic acetylcholine release as the final pathway.[1]

None of these above mechanisms were shown to be correct.

## CLINICAL APPLICATIONS

The pharmacological properties of MCP are now much better understood, and many effects aside from gastric prokinesis and anti-emesis have been identified. These are probably unrecognised by routine prescribers, and arise from a molecular configuration which allows for binding with many receptor subtypes. MCP has most affinity for the dopamine (DA)2 and serotonin (5HT)2 receptor subtypes, where it acts as an antagonist, but is also a DA1, alpha-2 and 5HT3 antagonist, and 5HT1 and 5HT4 partial agonist.[3]

The clinical utility of these effects in anaesthetic practise may be considered under the following headings:

### Propofol pain prevention

Pain on injection of propofol is a common problem, affecting around 70% of patients who are not given any preventative.[4] The mechanism which causes this pain remains unknown. However, the symptoms may be decreased by the addition of various drugs to either the propofol mixture itself or into the cannulated vein prior to injection. The drug traditionally used for this is lignocaine, although others have been trialled. An Australian study compared the addition of either 20 mg MCP or 10 mg lignocaine to the propofol mixture in 100 patients undergoing minor surgery. It found no difference in propofol pain scores between the groups, although it must be noted that the lignocaine dose was relatively low.[5] A later systematic review covered data from 6264 patients and analysed treatment with the following agents: lignocaine, opioids, MCP and heating or cooling the mixture. Results for different treatment modalities were compared by assessment of number needed to treat (NNT), which refers to the number of patients exposed to the active intervention before one will gain benefit. The best result was obtained when lignocaine 60mg was given into a vein occluded by tourniquet prior to propofol injection; this gave a NNT of 1.6. This increased to 2.4 if lignocaine 20 mg was mixed with propofol, a more common regime, and was in fact bettered by 10 mg MCP using a tourniquet technique (NNT 2.2). MCP was less effective if given without a tourniquet, where the NNT increased to 2.7, but it is obvious that both mechanisms are effective. Temperature adjustment or opiate use was minimally helpful, excepting pethidine with a tourniquet technique, which had a

NNT of 1.9, possibly because of a direct local anaesthetic effect. Other agents were noted to have been trialled in this systematic review, including ondansetron, droperidol, ketamine, aspirin, ketorolac and others. However, no meaningful conclusions could be drawn from their respective studies.[4]

**Gastric prokinetic effects**

These effects appear to be mediated via gastrointestinal 5HT4 receptors,[6, 7, 8] akin to the method of action of the drug cisapride. They promote an increase in lower oesophageal sphincter pressure tone, along with accelerated gastric emptying through both more frequent and more intense antral and duodenal contractions.[1] Cisapride has been shown to be seven times more potent than MCP in promoting rat gastric emptying, despite the fact that both drugs are 5HT4 agonists.[9] One possible reason why MCP is less effective is that 5HT1 activation with sumatriptan, a specific 5HT1 agonist, mediates suppression of gastric emptying, and may be mimicked by the weaker 5HT1 effects of MCP.[10] The side effect which led to cisapride being withdrawn in some countries (QT prolongation and ventricular dysrhythmias) has been experimentally related to class III anti-arrhythmic effects of the cisapride molecule which also blocks voltage dependant potassium channels, and not the 5HT4 receptor.[11]

The oesophageal and gastrokinetic effects of MCP are blocked with concurrent use of atropine 10 mcg/kg owing to the involvement of intramural cholinergic modulation in the prokinetic pathway.[6] The decrease in gastric emptying seen with opiates is a separate issue, and occurs due to increased resting tone within the gastrointestinal tract. It is likely that the increased lower oesophageal sphincter tone due to MCP will persist in this setting, although it does not appear to have been assessed specifically.

The time to onset of gastrointestinal effects, if measured using oesophageal sphincter tone as an end point, is 1-3 minutes after IV injection, 10-15 minutes after IM injection, and 30-60 minutes after oral ingestion. (Peak plasma levels measured after a 20 mg MCP dose were around 84 ng/ml po or 221 ng/ml iv.)[12]

MCP has been used effectively to treat delayed gastric emptying in patients with diabetic gastroparesis. One study using isotope labeled egg salad sandwiches showed an accelerated gastric clearance at sixty and ninety minutes following dosage with MCP, such that the residual volume dropped to around two thirds of that in the untreated group. However, the emptying rate remained less than that of normal subjects.[1]

Post surgical ileus has been treated with MCP with varying results. Chronic gastric emptying abnormalities following vagotomy or gastric resection may improve significantly, although acute postoperative ileus has not shown accelerated resolution, despite some degree of symptomatic improvement.[1, 13] However, it should be remembered that increased gut wall tension without distal outflow may be undesirable where anastomoses in the upper GIT have been newly created, as the risk of leakage or rupture is enhanced.

**Anti-emesis**

Despite routine use as a first line anti-emetic, there remains much controversy about this role. MCP demonstrates a proven anti-emetic effect in adults only at higher doses (1-4 mg/kg) as used during some chemotherapy regimes. This has been related to its low affinity 5HT3 receptor antagonism, and it appears that the drugs antidopaminergic properties are irrelevant here, except in producing dose limiting side effects, including

behavioural depression.[14] A Cochrane Review of the literature from 1966 to 1998, considering only randomized controlled trials, found data on eighteen different MCP regimes for postoperative nausea and vomiting (PONV) prophylaxis. These included an adult dosage range of either 5-30 mg or 0.25-0.5 mg/kg MCP, yet at all dosage levels the outcome was the same: there was no statistical difference between metoclopramide and placebo in the prevention of PONV. In children, however, MCP was found to be more effective than placebo at preventing vomiting (most studies did not assess nausea) with a number needed to treat (NNT) of 7.9 (95% confidence interval 4.6, 28). The authors conclusions included a recommendation for randomized dose finding studies, as there was a suggestion that higher MCP doses in adults could prevent vomiting, although this did not reach clinical significance in the studies considered. Interestingly, there were no significant differences between MCP and placebo in the incidence of all adverse reactions, despite a predicted 10% incidence overall with MCP.[15]

**Phaeochromocytoma detection**

A single paper suggested that the injection of 10 mg MCP intravenously could be a useful provocative test to screen for suspected phaeochromocytoma. A review of three such cases showed a mean systolic blood pressure rise of 15 mmHg, and an increase in serum arginine vasopressin levels following injection. This response returned to normal after tumour excision.[16] Most product information however, lists phaeochromocytoma as a contraindication, due to potential and unpredictable hormonal responses. Perhaps more than three cases need to be assessed before this procedure can be recommended. A more sensible use of this knowledge would be to consider the diagnosis of phaeochromocytoma in patients who develop hypertension instead of the expected drop in blood pressure following MCP injection, particularly when there is an associated history of paroxysmal hypertension.

**Hormone release**

Documented effects on other hormone levels include: a transient increase in serum aldosterone levels mediated by action on the 5HT4 receptor,[7, 17, 18] a transient increase in serum cortisol levels,[17,19] particularly with higher doses, and increased prolactin release.[1] A dose of 10 mg MCP tds has been advocated in obstetric practice to stimulate lactation by increasing prolactin levels, but the other hormonal effects described have not yet been shown to have any perioperative application.

**Migraine treatment**

Recent advances in migraine pathophysiology have implicated the release of a neuropeptide, calcitonin gene-related peptide (CGRP), by trigeminal nerve vascular afferents in the initiation and maintenance of migraine pain.[20] The resultant inflammation and accompanying meningeal blood vessel dilatation respond to treatment with drugs such as sumatriptan, a selective 5HT1 agonist. This newer class of antimigraine drugs, known as triptans, act via 5HT1b/1d receptors to cause both selective cranial blood vessel constriction and repression of CGRP secretion and transcription.[20, 21]

This knowledge has provided a basis for further research in migraine treatment, yet it is worth noting that an older treatment modality using aspirin, when combined with MCP has been shown to be similarly effective. A double blinded placebo controlled

trial in 1995, randomised 421 pts who fulfilled International Headache Society criteria for migraine to receive either aspirin 900 mg with MCP 10 mg, or sumatriptan 100 mg at headache onset. Similar improvement was found between both groups ($P=0.5$) when compared with placebo ($P<0.006$), using an outcome of a decrease in headache intensity; from severe or moderate to mild or no pain.[22] The mechanism for the first group was attributed to more rapid absorption of aspirin in nauseated migraine sufferers when given MCP, although the actions of MCP as a 5HT1 agonist are probably significant, given the specific 5HT1 actions on migraine pathology. 5HT2 mediated vasodilatation, which MCP also produces, is unlikely to interfere with this specific mechanism and is supported by the experimental finding that 5HT2 blockade by ketanserin does not interfere with sumitriptan induced contractions in isolated middle meningeal artery specimens.[23]

**Myocardial protection**

Clinical data in humans has shown that normal epicardial coronary arteries dilate in response to serotonin, probably by activation of endothelial 5HT1 receptors, which initiate the release of endothelium derived relaxing factors. On the other hand, diseased atherosclerotic coronary arteries constrict due to activation of 5HT2 receptors on smooth muscle cells.[24] In both cases, MCP will favour vasodilatation.

Human and animal laboratory studies go as far as to suggest a role for MCP in prevention of acute coronary artery thrombosis. Atherosclerotic plaque rupture is the usual initiating event in coronary thrombosis,[25] and may be precipitated by mechanical stress from blood pressure changes during intubation in susceptible individuals. This leads to clotting activation and platelet aggregation by the extrinsic pathway. The platelet response, which includes serotonin release, promotes arteriospasm and further platelet aggregation, but has been shown to be attenuated by both 5HT2 blockade and thromboxane A2 inhibition.[26] These effects are synergistic and the response has been shown to be useful in human radial and internal mammary artery specimens,[26] as well as in animal models of coronary artery occlusion.[27, 28]

The human studies have been directed specifically at preventing vasospasm related to platelet activity in the setting of coronary artery grafting, whereas animal studies have assessed experimental lesions and provocation of coronary thrombosis. A canine model of coronary artery thrombosis assessed cyclic flow variation (CFV) to demonstrate sequential thrombosis and embolisation in a critically stenosed circumflex coronary artery, with a controlled arterial lesion at the site of stenosis. The findings in this case were that the combination of aspirin 0.1 mg/kg and MCP 0.3 mg/kg lead to a drop in CFV from $6.7\pm0.5$ to $0.8\pm0.4$ cycles per hour ($P<0.01$), indicating a significant antithrombotic effect. Neither agent alone produced significant results, indicating synergism between the two.[28] A separate experiment involving a rat model of critical arterial stenosis, given hypotonic saline with serotonin to initiate thrombosis, showed significant antithrombotic effect from MCP, as well as synergism with aspirin.[27]

The human experimental data showed significant antispasmodic effect on both types of arterial specimen from 5HT2 blockade, particularly after removal of endothelium.[26] These results suggest that the metoclopramide alone, or in combination with aspirin, may decrease the risk of perioperative myocardial infarction caused by coronary plaque rupture. A definitive study in humans remains to be carried out. The potential risks of impaired coagulation during surgery when metoclopramide has been given do

not appear to have been investigated, but there are no reported problems of this nature in the literature.

**Pain relief**

A recent study investigated the possibility that MCP enhances analgesic efficacy. Fifty-six patients were randomised to receive either metoclopramide or placebo prior to undergoing laparoscopic tubal ligation, then assessed postoperatively using a numeric pain rating scale. Opioid requirements were also documented. The pain scores were similar in both groups at all three time intervals examined, but the morphine requirements were significantly higher in the placebo group ($P=0.031$). The data allowed a conclusion that the addition of metoclopramide was opiate sparing, but did not assess for direct analgesic efficacy.[29] The mechanism of opiate sparing was not determined although 5HT1 and 5HT4 receptor activation may be important if we consider that the analgesic drug tramadol acts partially by enhancing serotonin release,[30] but can be rendered less efficacious by concurrent strong 5HT3 blockade with ondansetron.[31, 32]

**ADVERSE REACTIONS**

Aside from these useful applications, MCP may also cause a number of unwanted side effects. Those which are not widely known will be considered, as well as the symptoms commonly known as dystonias, movement disorders or extrapyramidal reactions, which are infrequent but striking responses to the drug.

**Plasma cholinesterase inhibition**

MCP may prolong the effects of suxamethomium[12] and does prolong mivacurium induced neuromuscular block, by decreasing plasma cholinesterase activity. A recent study randomized 45 patients to receive either placebo, MCP 10 mg or MCP 20 mg prior to induction of a relaxant anaesthetic including mivacurium. Neuromuscular block was monitored with a force transducer using train of four stimulation and found the time to spontaneous recovery of 90% twitch height was increased from 32±9 min to 44±15 min following 10 mg MCP, or 57±10 min after 20 mg MCP. A slight but significant decrease in plasma cholinesterase activity was found in the 20 mg group.[33] This may be significant if reversal agents are to be avoided on the assumption of spontaneous reversal with mivacurium within the above time frames. The finding warrants confirmation of reversal with a nerve stimulator where MCP has been given.

**Raised intracranial pressure (ICP)**

Although the mechanism was not determined, a case report on one head injured patient with diffuse axonal injury, showed a reproducible increase in ICP from a 15-20 mmHg baseline up to 39 mmHg, following 10 mg IV MCP. The patient was given the same dose during transcranial doppler studies the next day; this gave rise to a similar increase in ICP and an associated blood velocity increase from 122 cm/s to 150 cm/s in the middle cerebral artery.[34] A possible explanation is that rapid injection of MCP does cause transient systemic hypotension due to vasodilatation,[35] the intracranial component of which would compress the intracranial contents further, causing the observed rise in ICP.

**Reversal of "renal dose" dopamine**

It has been questioned whether the use of MCP is appropriate in patients who are being treated with a so called renal dose of dopamine (around 3 mcg/kg/min), as the consequent dopamine blockade may reverse the desired effect of renal vascular dilatation. Such reversal has been shown to occur in animal studies, with a 44% drop in dopamine induced vasodilatation after MCP 1 mg/kg, despite the drug having a relatively low affinity for the DA1 receptor responsible.[36] A study of stable patients in intensive care however, measured renal blood flow and creatinine clearance during dopamine infusion at the rate above, and showed no detrimental effect from 10 mg MCP.[37] The relevance of this remains questionable, as no convincing data has shown dopamine to be effective for renal function preservation.[38]

**Serotonion syndrome with selective serotonin reuptake inhibitors**

Two patients were reported to have developed a serotonin syndrome after being given MCP whilst on the SSRI drugs, sertraline or venlafaxine. The symptoms were reproducible with re-administration in both cases and included agitation, dysarthria, movement disorder and diaphoresis. The symptoms met criteria for serotonin syndrome, and were thought to be due to a pharmacodynamic interaction, although a single drug effect from MCP could not be ruled out.[39] Concomitant use with SSRIs is not a listed contraindication to MCP use, but it would be prudent to avoid such use if alternatives exist.

**Acute movement disorders**

These are sometimes known as extrapyramidal symptoms; however, this terminology encompasses a range of descending spinal pathways and implies an anatomical basis for the problem. These disorders may occur with most antidopaminergic drugs, and are divided into three potential syndromes:[1]

1. A range of dystonias affecting primarily the facial muscles and producing trismus, torticollis, facial spasms, opisthotonos, or oculogyric crises;
2. A parkinsonian like syndrome with agitation and impaired mobility; or,
3. Akithisia, a compulsion toward constant motor activity.

(Tardive dyskinesia is associated only with chronic administration and will not be considered specifically).

The largest prospective study to date assessing first prescriptions of all forms of metoclopramide found a total incidence of movement disorders of 0.5% in 2557 patients.[40] However, perioperative data from a meta-analysis of randomised controlled trials found no acute movement disorder in 220 children given parenteral MCP doses from 0.15 to 0.5 mg/kg, and one isolated episode in an adult given 20 mg intravenously, from a total of 1177 subjects given either MCP or placebo.[15]

It is interesting to note that, despite the above results, both the MIMS[12] and *Australian Medicines Handbook*[41] do not recommend MCP use in children due to an increased risk of movement disorders. Some support for this notion is given by general practice data showing that high risk patients for these side effects include young children and females aged 12-19 years.[42] There have also been reports suggesting an increased risk in patients with the acquired immune deficiency syndrome, where it may be associated with a disease related decrease in central nervous system dopamine activity.[40]

Acute movement disorders may be unpleasant for the patient, but are not generally

dangerous. Fatal outcome has been reported due to pharyngeal muscle involvement, although it is extremely rare. It is recommended that these reactions be treated with parenteral benztropine 1-2 mg.[41]

Of related interest, a pH neutral form of MCP (ph 6.9-7 instead of 2.5-3.5) has been developed, and initial testing shows decreased DA2 affinity and increased 5HT3 affinity, suggesting the possibility of significantly improving the safety and efficacy of the drug.[15]

## CONCLUSION

It has now become obvious that a drug originally presented as a dopamine antagonist, in fact has no useful effects attributable to this mechanism, except perhaps prolactin release. It is probably better considered as a serotonin receptor modulating drug which has dose limiting anti-dopaminergic effects.

If one considers that the recommended uses of MCP are limited to treatment of nausea and vomiting, gastric stasis (post-surgical and diabetic) and difficult small intestinal intubation, with accepted indications of gastro-oesophageal reflux disease and gastrointestinal radiology, and that recommended adult doses are 10 mg 6-8 hourly for nausea and vomiting,[41] then it is surprising that it continues to be prescribed postoperatively at all. The likelihood of a useful response with these indications is small. The use of MCP in anaesthesia may be better confined to preoperative and intraoperative use, where it has a number of known and potential benefits including minimisation of propofol injection pain, increased lower oesophageal sphincter tone, enhanced analgesia, myocardial protection, and enhanced gastric emptying.

## REFERENCES

1. Albibi R, McCallum R. Metoclopramide: Pharmacology and clinical application. *Ann Int Med* 1983; 98:86-95.
2. Thorner MO. Dopamine is an important neurotransmitter in the autonomic nervous system. *Lancet* 1975; 1:662-665.
3. Murphy EJ. Intra-arterial injection of metoclopramide, midazolam, propofol and pethidine. *Anaesth Intens Care* 2002; 30(3):367-369.
4. Picard P, Tramer MR. Prevention of pain on injection with propofol: a quantitative systematic review. *Anesth Analg* 2000; 90:963-969.
5. Mecklem DWJ. Propofol injection pain: comparing the addition of lignocaine or metoclopramide. *Anaesth Intens Care* 1994; 22:568-570.
6. Goldhill J, Porquet MF, Angel I. Post-synaptic 5-HT4 receptor modulation of tachykinergic excitation of rat oesophageal tunica muscularis mucosae. *Europ J Pharmacol* 1997; 323:229-233.
7. Sommers DK, Snyman JR, van Wyk M. Effects of metoclopramide and tropisetron on aldosterone secretion possibly due to agonism and antagonism at the 5-HT4 receptor. *Eur J Clin Pharmacol* 1996; 50:371-373.
8. van Wyk M, Sommers DK, Moncrieff J. Effects of enhancement and antagonism of 5-hydroxytryptamine activity on the influence of metoclopramide on gastric emptying. *Digestion* 1993; 54:40-43.
9. Megens AA, Awouters FH, Niemegeers CJ. General pharmacology of the four gastrointestinal motility stimulants bethanechol, metoclopramide, trimebutine, and cisapride. *Arzneimittel-Forschung-Drug Research* 1991; 41:631-634.
10. Houghton LA, Fowler P, Keene ON, Read NW. Effect of sumatriptan, a new selective 5HT1-like agonist, on liquid gastric emptying in man. *Alim Pharmacol Therapeut* 1992; 6:685-691.
11. Tonini M, De Ponti F, Di Nucci A, Crema F. Review article: Cardiac adverse effects of gastrointestinal prokinetics. *Alim Pharmacol Therapeut* 1999; 13:1585-1591.
12. Maxolon (DermaTech Laboratories): product monograph in MIMS Annual 2000; 3-350–3-351.
13. Cheape JD. Wexner SD. James K. Jagelman DG. Does metoclopramide reduce the length of ileus after colorectal surgery? A prospective randomized trial. *Dis Colon Rectum* 1991; 34:437-441.

14. Tyers MB. Pharmacology and preclinical anti-emetic properties of ondansetron. *Seminars Oncol* 1992; 19(4 Suppl 10):1-8.
15. Henzi I, Walder B, Tramer MR. Metoclopramide in the prevention of postoperative nausea and vomiting: a quantitative systematic review of randomized, placebo-controlled studies. *Br J Anaesth* 1999; 8:761-771.
16. Chiodera P, Volpi R, Coiro V. Hyperresponsiveness of arginine vasopressin to metoclopramide in patients with pheochromocytoma. *Arch Intern Med* 1999; 159(21):2601-2603.
17. Staessen J, Fiocchi R, Bouillon R, Fagard R, Lijnen P, Lissens W, Amery A. Differential responses of plasma aldosterone, cortisol and adrenocorticotropin to two dopamine receptor antagonists. *Method Find Exp Clin Pharmacol* 1985; 7:523-527.
18. Rizzi CA, Mierau J, Ladinsky H. Regulation of plasma aldosterone levels by metoclopramide: a reappraisal of its mechanism from dopaminergic antagonism to serotonergic agonism. *Neuropharmacology* 1997; 36:763-768.
19. Coiro V, Volpi R, Capretti L et al. 5-HT1-, but not 5-HT2-serotonergic, M1-, M2-muscarinic cholinergic or dopaminergic receptors mediate the ACTH/cortisol response to metoclopramide in man. *Horm Res* 1990; 33:233-238.
20. Durham P, Russo A. New insights into the molecular actions of serotonergic antimigraine drugs. *Pharmacol Therapeut* 2002; 94:77.
21. Akerman S, Williamson D, Kaube H, Goadsby P. The effect of anti-migraine compounds on nitric oxide-induced dilatation of dural meningeal vessels. *Europ J Pharmacol* 2002; 452:223.
22. Tfelt-Hansen P, Henry P, Mulder LJ, Scheldewaert RG, Schoenen J, Chazot G. The effectiveness of combined oral lysine acetylsalicylate and metoclopramide compared with oral sumatriptan for migraine. *Lancet* 1995; 346:923-926.
23. Van den Broek RW, MaassenVanDenBrink A, Mulder PG et al. Comparison of contractile responses to donitriptan and sumatriptan in the human middle meningeal and coronary arteries. *Europ J Pharmacol* 2002; 443:125-132.
24. MacIntyre PD, Bhargava B, Hogg KJ, Gemmill JD, Hillis WS. Effect of subcutaneous sumatriptan, a selective 5HT1 agonist, on the systemic, pulmonary, and coronary circulation. *Circulation* 1993; 87:401-405.
25. Gudimetia S, Kusomoto FM. Coronary Artery Disease. In: Kusomoto FM, ed. Cardiovascular Pathophysiology. Fence Creek Publishing, LLC, Madison, Connecticut, 1999; 91-137.
26. Chardigny CI, Van der Perre K, Simonet S et al. Platelets and prostacyclin in arterial bypasses: implications for coronary artery surgery. *Ann Thorac Surg* 2000; 69:513-519.
27. Mobarok Ali AT, al-Humayyd MS. Enhancement of antithrombotic effect of aspirin by metoclopramide in rats. *Thromb Res* 1991; 61:361-367.
28. Duval N, Grosset A, O'Connor SE. Combination of aspirin and metoclopramide produces a synergistic antithrombotic effect in a canine model of coronary artery thrombosis. *Fundam Clin Pharmacol* 1997; 11:57-62.
29. Gibbs RD, Movinsky BA, Pellegrini J, Vacchiano CA. The morphine-sparing effect of metoclopramide on postoperative laparoscopic tubal ligation patients. *AANA Journal* 2002; 70:27-32.
30. Reimann W, Schneider F. Induction of 5-hydroxytryptamine release by tramadol, fenfluramine and reserpine. *Eur J Pharmacol* 1998; 349:199-203.
31. De Witte JL, Schoenmaekers B, Sessler DI, Deloof T. The analgesic efficacy of tramadol is impaired by concurrent administration of ondansetron. *Anesth Analg* 2001; 92:1319-1321.
32. Arcioni R, della Rocca M, Romano S, Romano R, Pietropaoli P, Gasparetto A. Ondansetron inhibits the analgesic effects of tramadol: a possible 5-HT(3) spinal receptor involvement in acute pain in humans. *Anesth Analg* 2002; 94:1553-1557.
33. Motamed C, Combes X, Dhonneur G, Duvaldestin P. Effect of metoclopramide on mivacurium-induced neuromuscular block. *Acta Anaesthes Scand* 2002; 46:214-216.
34. Deehan S, Dobb GJ. Metoclopramide-induced raised intracranial pressure after head injury. *J Neurosurg Anesth* 2002; 14:157-160.
35. Magnifico F, Pierangeli G, Barletta G et al. The cardiovascular effects of metoclopramide in multiple system atrophy and pure autonomic failure. *Clin Autonomic Res* 2001; 11:163-168.
36. Hahn RA, Wardell JR Jr. Antagonism of the renal vasodilator activity of dopamine by metoclopramide. *Naunyn Schmiedebergs Arch Pharmacol* 1980; 314:177-182.
37. Munn J, Tooley M, Bolsin S, Hronek I, Lowson S, Willcox J. Effect of metoclopramide on renal vascular resistance index and renal function in patients receiving a low-dose dopamine infusion. *Brit J Anaesth* 1993; 71:379-382.

38. Kellum JA. Use of dopamine in acute renal failure: a meta-analysis. *Crit Care Med* 2001; 29:1526-1531.
39. Fisher AA, Davis MW. Serotonin syndrome caused by selective serotonin reuptake-inhibitors-metoclopramide interaction. *Ann Pharmacol* 2002; 36:67-71.
40. van der Kleij FGH, de Vries PAM, Stassen PM, Sprenger HG, Gans ROB. Acute dystonia due to metoclopramide: increased risk in AIDS. *Arch Intern Med* 2002; 162; 358-359.
41. Metoclopramide prescribing information. Drugs for Nausea and Vomiting. In: Australian Medicines Handbook. Australian Medicines Handbook, Adelaide, South Australia, 2000; 12-22–12-23.
42. Bateman DN, Rawlins MD, Simpson JM. Extrapyramidal reactions with metoclopramide. *Brit Med J* (Clin Res Ed) 1985; 291:930-932.

# Gastro-oesophageal Reflux: What's the Big Deal?

DAVID J. SANDEMAN, MB, BS(Hons), BSc, FANZCA
Staff Specialist Prince of Wales Hospital and Sydney Children's Hospitals, Conjoint Lecturer University of NSW

*David Sandeman is a Staff Specialist at Prince of Wales and Sydney Children's Hospitals at Randwick. David's main interests are paediatric anaesthesia, risk management and evidence based medicine. David presented this paper at a Sydney CME meeting in September 2002.*

Gastro-oesophageal reflux (GOR) remains a concern for the anaesthetist because aspiration pneumonitis carries the potential for significant morbidity and mortality. Reflux is probably common, whilst aspiration during the peri-operative period is rare, and aspiration pneumonitis is exceedingly rare.

**Background**

General anaesthesia may predispose patients to aspiration of gastroesophageal contents because of depression of protective reflexes during loss of consciousness.[1] Some patients may be at increased risk of pulmonary aspiration because of retention of gastric contents caused by pain, inadequate starvation, or gastrointestinal pathology resulting in reduced gastric emptying and gastroesophageal reflux. The prevalence of aspiration in elective surgery is about one in three thousand anaesthetics and aspiration causes one death per 200,000 elective anaesthetics.[1, 2, 3]

**History**

The effects of aspiration of food and drink had been known since the time of Hippocrates, but it was John Hunter who performed the first scientific experiments investigating the pathophysiology of aspiration in 1781. The first documented death related to anaesthesia was most likely a result of the liquid administered during unconsciousness. In this case, Sir James Simpson identified pulmonary aspiration of the brandy and water that Hannah Greener, a 15-year-old girl, was given during chloroform anaesthesia since "her lips, which had been previously of good colour, became suddenly blanched, and sputtered slightly at the mouth as one with epilepsy."

John Snow's book "On Chloroform and other Anaesthetics" was published in 1858.[4] He comments regarding regurgitation and demonstrates concern about vomiting rather than aspiration:

"The only direction which is usually requisite to give beforehand, to the patient who is to inhale chloroform, is to avoid taking a meal previous to the inhalation; for chloroform is very apt to cause vomiting, if inhaled whilst there is a quantity of food in the stomach. The sickness is not attended with any danger, but it constitutes an unpleasantness and inconvenience, which it is desirable to avoid. The best time of all for an operation under chloroform is before breakfast, but the customs and

arrangements of this country do not often admit of that time being chosen, and it is unadvisable to make the patient fast beyond his usual hour. The most usual time for vomiting to commence is when the inhalation has been discontinued, and the effects of the chloroform are passing off. In many cases, it occurs before the patient has become quite conscious, and he does not know that it has occurred unless told. In a few cases, especially where there is a good deal of food in the stomach, the vomiting comes on before the operation is finished, or even before it is commenced. When vomiting comes on during an operation, it is apt to interfere with the inhalation, and it is difficult sometimes to prevent the patient from waking."

The dangers of regurgitation and aspiration seem to have been ignored at this early stage. GOR and vomiting were considered messy, a nuisance and interfering with the administration of anaesthesia. In the past, teaching on anaesthesia was usually in surgical texts. One such example was "A Manual of Surgical Treatment" published in 1912;[5] in this text recommended preoperative preparation included, "A purge of castor oil, compound liquorice powder, calomel, colocynth or compound rhubarb pill the night before operation and, if necessary, an enema on the morning of surgery." With regards to diet, the author wrote, "It is important that the stomach should be empty before the inhalation commences; but starvation may be carried too far, especially in the feeble." He added, "The best time for operating is the early morning, in which case no food need be given after supper on the previous evening. If the operation be fixed for the afternoon, a light breakfast should be taken not later than 8 am, and a cup of hot broth or beef tea, or even hot water alone, should be given not less than three hours before the actual time of operation."

The publication in 1946 of Mendelson's paper entitled "The aspiration of stomach contents into the lungs during obstetric anesthesia", began a new era in our understanding of regurgitation and aspiration.[6] Mendelson's data included 44016 unfasted parturients who received ether anaesthesia for delivery and operations associated with delivery. There were 66 cases of aspiration (0.15%) and two deaths from complete airway obstruction caused by large food particles. No patients died from aspiration pneumonitis despite cricoid pressure (Sellick's manoeuvre) not being utilised at this stage. Mendelson established the role of low pH of gastric secretions as an important mechanism involved in lung injury associated with aspiration. He described the changes that occurred as a result of gastric aspiration as well as the clinical symptoms. Based on these findings, recommendations for prevention and treatment of aspiration of gastric contents were proposed. These included: "withholding oral feeding during labour and substituting parental administration, wider use of local anaesthesia where feasible and alkalinisation of and emptying the stomach."

In 1961 there was another important publication in the *Lancet* by Brian Sellick from the Middlesex Hospital in London.[7] Sellick wrote, "When the contents of the stomach or oesophagus gain access to the air passages during anaesthesia, the consequences are disastrous. In spite of modern anaesthetic techniques, or sometimes regrettably because of them, regurgitation is still a considerable hazard during induction of anaesthesia, particularly for operative obstetrics and emergency general surgery. Cricoid pressure must be exerted by an assistant. Before induction, the cricoid is palpated and lightly held between the thumb and second finger; as anaesthesia begins, pressure is exerted on the cricoid cartilage mainly by the index finger. Even a conscious patient can tolerate moderate pressure without discomfort but as soon as consciousness is lost, firm pressure can be applied without obstruction of the patient's airway.

Pressure is maintained until intubation and inflation of the cuff of the endotracheal tube is complete."

Sellick was aware that cricoid pressure could not be relied on totally.

**Gastro-oesophageal reflux**

The stomach contracts every 20 seconds with pressure increases to as much as 50 cm $H_2O$; it can accommodate up to1500 ml. Fasted patients may have 200 ml of gastric fluid. The lower oesophageal sphincter (LES) is competent to about 30 cm $H_2O$ if normal. The difference between LES pressure and gastric pressure is the "barrier pressure" (BaP). Anaesthesia tends to decrease BaP and reduce protective reflexes, thus increasing the risk of GOR.[1, 8]

Various drugs decrease LES tone, including anticholinergics, thiopentone, opioids and volatile anaesthetics. Drugs increasing LES tone include antiemetics, cholinergics, succinylcholine and antacids, whilst non-depolarising muscle relaxants and histamine 2 receptor ($H_2R$) blockers have no effect on the LES tone.[1] Contrary to common belief, there is poor correlation between Body Mass Index (BMI), smoking, duration of fasting, alcohol consumption, gastric pH and gastric volume with GOR. However, pregnancy, ascites, obesity, gastrointestinal distension, acute pain, opioids, anxiety, trauma and active labour, delay gastric emptying increasing the probability of GOR.[1,3,8]

**Morbidity and mortality with aspiration**

The AIMS study published in 1999, included 133 cases of aspiration in 5000 incident reports.[9] Most patients had at least one predisposing factor for regurgitation, vomiting or aspiration. There were five deaths, all in high-risk patients (two ASA 3, two ASA 4 and one ASA 5). These results compare with other series that show a mortality rate of about 4% from aspiration, and that death from aspiration occurs in high-risk patients.[1-3] The airways present at the time of aspiration were ETT,[8] Hudson mask,[5] LMA[27] and Face Mask[9]. Amongst the patients aspirating, 20% were under 14 years; aspiration in the 0 to 9 year age group was three times that in the 20 to 49 year old group.[9] Most patients have prolonged hospital stay, or ICU admission with varying levels of respiratory support.

**Methods to minimise reflux and aspiration**

*Minimise gastric content*

Gastric content may be reduced by appropriate fasting, pre-operative emptying of the stomach (in selected cases using a gastric tube), pro-kinetic agents, and by inhibiting gastric secretion. Metoclopramide is an effective pro-kinetic agent, which acts centrally and peripherally to increase gastric peristalsis, increase LES tone and relax the pylorus.[8] Intravenous metoclopramide (10-20 mg) decreases gastric volume, without any effect on gastric pH within 30 minutes; the effect persists for about three hours.[8, 10]

*Minimising GOR*

The hazards of GOR may be minimised by reducing gastric content, using a pro-kinetic agent or cricoid pressure (CP). There is debate regarding the efficacy of CP arising from a lack of convincing evidence for a reduced incidence of aspiration or mortality. However, this reflects both relatively few studies and the infrequency of aspiration, rather than ineffectiveness of CP. Effective CP does reduce passive

regurgitation of gastric contents to the laryngopharynx and should therefore continue to be used in all cases where there is a significant risk of regurgitation or aspiration.

*Prevention of pulmonary aspiration*

Aspiration of regurgitated gastric material is nearly always prevented by CP and by securing the airway. Tracheal intubation with a cuffed tube offers most protection, but is imperfect. High volume/low-pressure cuffs are most effective for sealing the airway, but leaks can still occur around the cuff.[1] Awake endotracheal intubation may be indicated when there is high regurgitation risk combined with an anticipated difficult intubation.

*Hiatus hernia*

Hiatus Hernia (HH) can be viewed as a progressive disruption of the oesophago-gastric-junction. The LES tone is the major component of the anti-reflux barrier. Anything disrupting the LES will increase GOR. The larger the hiatus hernias, the greater the risk of GOR in both adults and children.[11] HH is common and estimates of prevalence vary from 10 to 80% of the adult population in North America. In most patients, HH is asymptomatic and there is poor correlation between HH and GOR. The safest approach is to treat all patients with symptomatic HH as at risk for GOR. Management of patients with asymptomatic HH is controversial. Patients with asymptomatic HH who do not have other risk factors for GOR, do not need any specific anti-reflux therapy.

**Pathogenesis of lung injury following aspiration**

Lung injury may result from particulate, acid and bacterial soiling of the lung. Extrapolation from rhesus monkey studies led to the now much-debated conclusion that gastric volume greater than 25 ml (or 0.4 ml/kg) and pH less than 2.5 worsens outcome from aspiration.[12] There are no data to show improved outcome after aspiration following the prophylactic use of antacids, $H_2R$ blockers, PPIs or prokinetics. Therefore, when considering that there is about one death per 200,000 anaesthetics from aspiration, it seems that the routine use of antacid/prokinetic therapies is not justified.[8]

**Prophylaxis against acid aspiration**

$H_2R$ blockers and PPIs are effective at reducing secretion of gastric acid. $H_2R$ blockers have been shown to be more effective compared with PPIs.[13] A single oral dose of ranitidine or famotidine three hours before surgery provides more effective volume and pH control of gastric secretion than omeprazole.

Particulate antacids should never be used because the particles can cause a pneumonitis.[14] Non-particulate antacids such as 30 ml of 0.3 M sodium citrate are effective at increasing gastric pH to greater than 2.5 in most patients within a few minutes and continue to be effective for up to three hours. Antacids have no significant effect on gastric volume.

Pro-kinetics may be useful for reducing gastric volume prior to induction.

**Management of aspiration once it has occurred**

Useful measures include bronchoscopy, suction, physiotherapy, specific antibiotics for established infection, supportive treatments including supplemental oxygen,

respiratory support, monitoring in a high dependency environment and bronchodilators, if indicated.[1] Broad-spectrum antibiotics and prophylactic steroids have not been shown to improve morbidity or mortality. Lavage should only be used if particulate material is aspirated. Lavage extends acidic fluid deeper into bronchial tree, worsening outcome.[8]

**Laryngeal mask airway**

The effect of the LMA on barrier pressure and LES tone is a subject of controversy. The LMA probably reduces BaP at the LES. Based on pH probe measurements, patients probably regurgitate to the lower oesophagus more often with spontaneous or positive pressure ventilation (PPV) using a LMA compared with a facemask.[15] The Classic LMA does not protect the airway if regurgitation to the proximal oesophagus occurs. The Pro-seal LMA may provide better airway protection.

In 1995, Brimacombe and Berry published a meta-analysis detailing 12,901 uses of the LMA.[16] There were three episodes of pulmonary aspiration giving a prevalence of 1 in 4000. Generally, cases had one or more predisposing factors. There was no permanent disability or death.

*LMA and Caesarean Section*

The LMA has been used safely for elective Caesarean section in one prospective study of 1067 cases.[17] In this study, all patients were fasted and given ranitidine and sodium citrate. A rapid sequence induction was performed with thiopentone and suxamethonium. Cricoid pressure was maintained until delivery, but was relaxed if insertion or ventilation was difficult. An effective airway was obtained in 1060 (99%) patients, 1051 (98%) at the first attempt and nine (1%) at the second or third attempt. There were no episodes of hypoxia ($SpO_2$ <90%), aspiration, regurgitation, laryngospasm, bronchospasm or gastric insufflation. In this study, the LMA was both safe and effective, even though such use is unorthodox. The report offers reassurance that, in the event of a failed intubation, the LMA makes an adequate second line airway.

*Proseal LMA (PLMA)*

*Does the PLMA airway prevent aspiration of regurgitated fluid?*

The PLMA has a drainage tube, which directs regurgitated fluid away from the laryngeal inlet. The drainage tube also allows the passage of a gastric tube, which can be utilised to decrease gastric volume. In a cadaver model, a correctly placed PLMA allows fluid in the oesophagus to bypass the pharynx and mouth, when the drainage tube is open. The PLMA provides about twice the seal at 40 ml cuff volume compared with the Classic LMA.[18]

The place of the PLMA is yet to be established. On the evidence to hand, the PLMA may provide more airway protection in situations of moderate to high GOR risk compared with the Classic LMA. It may be suitable for routine use in patients with mild to moderate GOR. Where there is failed endotracheal intubation in a situation of high GOR risk, the PLMA (with or without a gastric tube) is likely to provide a greater margin of safety when compared with the Classic LMA.

**Conclusion**

GOR remains a concern for the anaesthetist. It is common, probably happening more often then we realise. Pulmonary aspiration of gastric contents is fortunately a

rare event, yet carries significant risk of morbidity and mortality, particularly in high-risk patients. The role of the LMA, particularly the PLMA in patients who are at increased risk for GOR is yet to be resolved.

## References


1. Ng A, Smith G. Gastroesophageal reflux and aspiration of gastric contents in anesthetic practice. *Anesth Analg* 2001; 93:494-513.
2. Tiret L, Nivoche Y, Hatton F, Desmonts JM, Vourc'h G. Complications related to anaesthesia in infants and children. *Brit J Anaesth* 1988; 61:263-269.
3. Olsen GL, Hallen B, Hambraeus-Jonzon K. Aspiration during anesthesia: *Acta Anaesthesiol Scand* 1986; 30:84.
4. Snow J. On Chloroform and other Anaesthetics: John Churchill London 1858; 74-75, 102-103
5. Cheyne WW, Burghard FF, Legg TP and Edmunds A. A Manual of Surgical Treatment. Longmans, Green and Co, London 1912; 443-445.
6. Mendelson CL. The aspiration of stomach contents into the lungs during obstetric anesthesia. *Am J Obstet Gynecol* 1946; 52:191-205.
7. Sellick BA. Cricoid pressure to control regurgitation of stomach contents during induction of anaesthesia: preliminary communication. *Lancet* 1961; 2:404-406.
8. Engelhardt T, Webster NR. Pulmonary aspiration of gastric contents in anaesthesia. *Br J Anaesth* 1999; 83:453-460.
9. Kluger MT, Short TG. Aspiration during anaesthesia: a review of 133 cases from the Australian Incident Monitoring Study (AIMS). *Anaesthesia* 1999; 54:19-26.
10. Cohen SE, Jasson J, Talafie ML, Chauvelot-Moachon L, Burrier G. Does metoclopramide decrease the gastric volume in patients undergoing cesarean section? *Anesthesiology* 1984; 61:604-607.
11. Jones MP, Sloan SS, Rabine JC, Ebert CC, Huang CF, Kahrilas PJ. Hiatal henia size is the dominant determinant of esophagitis presence and severity in gastroesophageal reflux disease. *Am J Gastroenterol* 2001; 96(6):1711-1712.
12. Raidon DM, Roche DA, Brock-Utne JG, Marszalek A, Engelbrecht HE. Critical Volume for pulmonary acid aspiration: reappraisal in a primate model. *Brit J Anaesth* 1990; 65:248-250.
13. Boulay K, Blanloeil Y, Bourveau M, Geoy G, Malinovsky JM. Effects of oral ranitidine, famotidine and omeprazole on gastric volume and pH at induction and recovery from general anaesthesia. *Brit J Anaesth* 1994; 73:475-478.
14. Eyler SW, Cullen BF, Murphy ME, Welch WD. Antacid aspiration in rabbits: a comparison of Mylanta and Bicitra. *Anesth Analg* 1982; 61(3):288-292.
15. Owens TM, Roberston P, Twomey C. The incidence of Gastroesophageal reflux with the laryngeal mask. *Anesth Analg* 1995; 80:980-984.
16. Brimacombe JR, Berry A. The incidence of aspiration associated with the laryngeal mask airway: a meta-analysis of published literature. *J Clin Anesth* 1995; 7:297-305.
17. Han TH, Brimacombe J, Lee EJ, Yang HS. The laryngeal mask airway is effective (and probably safe) in selected healthy parturients for elective Caesarean section: a prospective study of 1067 cases. *Can J Anaesth* 2001; 48:1117-1121.
18. Keller C, Brimacombe J, Kleinsasser A, Loeckinger A. Does the ProSeal laryngeal mask airway prevents aspiration of regurgitated fluid? *Anesth Analg* 2000; 91:1017-1020.

# Standard Base Excess

T. J. MORGAN
Senior Specialist, Deputy Director,
Adult Intensive Care, Mater Misericordiae Hospitals, South Brisbane, Queensland

*Thomas (John) Morgan is an examiner for the Joint Faculty of Intensive Care Medicine, and serves on the editorial board of Critical Care and Resuscitation. Dr Morgan has produced original publications in the areas of acid-base, crystalloid design, tissue oxygenation, haemoglobin-oxygen affinity and evaluation of perioperative risk.*

The history of arterial blood gas analysis stretches back over more than forty years. Throughout this time, differences of opinion and entrenched viewpoints have characterized the evaluation of acid-base disturbances. Controversy persists to this day. Most define acidaemia as arterial pH<7.35, and alkalaemia as pH>7.45. Most agree that in isolated respiratory acidosis, $PaCO_2$ exceeds 45 mmHg, and in isolated respiratory alkalosis, $PaCO_2$ is less than 35 mmHg. It is also accepted that metabolic (non-respiratory) acid-base abnormalities manifest on blood gas analysis as a disturbed pH/$PaCO_2$ relationship (Figure 1). The major source of disagreement is how to identify and quantify these metabolic acid-base disturbances. In other words, at any given $PaCO_2$ and pH, what is the best tool to delineate the separate respiratory and metabolic contributions to the overall acid-base status?

The ideal bedside metabolic acid-base index has the following characteristics:

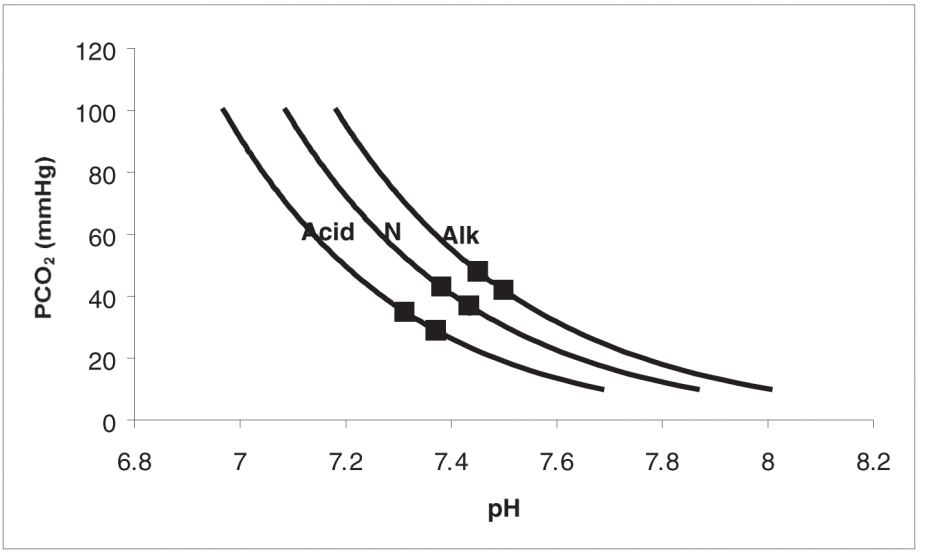


**Figure 1.** $PaCO_2$/pH relationships in vivo. Illustrated are the normal curve (N) and examples of metabolic acidosis (Acid) and metabolic alkalosis (Alk). The black squares delineate the appropriate $PaCO_2$ range for each curve.

1. Simple and "user friendly".
2. Independent of $PaCO_2$ ($CO_2$-invariant).
3. Stoichiometric *in vivo*. This means that the index should quantify the amount of strong acid or base (expressed as mmol/l extracellular fluid) which would correct any metabolic acid-base disturbance.

There are three contenders:

1. The $PaCO_2/[HCO_3^-]$ "rules" of the Boston school.
2. Strong ion difference (Stewart's physical chemical approach).
3. Standard base excess (Copenhagen approach).

We will discuss each of these in turn. On evaluation it should become evident that standard base excess is closest to the ideal, although all three are merely different ways of describing the same process.[1]

**The Boston "rules"**

This is a $PaCO_2/[HCO_3^-]$ based approach developed by Schwartz, Relman and colleagues at Tufts University.[2,3] Six equations were derived using in vivo data from human volunteers and patients (Table 1). These calculate either the appropriate [HCO3-] in primary respiratory disturbances or the appropriate $PaCO_2$ in primary metabolic disturbances (Table 1). If there is a difference between the actual and the expected $[HCO_3^-]$ for a given $PaCO_2$, a metabolic acid-base disturbance is present.

**Table 1**
The six $PaCO_2/[HCO_3^-]$ rules ($PaCO_2$ values are in mmHg, $[HCO_3^-]$ in mmol/l)

| Condition | Rule |
|---|---|
| Acute respiratory acidosis. | Expected $[HCO_3^-]=24+(PaCO_2-40)/10$ |
| Chronic respiratory acidosis. | Expected $[HCO_3^-]=24+4\times(PaCO_2-40)/10$ |
| Acute respiratory alkalosis. | Expected $[HCO_3^-]=24-2\times(40-PaCO_2)/10$ |
| Chronic respiratory alkalosis. | Expected $[HCO_3^-]=24-5\times(40-PaCO_2)/10$ |
| Metabolic acidosis. | Expected $PaCO_2=1.5\,[HCO_3^-]+8$ |
| Metabolic alkalosis. | Expected $PaCO_2=0.9\,[HCO_3^-]+9$ |

The Boston "rules" require considerable bedside "mental gymnastics".[1] Mistakes and memory lapses could lead to adverse clinical consequences. It is true that the equations apply throughout the normal "working" range of $PaCO_2$. This means that there is $CO_2$-invariance. It is also true that a significant deviation from the expected $[HCO_3^-]$ for any given $PaCO_2$ signals a metabolic acid-base perturbation. However, this offset does not quantify the amount of strong acid or base required to correct the metabolic disturbance, even in vitro but most especially in vivo.[4] In other words, $[HCO_3^-]$ offsets are not stoichiometric.

Nevertheless, devotees abound. Some have gone as far as removing "base excess" calculations from analyser printouts. This is an over-reaction.

**Strong ion difference**

Peter Stewart's semi-quantitative analysis using principles of physical chemistry was a major conceptual shift.[5] In this type of analysis, pH and $[HCO_3^-]$ are dependent variables determined by three independent variables, which are $PaCO_2$, strong ion difference ([SID]), and the total concentration of non-volatile weak acid buffer ($[A_{TOT}]$).

*What are strong ions and what is [SID]?*

Strong ions are those essentially fully ionised under all physiologic acid-base conditions. Quantitatively, any unionised fraction is insignificant. Examples of strong ions include $Na^+$, $K^+$, $Ca^{++}$, $Mg^{++}$, $Cl^-$ and lactate. [SID] is [strong cations] – [strong anions]. Since this is a difference in charge concentration, [SID] is expressed in mEq/l. In normal plasma, [SID] is approximately 40 mEq/l. To preserve electrical neutrality, the SID "space" is filled passively by the buffer base.[6] The buffer base components of quantitative importance are $HCO_3^-$ and "non-volatile buffer anions" ($A^-$). Buffer base therefore=($[A^-]+[HCO_3^-]$), and is numerically the same as [SID]. The ions are weak because at physiological pH they combine reversibly with protons to form their conjugate parent molecules as follows:

$$H^+ + A^- \longleftrightarrow HA$$
$$H^+ + HCO_3^- \longleftrightarrow H_2CO_3 \longleftrightarrow CO_2 + H_2O$$

In fact $A^-$ is not a true anion in the strict sense. In plasma, $A^-$ is made up primarily of the net negative charge on albumin, (inorganic phosphate also makes a small contribution). In erythrocytes $A^-$ is the net negative charge on haemoglobin. Both albumin and haemoglobin have a number of ionisable groups with a range of pK values and charges, so that the net molecular negative charge alters with pH. In Stewart's terminology this process is described simplistically as $A^-$ combining with protons to form the conjugate base HA. We will continue with this convention although it is technically incorrect.

If [SID] is reduced, the buffer base concentration must be reduced by an identical amount, since [SID] is the independent variable setting the buffer base boundaries.[5] The buffer base equilibria therefore shift to the right, reducing $[A^-]$ and $[HCO_3^-]$. The $[HCO_3^-]$ reduction shifts the $PaCO_2$/pH curve to the left, producing a metabolic acidosis (Figure 1). Similarly an isolated increase in [SID] causes a metabolic alkalosis.

The focus of Stewart and his followers has always been on plasma [SID] rather than whole blood or extracellular [SID], which are harder to quantify. The easiest way to calculate plasma [SID] is as the buffer base concentration ($[HCO_3^-]+[A^-]$), rather than by laborious direct measurement of all strong anions and cations.[7] It can be argued that although [SID] defines, controls and is identical to the buffer base concentration, it is buffer base which determines metabolic acid-base status by setting the relationship between $PaCO_2$ and $[HCO_3^-]$ and thus that between $PaCO_2$ and pH (Figure 1).[6]

*What is $[A_{TOT}]$?*

Using Stewart's (over-simplified) terminology, $[A_{TOT}]$ is ($[A^-]+[HA]$).[5] In plasma, $[A_{TOT}]$ is comprised largely of albumin and inorganic phosphate, and in erythrocytes $[A_{TOT}]$ consists primarily of haemoglobin.[1] Alterations in $[A_{TOT}]$ will cause the $[A^-]$ component of buffer base to vary in parallel. If [SID] is held constant while $[A_{TOT}]$ is varied, the total buffer base concentration ($[HCO_3^-]+[A^-]$) must remain constant since it is defined and controlled by [SID]. Isolated $[A_{TOT}]$ reductions therefore reduce $[A^-]$ and increase $[HCO_3^-]$, causing a metabolic alkalosis. Conversely isolated $[A_{TOT}]$ elevations cause a metabolic acidosis.[8]

In the physical chemical approach, there are thus two determinants of metabolic acid-base status, [SID] and $[A_{TOT}]$. Both qualify as $CO_2$-invariant, since they are independent variables. The only way that [SID] can used as a stoichiometric index of metabolic acid-base status is by converting it to SIDex. SIDex is the change in whole

blood [SID] at a given [$A_{TOT}$] required to bring the pH to 7.4, while the $PaCO_2$ is maintained at 40 mmHg.[1] This conversion is really a pointless exercise, since SIDex is virtually identical to base excess (Figure 2). Unlike base excess it cannot be calculated from blood gas analysis alone, which is a major disadvantage.

The strength of the Stewart approach lies mainly in its conceptual insights. For example Stewart's principles can be put to use in the design of crystalloid resuscitation fluids with specific acid-base effects.[9] More direct application to bedside acid-base quantification is impractical.

**Standard base excess**

The evolution of the standard base excess concept was gradual.[10] It began with "standard bicarbonate", devised by Jørgensen and Astrup.[11] They used tonometry to convert small volumes of blood to fully oxygenated specimens with a $PCO_2$ of 40 mmHg at 38°C, and then measured the plasma pH. "Standard bicarbonate" concentration could then be determined by substituting the $PCO_2$ (40 mmHg) and the measured pH value in the Henderson-Hasselbalch equation. $CO_2$-invariance was thus ensured by physically returning the $PCO_2$ to 40 mmHg. Any offset of the standard bicarbonate from 24.4 mmol/l signified a metabolic acid-base process.

However, there are two problems with this approach. The first is that [$HCO_3^-$] indices are not stoichiometric. The second is that the in vitro blood tonometry does not replicate the in vivo pH/$PaCO_2$ relationship.

*The answer to the stoichiometry problem*

Ole Siggaard-Andersen and colleagues solved the stoichiometry problem by introducing a new parameter called base excess (BE).[12] BE can be defined as the concentration of strong acid or base required to return the pH of an in vitro specimen of whole blood to 7.4 while maintaining $PCO_2$ at 40 mmHg at 37°C by equilibration. If

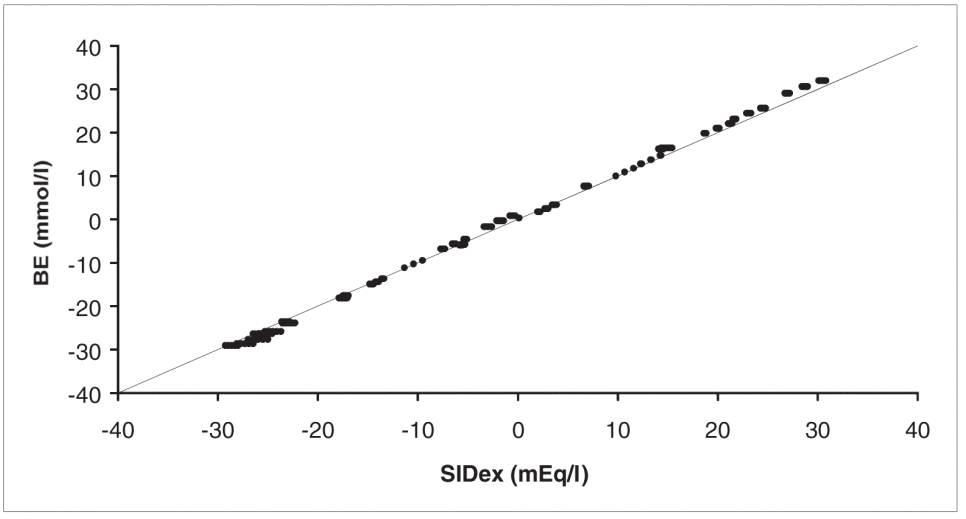


**Figure 2.** Relationship between whole blood base excess (BE) and whole blood strong ion difference excess (SIDex). There are 271 data points, with [haemoglobin] ranging from 50 to 150 g/l, [albumin] from 10 to 40 g/l, $PaCO_2$ from 10 to 60 mmHg and pH from 6.9 to 7.6. The line of identity is shown. By inspection, values are almost identical. $R^2$=0.999.

plasma pH is >7.4 once the $PCO_2$ has been returned to 40 mMHg, BE is the mmol/l of strong acid which brings the pH to 7.4 (while $PCO_2$ continues to be maintained at 40 mmHg). BE then has a positive sign. If pH<7.4, BE is quantified by the strong base required (BE then has a negative sign). A negative BE is sometimes referred to as a "base deficit".

To overcome the need for individual bench top tonometry and acid-base titrations, a series of in vitro experiments was carried out. These recorded the effects on plasma pH of adding known amounts of acid or base to blood maintained by tonometry at various $PCO_2$ values at 37°C.[13, 14] The experiments were repeated with a wide range of haemoglobin concentrations, which is a way of factoring in the effects of varying $[A_{TOT}]$. From the data, an "alignment nomogram" was created. With this nomogram it was possible to determine BE from a single measurement of pH, $PCO_2$ and haemoglobin concentration at 37°C, without tonometry or acid-base titrations. By the late 1960s, blood gas analysers could be programmed with the BE nomogram in the form of look-up tables or equations.

However, BE can also be defined more simply as the offset in buffer base concentration. In other words BE is really ($\Delta A^- + \Delta[HCO_3^-]$). Plasma BE is thus easy to derive by non-empiric means. All that is required to calculate $\Delta A^-$ is the pH and a value for the buffering capacity of albumin, and $\Delta[HCO_3^-]$ is simply ($[HCO_3^-]$–24.4) mmol/l. For whole blood the non-empiric method is more complicated, since whole blood is plasma mixed with red cells, and haemoglobin is the predominant buffer. In 1977 Siggaard-Andersen published the Van Slyke equation,[15] which enables whole blood BE to be calculated from pH, $PaCO_2$, and haemoglobin concentration using linking equations for plasma and intra-erythrocytic buffering and Gibbs-Donnan ionic distributions. The Van Slyke equation can be written:

$$BE = \{[HCO_3^-] - 24.4 + (2.3 \times [Hb] + 7.7) \times (pH - 7.4)\} \times (1 - 0.023 \times [Hb])$$

where $[HCO_3^-]$ and pH are plasma values and [Hb] is the blood haemoglobin concentration expressed in mmol/l. On publication, Siggaard-Andersen reported close agreement between this equation and its empiric predecessor — the whole blood BE nomogram.[15] Our group has re-evaluated the in vitro accuracy of the Van Slyke equation. We confirmed that it quantifies whole blood metabolic acid-base change with acceptable degrees of precision and minimal bias.[16] We found that the accuracy was maintained at very low haemoglobin concentrations, and was little affected by large simultaneous alterations in $PCO_2$.

*The in vivo versus in vitro equilibration problem*

In 1963, Schwarz and Relman pointed out that BE is not $CO_2$-invariant in vivo.[2] This is because for any specimen of arterial blood the in vitro plasma pH/$PaCO_2$ equilibration curve differs from the in vivo curve, since in vivo $CO_2$ equilibration occurs throughout the total extracellular compartment. An isolated in vivo change in $PaCO_2$ shifts whole blood BE in the opposite direction although no overall extracellular metabolic acid-base alteration has occurred (Figure 3).

This phenomenon is due to ionic shifts between intravascular and interstitial compartments. When $PaCO_2$ rises, $HCO_3^-$ is generated maximally within the erythrocytes, where buffering capacity is greatest due to haemoglobin. It then diffuses down a concentration gradient to the plasma (as occurs in vitro) but continues on into the interstitial fluid. $Cl^-$ moves in the opposite direction, with final ionic distributions

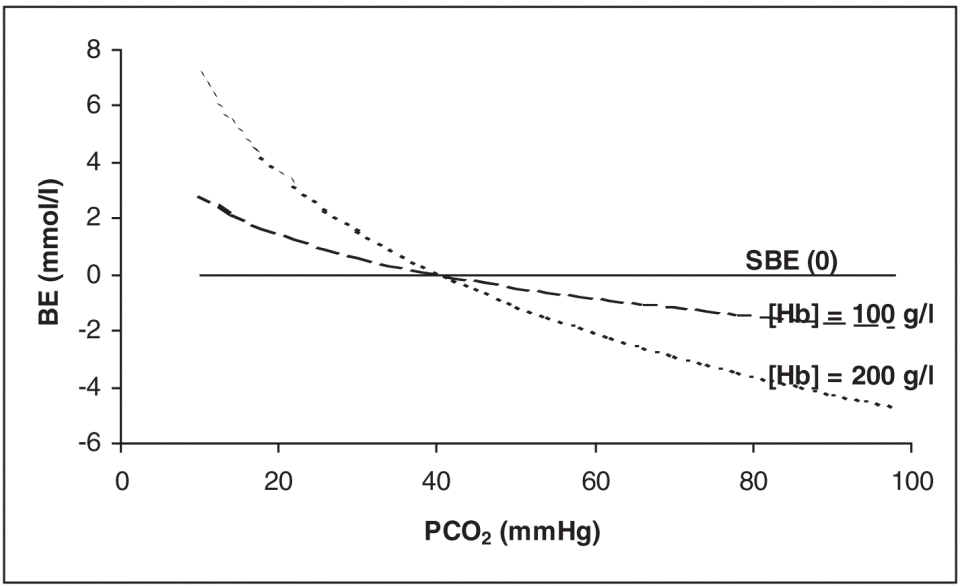


**Figure 3.** Relationship in vivo between BE and $PaCO_2$ for haemoglobin concentrations of 100 g/l and 200 g/l. A change in $PaCO_2$ causes BE to change in the opposite direction. The effect is more marked at higher haemoglobin concentrations. In contrast, standard base excess (SBE) does not change with $PaCO_2$ (see text).

determined according to Gibbs-Donnan equilibria and the laws of chemical equilibrium and electroneutrality. Thus the rise in plasma [$HCO_3^-$] during hypercapnia will not be as high in vivo as in vitro. Or put in physical chemical terms, when $PCO_2$ rises in an in vitro specimen of whole blood, erythrocytic [SID] decreases, plasma [SID] increases but whole blood [SID] and thus BE do not change. When $PaCO_2$ rises in vivo, whole blood [SID] and thus BE decrease, interstitial [SID] increases but extracellular [SID] is unaffected.

The answer to the objection of Schwarz and Relman is remarkably simple. If BE is calculated from the measured plasma pH and $PCO_2$ as originally described, but at a haemoglobin concentration of 50 g/l (the approximate mean extracellular haemoglobin concentration), total extracellular buffering is emulated successfully.[6] BE calculated in this way is termed standard BE (SBE). "Rules of thumb" equations derived by meta-analysis of published numerical and graphical data now describe the appropriate SBE responses to acute and chronic respiratory acid-base disturbances, and the appropriate $PaCO_2$ responses to primary disturbances in SBE.[17] They are similar in nature to the Boston rules, and are set out in Table 2. Importantly, the meta-analysis showed that

**Table 2**
The four $PaCO_2$/SBE rules (SBE in mmol/l, $PaCO_2$ in mmHg)

| Condition | Rule |
|---|---|
| Acute respiratory acidosis and alkalosis | $\Delta SBE = 0 \times \Delta PaCO_2$ |
| Chronic respiratory acidosis and alkalosis | $\Delta SBE = 0.4 \times \Delta PaCO_2$ |
| Metabolic acidosis | $\Delta PaCO_2 = \Delta SBE$ |
| Metabolic alkalosis | $\Delta PaCO_2 = 0.6 \times \Delta SBE$ |

SBE is $CO_2$-invariant in vivo (see Table 2, Rule 1), confirming that although SBE is only an approximation of total extracellular buffering, it successfully replicates the in vivo situation.

**Clinical application of standard base excess**

In most laboratories, the reference range for SBE is approximately –3 mmol/l to +3 mmol/l. Before applying the SBE rules, primary acid-base disturbances should be detected first by inspection of the pH and $PaCO_2$ (Table 3). A primary process is one that if unopposed will shift the pH out of the normal range, and can be either respiratory ($PaCO_2$) or metabolic. There can be one, two or less commonly more than two primary processes operating at once. A primary metabolic process should stimulate immediate respiratory compensation, and a primary respiratory process normally results in metabolic compensation by renal [$HCO_3^-$] adjustment within hours. However compensation rarely returns the pH to normal except in the case of chronic respiratory alkalosis. With this exception, it can be stated that if only one of either $PaCO_2$ or pH is normal, two primary processes must be operating simultaneously.

**Table 3**
Primary processes revealed by initial examination of $PaCO_2$ and pH

| $PaCO_2$ | pH | Primary processes |
|---|---|---|
| Normal | Normal | None |
| Normal | High | Metabolic alkalosis, Respiratory alkalosis |
| Normal | Low | Metabolic acidosis, Respiratory acidosis |
| High | Normal | Respiratory acidosis, Metabolic alkalosis |
| High | High | Metabolic alkalosis |
| High | Low | Respiratory acidosis |
| Low | Normal | Chronic respiratory alkalosis |
| Low | High | Respiratory alkalosis |
| Low | Low | Metabolic acidosis |

After scanning for primary processes, SBE can then be combined with the $PaCO_2$/SBE rules in Table 2 to quantify:

1. The in vivo metabolic component of any acid-base disturbance.
2. The appropriateness of the metabolic response to any respiratory acid-base derangement.
3. The appropriateness of the respiratory response to any metabolic acid-base derangement.

For example, if the pH, $PaCO_2$ and SBE are all elevated, the primary process is a metabolic alkalosis. The next step is to quantify the severity of the metabolic alkalosis by determining the elevation of SBE above the normal range. It then remains to decide whether the accompanying respiratory acidosis is purely compensatory. If so, the $PaCO_2$ should be within a few mmHg of $(40+0.6\times\Delta SBE)$ mmHg. If this is not so, there are two primary acid-base disturbances (metabolic and respiratory).

*Concealed metabolic acidosis*

It is important to understand that a normal SBE can conceal complex metabolic acid-base disturbances. The example given by Schlichtig and colleagues concerns a patient with chronic lung disease and a compensated respiratory acidosis who develops

septic shock and lactic acidosis.[1] In this case SBE and [SID], originally elevated by the compensatory metabolic alkalosis, are reduced acutely to normal by an increased concentration of lactate (a strong anion). The normal SBE and elevated $PaCO_2$ give the appearance on blood gas analysis of a simple acute respiratory acidosis. Here the truth can only be revealed by attention to history and examination, supplemented by the anion gap (which is likely to be elevated) and a measured plasma lactate.

**In vivo evaluation of SBE**

In virtually all published in vivo studies, it is BE rather than SBE which has been evaluated. The difference between BE and SBE is usually quite small.

*Some animal studies*

In 1991, Davis and colleagues bled 16 pigs to 40% of their blood volumes over 30 minutes, waited a further 30 minutes, then resuscitated the animals with crystalloid and blood.[18] With the onset of hypovolemia, mean arterial pressure and mixed venous oxygen saturations decreased promptly and arterio-venous oxygen extraction increased, but all began to improve from compensatory homeostasis before resuscitation was commenced. BE decreased just as promptly, but improvement was delayed until volume replacement. The authors commented that BE was more reflective of the true volume deficit in compensated hypovolemic shock. In a related study, BE correlated significantly with arterial lactate concentrations ($R=0.81$).[19]

Another report in 1991 came from Dunham and colleagues, who subjected 63 dogs to haemorrhagic shock titrated to predetermined oxygen debts.[20] Baseline oxygen consumption was measured prior to controlled bleeding over one hour to oxygen debts ranging from 60 ml/kg to 120 ml/kg. At 60 minutes there was no statistically significant difference in shed blood volumes among the groups, and neither blood pressure nor cardiac output discriminated well between the degrees of oxygen debt. However plasma lactate and whole blood BE were both highly significant discriminators of oxygen debt, with BE having the highest explained variability. The fall in BE prior to resuscitation was almost twice the rise in plasma lactate, perhaps reflecting additional dysoxic metabolites.

*Some clinical studies*

In a retrospective analysis, it was possible using the admission BE to stratify 209 patients with blunt or penetrating trauma into groups with very similar initial mean arterial pressures, trauma scores, blood and total fluid requirements.[21] A decreasing BE was associated with ongoing hemorrhage in 65% of cases. Another retrospective review of 3223 blunt trauma patients revealed the admission BE as the single most important predictor of the need for laparotomy (odds ratio 6.2).[22] The less successful predictors were admission hypotension (odds ratio 5.1), chest injury (odds ratio 4.8), pelvic fracture (odds ratio 3.3) and pre-hospital hypotension (odds ratio 3.1). The authors recommended that in cases of blunt abdominal trauma an admission BE of $<-6.0$ mmol/l should be considered an indication for diagnostic peritoneal lavage or abdominal CT. In burns, the admission BE predicts resuscitation volume requirements, and when $\leq -6.0$ mmol/l it portends a greatly increased mortality rate (9% versus 72%, $P<0.001$).[23]

Siegel and colleagues studied admission data from 185 consecutive patients with hepatic trauma.[24] By linear logistic modeling, BE predicted death as successfully as the

24 hour blood transfusion requirement, and better than arterial lactate, the Glasgow Coma Score and the Injury Severity Score. A predictive model incorporating BE and the Glasgow Coma Score was then tested successfully against 323 additional patients with pelvic fracture as their index injury. In one recent study of general medical/surgical intensive care patients, the admission BE was strongly linked with mortality.[25] By combining the admission BE with the admission plasma lactate concentration, it was possible to predict mortality with a sensitivity of 80.3% and a specificity of 58.7%. Failure to improve BE significantly over 24 hours in this population was an adverse prognostic factor. In shocked trauma patients, failure to increase BE to >−6 mmol/l in the first 24 hours predicted an increased frequency of acute respiratory distress syndrome, multiple organ failure and death.[26] There are other similar studies.[27, 28]

In summary, SBE is a simple, user-friendly $CO_2$-invariant index of extracellular metabolic acid-base status. BE itself is stoichiometric on in vitro evaluation in whole blood. From studies in shock, trauma and in ICU patients in general, BE appears to have clinical meaning. It can be a guide to the overall severity of shock and trauma, and is closely related to oxygen debt in experimental hemorrhagic shock. It tracks elevated plasma lactate concentrations, and at times adds information to lactate analysis. BE may have some utility as a guide to resuscitation or even surgical intervention. Importantly, SBE and the Copenhagen approach are quite consistent with the physical chemical analysis of acid-base status, and complement rather than antagonize the Stewart approach.

**References**


1. Schlichtig R, Grogono AW, Severinghaus JW. Current status of acid-base quantitation in physiology and medicine. *Anesthesiology Clinics of North America* 1998; 16:211-233.
2. Schwartz WB, Relman AS. A critique of the parameters used in the evaluation of acid-base disorders. *N Engl J Med* 1963; 268:1382-1388.
3. Narins RG, Emmett M. Simple and mixed acid-base disorders: A practical approach. *Medicine* 1980; 59:161-187.
4. Severinghaus JW. Siggaard-Andersen and the "Great Trans-Atlantic Acid-Base Debate". *Scand J Clin Lab Invest* 1993 ;53 Suppl 214:99-104.
5. Stewart PA. Modern quantitative acid-base chemistry. *Can J Physiol Pharmacol* 1983; 61:1444-1461.
6. Siggaard-Andersen O, Fogh-Andersen N. Base excess or buffer base (strong ion difference) as measure of a non-respiratory acid-base disturbance. *Acta Anesth Scand* 1995; 39: Suppl 107, 123-128.
7. Figge J, Mydosh T, Fencl V. Serum proteins and acid-base equilibria: a follow-up. *J Lab Clin Med* 1992; 120:713-719.
8. Rossing TH, Maffeo N, Fencl V. Acid-base effects of altering plasma protein concentration in human blood in vitro. *J Appl Physiol* 1986; 61:2260-2265.
9. Morgan TJ, Venkatesh B, Hall J. Crystalloid strong ion difference determines metabolic acid-base change during in vitro hemodilution. *Crit Care Med* 2002; 30:157-160.
10. Severinghaus JW. AHA! In Astrup P, Severinghaus JW (ed.) The History of Blood Gases, Acids and Bases. Copenhagen: Munksgaard, 1986; pp 264-276.
11. Jørgensen K, Astrup P. Standard bicarbonate, its clinical significance, and a new method for its determination. *Scand J Clin Lab Invest* 1957; 9:122.
12. Siggaard-Andersen O, Fogh-Andersen N. Base excess or buffer base (strong ion difference) as measure of a non-respiratory acid-base disturbance. *Acta Anesth Scand* 1995; 39:Suppl 107, 123-128.
13. Siggaard-Andersen O, Engel K. 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.
14. Astrup P, Jorgensen K, Siggaard-Andersen O et al. Acid-base metabolism: New approach. *Lancet* 1960; 1:1035-1039.
15. Siggaard-Andersen O. The Van Slyke equation. *Scand J Clin Lab Invest* 1977; 37 Suppl 146:15-20.
16. Morgan TJ, Clark C, Endre ZH. The accuracy of base excess — an in vitro evaluation of the Van Slyke

equation. *Crit Care Med* 2000; 28:2932-2936.
17. Schlichtig R, Grogono AW, Severinghaus JW. Human $PaCO_2$ and standard base excess compensation for acid-base imbalance. *Crit Care Med* 1998; 26:1173-1179.
18. Davis JW, Shackford SR, Holbrook TL. Base deficit as a sensitive indicator of compensated shock and tissue oxygen utilization. *Surg Gynecol Obstet* 1991; 173:473-476.
19. Davis JW. The relationship of base deficit to lactate in porcine hemorrhagic shock and resuscitation. *J Trauma* 1994; 36:168-172.
20. Dunham CM, Siegel JH, Weireter L et al. Oxygen debt and metabolic acidemia as quantitative predictors of mortality and the severity of the ischaemic insult in hemorrhagic shock. *Crit Care Med* 1991; 19:231-243.
21. Davis JW, Shackford SR, Mackersie RC et al. Base deficit as a guide to volume resuscitation. *J Trauma* 1988; 28:1464-1467.
22. Davis JW, Mackersie RC, Holbrook TL et al. Base deficit as an indicator of significant abdominal injury. *Ann Emerg Med* 1991; 20:842-844.
23. Kaups KL, Davis JW, Dominic WJ. Base deficit as an indicator of resuscitation needs in patients with burn injuries. *J Burn Care Rehabil* 1998; 19:346-348.
24. Siegel JH, Rivkind AI, Dalal S et al. Early physiologic predictors of injury severity and death in blunt multiple trauma. *Arch Surg* 1990; 125:498-508.
25. Smith I, Kumar P, Molloy S et al. Base excess and lactate as prognostic indicators for patients admitted to intensive care. *Intensive Care Med* 2001; 27:74-83.
26. Davis JW, Kaups KL, Parks SN. Base deficit is superior to pH in evaluating clearance of acidosis after traumatic shock. *J Trauma* 1998; 44:114-118.
27. Davis JW, Parks SN, Kaups K et al. Admission base deficit predicts transfusion requirements and risk of complications. *J Trauma* 1996; 41:769-774.
28. Davis JW, Kaups KL. Base deficit in the elderly: a marker of severe injury and death. *J Trauma* 1998; 45:873-877.

# Herbal Medicine and Perioperative Care — An Australian Perspective

DR ROBERT GRAUER, MB, BS(Hons), FANZCA
Staff Anaesthetist at St. Vincent's Hospital, Melbourne, Victoria
Sessional (VMO) Anaesthetist at Box Hill Hospital, Victoria

*Dr Grauer is a Staff Anaesthetist, St Vincent's Hospital, Melbourne and a Visiting Medical Officer, Box Hill Hospital, having trained at both these institutions. In 2000, he worked as a Staff Anaesthetist in Munster, Germany. Neuroanaesthesia is a particular interest.*

## INTRODUCTION

In recent years, there has been a definite rise in the popularity and use of complementary and alternative medicines (CAM), including herbal medicines, in developed countries.[1] A "herbal medicine" is defined as a plant-derived product used for medicinal or health purposes.[1] Also known as nutraceuticals or natural remedies, they are perceived by many patients as being "natural" and therefore safe.[2]

As anaesthetists, we cannot ignore the impact that these medicines have on anaesthetic practice. This includes the potential interaction that such products have either with a patient's regular prescription medications or with anaesthetic drugs. Anaesthetic techniques (e.g. regional anaesthesia) may need to be reconsidered if patients have been taking herbal medicines.

## HISTORY AND CULTURE

The earliest evidence of human use of plants for healing dates back to the Neanderthal period. In the 16th century, botanical gardens were created to grow medicinal plants for medical schools[3] and herbal remedies flourished until the 17th century when more "scientific" pharmacological remedies came into favour. At that time, though, health care was commonly provided by women using home-made botanical remedies and purchasing similar products as "patent medicines". Scientific methods became more advanced and preferred in the 19th century and the practice of botanical healing was dismissed as quackery. However, by the 1960s, with concerns over the iatrogenic effects of conventional medicine and desire for self-reliance, interest in "natural health" and the use of herbal products increased.[3]

In 1985 the World Health Organisation estimated that up to 80% of the world's population still relied on botanicals for their primary health needs.[4, 5] Many cultures espouse the use of botanicals in their traditional healing practices, such as Chinese medicine, Ayurveda (a holistic system originated in the Vedic civilization of India), Curanderismo (a Mexican-American healing tradition), or Kampo medicine in Japan.[6, 7] This widespread use is not restricted to developing countries; up to 30% of doctors in Germany and France use herbs in practice.[5] In the pharmacopoeias of developed countries, 25% of drugs are substances first isolated from plants and a further 25% are modifications of chemicals first found in plants.[1]

## HERBAL PREPARATIONS

Herbs are available in many forms, depending on the type of plant and its use. They may include fresh, dried, chopped or whole herbs and can be steeped as teas (infusions) or simmered over low heat (decoctions). Other preparations include tinctures (fresh or dried herbs preserved in alcohol), vinegar extracts (acetracts), syrups, glycerites (in vegetable glycerine) or miels (in honey).[8, 9] Freeze-dried or herbal powders also come in bulk, tablet, capsule, paste, or concentrate (4-6 times regular strength) forms. Other ways to administer herbs may be by suppositories, creams, gels, liniments, oils, compresses, steams, aromatics (oils) or baths.[9]

## WHO AND WHY DO PEOPLE USE HERBAL MEDICINES

Virtually all surveys on this subject agree on several characteristics associated with the typical user of herbal medicines.[6] They are more likely to be female, 40-60 years old, better educated, have a higher income and be employed.[6, 10, 11, 12] The reasons for consumers turning to herbal medicine are complex and differ according to many factors. A patient with a life-threatening condition has different motivations for trying herbal medicines than a person who is essentially healthy but wants to invest in well-being or hopes to prevent future illness.[6] Contrary to common belief, disappointment with orthodox medicine is not a prime reason for people to turn towards alternative medicines. Potential reasons for trying herbal medicines include: perceived effectiveness and safety, emphasis on holism and natural products, control over treatment, affluence, accessibility, rejection of science and technology, desperation and dissatisfaction with conventional pharmacological treatment.[6]

## ANAESTHETISTS KNOWLEDGE OF AND ATTITUDE TOWARDS HERBAL MEDICINES

Do physicians and, in particular anaesthetists actually know much about the significance of these herbal medicines? What is the point of patients reporting herbal medicine usage if we are not aware of the clinical importance of these preparations?[13] In a study by Lennox, of the 28 anesthesiologists who participated by completing a 'herbal medicine' questionnaire containing 27 yes/no type responses, only 32% of the questions were answered correctly.[13]

Attitudes of physicians, in either being dismissive of herbal medicines or considering them as no more relevant than harmless placebos,[3] do not foster knowledge or understanding. Two studies highlighted the low percentage of documentation of herbal medicine use on the anaesthetic chart, even after the patient had reported their usage[12, 14] However, attitudes to herbal medicines are evolving. The Australian and New Zealand College of Anaesthetists, American Society of Anesthesiologists and the Australian Medical Association have all produced statements or papers on herbal medicines.[15, 9, 16] The recent anaesthetic literature contains numerous articles on "anaesthesia and herbal medicines". Multiple choice questions on herbal medicines even appeared in the College Part 2 Exams last year.

## USAGE AND COST OF HERBAL MEDICINES

A 1993, a South Australian survey of 3004 subjects found that 12% had used herbal medicines in the previous year.[17] Although current Australian expenditure on herbal remedies is unknown, $621 million is spent per year on CAM.[18] In the US, recent data estimated that 49% of the adult population have used a herbal product during the

previous year, with 24% of the population regular users and US$5.1 billion spent on such remedies.[19] The European market is valued at US$6 billion per year.[5] A conservative estimate of US$10 per month is being spent on herbs per patient.[20] Sales of herbal medicines are growing by 20% a year and are the largest growth area in retail pharmacy, far exceeding growth in the conventional drug category.[3] Sales have moved from speciality stores to mainline shopping environs.[21]

Numerous studies have investigated the use of herbal remedies by patients about to undergo surgery. In Australia, a 2002 study of 1102 patients (about to be published), revealed the prevalence rate for perioperative herbal usage to be 14.3%. Individual studies from Canada, England and Ireland have found a rate of 33.4%, 4.8% and 12.1% respectively.[13, 12, 22] In seven recent studies from the USA, the prevalence of perioperative herbal usage lies between 9.7%-37.0%, with a mean of 25.6%.[2, 11, 20, 23-26]

Failure to disclose patients' use of herbal medications to the treating doctor is common. Explanations for this include patient-held beliefs that physicians are not knowledgeable about herbal medications or are prejudiced against them, patients not considering these substances to be medications and others' fearing admitting to their physicians use of unconventional therapies.[27] Importantly, failure of the doctor to specially ask the patient about their use of herbal medications discourages disclosure. A South Australian survey of 3,027 people in 2000, found that only 42.8% told their doctor of their use of alternative medicines.[10] A survey of 325 patients attending a Sydney teaching hospital emergency department in 1994 revealed that 35.5% of users had informed their medical practitioner about any use of CAM.[17] This low rate of disclosure is similarly reflected in the surgical population.[1] In Crowe's study, 83% of the surgical team were unaware that the patient took herbal medications.[22] The rate of disclosure of herbal medications usage by patients to their anaesthetists is also low, being 30% and 43.6% in two US studies.[20, 23] An Australian 2002 study, found that only 27.8% of preoperative patients informed the hospital's doctors or other staff of their usage of herbal remedies, whereas 41.8% had told their family doctor.

## REGULATION OF HERBAL MEDICINES IN AUSTRALIA

Australia has one of the more stringent regimes for regulation of alternative medications production and labelling.[15] Products for medicinal use (prescription, scheduled proprietary medicines and all complementary medicines) must be placed on the Australian Register of Therapeutic Goods (ARTG), in one of two categories — 'listed' or "registered".[17, 28] Formulations can be "listed" for a small fee if they contain substances regarded by the Therapeutic Goods Administration (TGA) as being of low public health concern and comply with the TGA advertising code. This restricts wording of claims to "assist" rather than to "treat" and limits indications to minor self-limiting conditions.[17] Manufacturers and sponsors are only required to "hold the evidence" for preparations for which minor therapeutic claims are being made. These claims are not routinely evaluated for a listable preparation.[28] About 4,500 plant-based products are listed; these are given an "AUST L" number.[17]

Registered products, which bear an "AUST R" number, contain herbs that either make substantial efficacy claims (e.g. disease prevention, modification or management) or those that the TGA specify as being of some health concern or are a restricted substance.[17] Registered medicines are individually evaluated for safety, quality and efficacy before they are released onto the market.[29] For registration, which is much more costly, appropriate documentation outlining clinical trial work must be submitted

to the Complementary Medicines Evaluation Committee (CMEC), which advises the TGA. Less than five CAM products have been evaluated in this way.[17]

Adverse drug reporting (ADR) is as essential for CAM products as it is for pharmaceuticals in providing post-marketing surveillance. Given the widespread use of CAM, the low number of (voluntary) adverse reports to Adverse Drug Reactions Advisory Committee (ADRAC) suggests that either CAM have a low risk of adverse effects or such effects are significantly under-reported.[17] The latter is likely for the following reasons: the public perception that "natural" products are safe, the use of CAM is not routinely included in patients' drug histories and that, until recently, ADRAC has not actively encouraged the reporting of adverse reactions by practitioners and consumers of alternative medicines.[17]

## EFFICACY AND SAFETY STUDIES

Although herbal medicines have been used for hundreds of years, traditional use is a poor indication for either efficacy or safety.[6] Randomised clinical trials (RCT), the accepted standard for testing efficacy, have had limited impact in regards to herbal medicines for several reasons. Since herbal remedies (i.e. plants) cannot be patented, and as they are unlikely to recoup the estimated US$350 million it costs to prove a new drug effective and safe, there is little motivation for manufacturers to conduct RCTs.[3] Despite this, RCTs of medicinal products have been conducted but, more often than not, trials assessing a certain treatment for a given condition differ in their methodological details and findings.[6] Therefore, the majority of literature on herbs is based on case reports, uncontrolled case series and studies on compounds that cannot be consistently reproduced.[7]

In 1978, the German Federal Health Agency established Commission E, which comprised an expert panel of physicians, pharmacists, toxicologists and industrial representatives. They were charged with the task of regulating and researching the safety and efficacy of herbal medicines. Until 1994, Commission E independently evaluated all available scientific data on herbs and published results in the form of monographs. These have been translated, reviewed and are widely used as reliable sources of information on herbs.[30]

## ADVERSE EFFECTS OF HERBAL MEDICINES

The uncritical acceptance of the mantra "natural=safe" by both the public and doctors has been a major problem. The media have also helped to perpetuate this myth.[31] Natural plant products are perceived to be "healthier" than manufactured medicines.[3] In one study, 92% of patients attending a Sydney emergency department felt that CAM agents were completely safe.[32] In fact, there are adverse effects of herbal medications; these may be either intrinsic or extrinsic.

### Intrinsic Effects

Intrinsic effects are those of the herb itself and are characterised as for pharmaceuticals, as type A (predictable, dose-dependent) and type B (unpredictable, idiosyncratic).[17]

### Extrinsic Effects

Extrinsic effects are not related to the herb itself, but to problems in commercial manufacture or extemporaneous compounding.[17]

*Misidentification:* Plants can be named four different ways, the common English name, the translated name, the Latinised pharmaceutical name and the scientific name. For example the Chinese herb "dong quai" is also known as "dong guai", "danggui", "tang kuei", "Angelica polymorpha" and "Radix angelica".[17]

*Lack of Standardisation:* The therapeutic and toxic components of plants vary depending on the part of the plant used, stage of ripeness, geographical area where the plant is grown and storage conditions. In one study the biological activity of ginseng was examined in fifty commercial brands sold in 11 countries. In 44 of these products, the concentration of ginsenoside (the active component of ginseng) ranged from 1.9% to 9% w/w; six products contained no ginsenoside and one product contained large amounts of ephedrine.[17]

*Contamination:* During growth and storage, crude plant material can become contaminated by pesticide residues, microorganisms, aflatoxins, radioactive substances and heavy metals; lead, cadmium, mercury and arsenic have been reported as contaminants of some overseas herbal preparations.[17]

*Substitution and Adulteration:* In 2001, the TGA released a "Practitioner Alert" advising about the risk of Aristolochia fangchi, containing the nephrotoxic component aristolochic acid, instead of Stephania tetrandra in a commercial preparation of a slimming treatment. The condition known as "Chinese herb nephropathy" affected 100 women overseas, 30 of whom developed terminal renal failure.[6, 17, 33]

*Incorrect Preparation/Dosage:* A West Australian patient had a heart attack when he failed to follow a herbalist's instructions.[17]

*Inappropriate Labelling/Advertising:* In the US, a 1999 study found that only 43% of dietary supplement products (including herbs) available for sale in stores, catalogues and on the Internet were correctly labelled as a "Supplement".[7] Some products labelled as ginseng actually also contained mandrake (scopolamine) or snake-root (reserpine) because of the high cost of pure ginseng.[3]

## WITHDRAWAL OF HERBAL MEDICINES BEFORE SURGERY

Currently the American Society of Anesthesiologists advises patients to stop taking herbal medications at least 2 to 3 weeks before surgery to allow time for the herbals to be cleared from the body. If the time span is shorter than two weeks, patients should be advised to bring the product in its original container to the hospital, so that the anaesthetist can see exactly what the patient is taking and what the ingredients are.[21, 34] However, a review of the literature favours a more targeted approach. Pharmacokinetic data on selective active constituents indicate that some herbal medicines are eliminated quickly and may be discontinued closer to the time of surgery.[27]

## MEDICINAL HERBAL USE IN AUSTRALIA

The ten most popular herbs used for medicinal purposes in descending order, from an Australian study, are: Garlic, Evening Primrose, Gingko, St. John's Wort, Echinacea, Valerian, Horseradish, Ginseng, Cranberry and Aloe. The top five to eight herbs consistently reappear in similar studies conducted throughout the world.[11, 12, 13, 20, 30, 35] Although there are a large number of different herbal medications available (over 1500 in the US), the top eight herbs account for more than 50% of single herb preparations sold.[27]

The possibility of drug interactions between herbal and orthodox medications is of

major concern.[32] Twenty to 34% of herbal medicine users also report concurrent prescription or over-the-counter medication use.[4, 7, 22]

Typical popular herbal treatments and their indications, adverse effects and possible interactions are listed in Appendix 1. Specific perioperative considerations are detailed in Appendix 2.

## CONCLUSION

Herbal medicine is here to stay. For the benefit of the patient and the physician, it is important for the anaesthetist to inform themselves about the potential benefits, drug interactions and adverse effects of herbal medications. Although there is a vast array and confusing nomenclature of different medicinal herbs, familiarity with the common 15-20 herbs, which account for the great majority that are sold, is all that really is required. Additional information if necessary, can always be sourced from a reliable reference. The question "Do you take any herbal medications or other substances" should be a routine part of every anaesthetist's pre-operative assessment.

## APPENDIX 1: COMMON HERBAL TREATMENTS

**Garlic** (Allium sativum)

*Indications:*

| | |
|---|---|
| Hypertension (moderate effect) | Level 1 evidence[1, 27, 36] |
| Hypercholesterolaemia (moderate effect) | Level 1 evidence[1, 36] |
| Atherosclerosis (moderate effect) | Level 4 evidence[27] |

*Significant Adverse Effects:*
Risk of spontaneous (including epidural haematoma) and post operative bleeding[1, 27, 36]

*Drug Interactions:*
Increased INR levels in patients previously stabilised on warfarin[36]
Avoid in combination with aspirin, NSAID because of enhanced antiplatelet activity[1]

*Cessation prior to Surgery:*
At least 7 days before surgery[27, 37]

**Evening Primrose** (Oenothera biennis)

*Indications:*
Premenstrual syndrome, menopause (No proven benefit) Level 1 evidence[6]

*Significant Adverse Effects:*
None

*Drug Interactions:*
Potential to lower seizure threshold with anti-epileptic medications[6]

*Cessation prior to Surgery:*
Probably not required

**Gingko** (Gingko biloba)

*Indications:*
Cognitive impairment (memory and concentration ability) in patients with Alzheimer's disease and multi-infarct dementia — Level 1 evidence[1, 36]
Intermittent claudication — Level 2 evidence[1, 31]
Tinnitus (vascular origin) — Level 2 evidence[31]

*Significant Adverse Effect:*
Intracranial haemorrhage, postoperative bleeding (case reports)[1, 27, 31]
Seizures — seen in children after excessive ingestion of seeds[31]

*Drug Interactions:*
Concomitant use with aspirin, NSAID and anticoagulants not advised[36]
May diminish the efficacy of anticonvulsants[36]

*Cessation prior to Surgery:*
At least 2 days (probably safer 7 days) before surgery[27, 37]

**St John's Wort** (Hypericum perforatum)

*Indications:*
Mild to moderate depression — Level 1 evidence[1, 36]

*Significant Adverse Effects:*
Photosensitivity (high doses in fair skinned people)[36]

*Drug Interactions:*
SSRI, TCA may precipitate the "serotonin syndrome"[1, 36]
MAOI potentiates side effects[36]
Induces hepatic enzymes —>decrease level of cyclosporin, warfarin, theophylline, oral contraceptive pill, digoxin, antiretroviral agents, anticonvulsants[1, 27, 31]

*Cessation prior to Surgery:*
Discontinue at least 5 days before surgery[27, 37]

**Echinacea** (Echinacea species)

*Indications:*
To decrease the severity and duration of URTI — Level 2 evidence[1,31]

*Significant Adverse Effects:*
Hepatitis, 3 cases reported to ADRAC[31]
May cause an overactive immune response in patients with HIV, SLE, active TB, multiple sclerosis[1]
Increased risk of allergic reactions (anaphylaxis, acute asthma) in patients with a history of asthma, atopy, allergic rhinitis[27, 38]

*Drug Interactions*:
May interfere with immunosuppressant drugs (e.g. steroids)[1, 27]
May potentiate hepatotoxicity in combination with anabolic steroids, amiodarone, methotrexate and ketoconazole[36]

*Cessation prior to Surgery:*
Discontinue as far in advance as possible in patients with liver dysfunction[27, 37]

**Valerian** (Valeriana officinalis)

*Indications:*
Sedation, insomnia Level 2 evidence[1, 36]

*Significant Adverse Effects:*
Withdrawal syndrome (similar to benzodiazepine)[27]

*Drug Interactions:*
Potentiates the sedative effects of barbiturates, benzodiazepines, anaesthetics[27, 36]

*Cessation prior to Surgery:*
If possible, taper dose weeks before surgery.[27, 37] If not, continue medication during the perioperative period[37]

**Horseradish** (Armoracia rusticana)

*Indications:*
Cough, bronchitis Level 3 evidence[39]

*Significant Adverse Effects:*
None[39]

*Drug Interactions:*
None known[39]

*Cessation prior to Surgery:*
Probably not required

**Ginseng** (Panax ginseng)

*Indications:*
Treatment of Type 2 Diabetes (decrease glucose levels and weight loss, improved subjective ratings) Level 2 evidence[36]
To improve stamina, vigilance, well-being (No proven benefit) Level 2 evidence[31]

*Significant Adverse Effects:*
Hypoglycaemia in fasting patients, Steven Johnson syndrome, increased risk of bleeding, nervousness, insomnia[27, 36]

*Drug Interactions:*
Increased bleeding risk with aspirin, NSAID, anticoagulants[36]
Increased risk of headaches, tremulousness, mania with phenelzine[36]

*Cessation prior to Surgery:*
At least 7 days before surgery[27, 37]

**Cranberry** (Vaccinium macrocarpon)

*Indications:*
Treatment and prevention of urinary tract infections — Level 3 evidence[40,41]

*Significant Adverse Effects:*
None

*Drug Interactions:*
Increased absorption of Vitamin B12 in patients taking omeprazole[41]

*Cessation prior to Surgery:*
Probably not required

**Aloe** (Aloe vera)

*Indications:*
Psoriasis — Level 3 evidence[42]
Wound healing (No proven benefit) — Level 3 evidence[42, 43]
Constipation — Level 4 evidence[42]

*Significant Adverse Effects:*
Topical/Oral — hypersensitivity reactions, dermatitis[42, 43]
Oral — hypokalaemia[42]

*Drug Interactions:*
May potentiate hypoglycaemia in combination with glibenclamide[43]

*Cessation prior to Surgery:*
Probably not required

## APPENDIX 2: PERIOPERATIVE IMPLICATIONS OF HERBAL MEDICINES

**Adverse Cardiovascular System Effects**

*Ginseng:* has been noted to cause tachycardia and hypertension during anaesthesia[1, 4]
*Ephedra (Ma-Huang):* associated with numerous reports of cerebral bleeds, CVA, panic attacks, palpitations, arrhythmia, chest pain and myocardial infarction[1, 4, 15, 33]

**Adverse Coagulation Effects**

Ginger, Garlic, Gingko, Ginseng*, and Feverfew:* all possess antiplatelet activity and there have been reports of unanticipated excessive surgical bleeding with them[1, 7, 8, 15]
Theoretical increase risk of epidural haematoma following neuraxial blockade, although no reported cases so far[1]
*Dong quai:* contains coumarin, has caused increased INR level and widespread bleeding[44, 45]

**Renal/Hepatic/Electrolyte Disturbance Effects**

*Liquorice:* due to its mineralocorticoid effect it may exacerbate renal insufficiency and has caused hypertension, hypokalaemia, sodium retention and oedema[4, 7]
*Goldenseal:* is an aquaretic (not a diuretic), thus it may decrease the potential benefits of anti- hypertensive medications[4]
*Echinacea:* may be hepatotoxic with chronic usage[4]

**Adverse Immunological Effects**

*Echinacea:* (long term use >8 weeks) could cause immunosuppression which may result in poor wound healing and increased infection risk[1]

**Prolongation of Anaesthesia**

*Kava:* potentiates barbiturate and benzodiazepine effect, resulting in prolonged sedation[1]

*Valerian:* via its effect upon the GABA neurotransmitter, potentiates the sedative effects of both barbiturates and benzodiazepines[1, 4]

## APPENDIX 3: DEFINITIONS/KEY TERMS

**Complementary and Alternative Medicines (CAM):** a term used to describe non orthodox (i.e. neither prescription nor "over the counter" regulated medications) consisting of herbs, dietary supplements, vitamins and homeopathic medicines

**Dietary Supplement:** a product not represented as a conventional food or as a sole meal item that increases the total dietary intake. They may contain parts, entire molecules or concentrated forms of vitamins, minerals, proteins or enzymes[46]

**Herbs:** plant-derived product used for medicinal or health purposes

**Nutraceuticals:** herbal medicines and other supplements

## REFERENCES


1. Hodges PJ, Kam PC. The peri-operative implications of herbal medicines. *Anaesthesia* 2002; 57:889-899.
2. Norred CL. Complementary and alternative medicine use by surgical patients. *AORN J* 2002; 76:1013-1021.
3. Winslow LC, Kroll DJ. Herbs as medicines. *Arch Intern Med* 1998; 158:2192-2199.
4. Leak JA. The perioperative hazards of nutraceuticals. In: Annual refresher course lectures-American Society of Anesthesiologists. Lippincott, Philadelphia, USA, 2001; 176 (1-5).
5. Soosay I. Western herbal medicine — an introduction. *Current Therapeutics* 2000; 41(6):90-95.
6. Ernst E, Pittler MH. Herbal Medicine. *Med Clin North Am* 2002; 86:149-161.
7. Leak JA. Perioperative considerations in the management of the patient taking herbal medicines. *Curr Opin Anaesthesiol* 2000; 13:321-325.
8. Leak JA. The perioperative hazards of herbs and supplements. In: Annual refresher course lectures-American Society of Anesthesiologists. Lippincott, Philadelphia, USA, 2002; 176 (1-7).
9. American Society of Anesthesiologists. What you should know about herbal use and anesthesia. American Society of Anesthesiologists; www.ASAhq.org accessed January 15, 2003.
10. MacLennan AH, Wilson DH, Taylor AW. The escalating cost and prevalence of alternative medicine. *Prev Med* 2002; 35:166-173.
11. Tsen LC, Segal S, Pothier M, Bader AM. Alternative medicine use in presurgical patients. *Anesthesiology* 2000; 93:148-151.
12. Skinner CM, Rangasami J. Preoperative use of herbal medicines: a patient survey. *Br J Anaesth* 2002; 89: 792-795.
13. Lennox PH, Henderson CL. Herbal medicine use is frequent in ambulatory surgery patients in Vancouver Canada. *Can J Anesth* 2003; 50:21-25.
14. Cohen RJ, Ek K, Pan CX. Complementary and alternative medicine (CAM) use by older adults: a comparison of self-report and physician chart documentation. *J Gerontol A Biol Sci Med Sci* 2002; 57: 223-227.
15. McCulloch T, Wilson L. "Which doctors?": — complementary and alternative medicines. ANZCA *Bulletin* 2001; 10:33-36.
16. Australian Medical Association. Position statement — complementary medicine. *Australian Medicine* 2002; March: 8.

17. Drew AK, Myers SP. Safety issues in herbal medicine: implications for health professions. *Med J Aust* 1997; 166:538-541.
18. Therapeutic Goods Administration — Office of Complementary Medicines. Information on complementary medicines. Therapeutic Goods Administration; www.health.gov.au/tga/cm accessed March 5, 2003.
19. Hepner DL, Harnett M, Segal S, Camann W, Bader AM, Tsen LC. Herbal medicine use in parturients. *Anesth Analg* 2002; 94:690-693.
20. Leung JM, Dzankic S, Manku K, Yuan S. The prevalence and predictors of the use of alternative medicine in presurgical patients in five California hospitals. *Anesth Analg* 2001; 93:1062-1068.
21. Flanagan K. Preoperative assessment: safety considerations for patients taking herbal products. *J Perianesth Nurs* 2001; 16:19-26.
22. Crowe S, Fitzpatrick G, Jamaluddin MF. Use of herbal medicines in ambulatory surgical patients. *Anaesthesia* 2002; 57:203-204.
23. Kaye AD, Clarke RC, Sabar R et al. Herbal medicines: current trends in anesthesiology practice — a hospital survey. *J Clin Anesth* 2000; 12:468-471.
24. Norred CL, Zamudio S, Palmer SK. Use of complementary and alternative medicines by surgical patients. *AANA J* 2000; 68:13-18.
25. Norred CL. A follow-up survey of the use of complementary and alternative medicines by surgical patients. *AANA J* 2002; 70:119-125.
26. Wang SM, Peloquin C, Kain ZN. Attitudes of patients undergoing surgery toward alternative medical treatment. *J Altern Complement Med* 2002; 8:351-356.
27. Ang-Lee MK, Moss J, Yuan CS. Herbal medicines and perioperative care. *JAMA* 2001; 286:208-216.
28. Smith A. It's natural so it must be safe. *Aust Prescr* 2002; 25:50-51.
29. Briggs DR. The regulation of herbal medicines in Australia. *Toxicol* 2002; 181-182:565-570.
30. Murphy JM. Preoperative considerations with herbal medicines. *AORN J* 1999; 69:173-182.
31. Ernst E. The risk-benefit profile of commonly used herbal therapies: Gingko, St John's Wort, Ginseng, Echinacea, Saw Palmetto and Kava. *Ann Intern Med* 2002; 136:42-53.
32. Pinn G. Herbal medicine — an overview. *Aust Fam Physician* 2000; 29:1059-1062.
33. Kam PC, Liew S. Traditional Chinese herbal medicine and anaesthesia. *Anaesthesia* 2002; 57:1083-1089.
34. American Society of Anesthesiologists. Anesthesiologists warn: if you're taking herbal products, tell your doctor before surgery. American Society of Anesthesiologists; www.ASAhq.org accessed January 15, 2003.
35. MacLennan AH, Wilson DH, Taylor AW. Prevalence and cost of alternative medicine in Australia. *Lancet* 1996; 347:569-573.
36. Miller LG. Herbal medicinals. *Arch Intern Med* 1998; 158:2200-2211.
37. Easton J. New recommendations proposed for safe use of herbal medications by patients having surgery. EurekAlert; www.eurekalert.org.pub accessed June 18, 2002.
38. Mullins RJ, Heddle R. Adverse reactions associated with echinacea: the Australian experience. *Ann Allergy Asthma Immunol* 2002; 88:42-51.
39. Gruenwald J, Brendler T, Jaenicke C, eds. Physician Desk Reference for Herbal Medicines. 2nd Ed. Medical Economics Company, Montvale, USA, 2000; 408-409.
40. Pinn G. Herbal medicine in renal and genitourinary disease. *Aust Fam Physician* 2001; 30:974-977.
41. Drew A. Cranberry. Current Therapeutics 2000; 41(10):80-81.
42. Gruenwald J, Brendler T, Jaenicke C, eds. Physician Desk Reference for Herbal Medicines. 2nd Ed. Medical Economics Company, Montvale, USA, 2000; 16-20.
43. Drew A. Aloe. *Current Therapeutics* 2000; 41(9):80-81.
44. Fugh-Berman A. Herb-drug interactions. *Lancet* 2000; 355:134-138.
45. Myers SP. Interactions between complementary medicines and warfarin. *Aust Prescr* 2002; 25:54-56.
46. Wren KR, Kimbrall S, Norred CL. Use of complementary and alternative medications by surgical patients. *J Perianesth Nurs* 2002; 17:170-177.

# Monitoring Cerebral Oxygenation: Recent advances

DR BALASUBRAMANIAN VENKATESH, MB, BS, MD(IntMed), FRCA, FFARCSI, MD, FJFICM
Associate Professor in Intensive Care, Royal Brisbane Hospital, Herston 4029, Brisbane, Queensland

DR ANDREA BEINDORF
Research Fellow in Intensive Care, Royal Brisbane Hospital, Herston 4029, Brisbane, Queensland

*Professor Bala Venkatesh has research interests in tissue oxygenation, aptoptosis, acid-base physiopathology, adrenal dysfunction, burns and neurointensive care. He has published widely in international journals. Dr Andrea Beindorf is an Intensive Care Medicine Specialist from Italy. She is currently a research fellow at Royal Brisbane Hospital and has a special interest in neurointensive care.*

## INTRODUCTION

The goals of head injury management are limitation of disability and reduction of mortality. Although the primary damage that occurs at the time of the injury cannot be modified, secondary brain insults contribute to the extension of the area of injury. Many studies have also demonstrated that, after traumatic brain injury, the cerebral autoregulation is impaired and the injured brain is more susceptible to secondary insults. Therefore understanding the mechanisms of secondary brain insults is vital to the institution of appropriate therapy. A landmark study by Chesnut et al confirmed the importance of hypoxia and hypotension, the two most common causes of secondary insults, in determining outcome from neurotrauma.[1] The final common pathway of secondary insult mediated brain damage appears to be through reductions in cerebral oxygen delivery. In support of this concept are data from other studies on cerebral blood flow and brain tissue oxygen measurements in patients with head injury.[2, 3] Monitoring cerebral perfusion and oxygenation is therefore becoming commonplace in neurosurgical critical care practice.

The purpose of this review is to examine the physiology and pathophysiology of cerebral oxygenation, review the various modes of monitoring cerebral oxygenation, critically review the literature concerning their use in day to day intensive care practice, outline their limitations and define possible indications for their use.

## PHYSIOLOGY AND PATHOPHYSIOLOGY OF CEREBRAL OXYGENATION AND ISCHAEMIA

The brain has unique anatomical and physiological characteristics, which render it susceptible to ischaemia. The predominant substrates of cerebral energy metabolism are glucose and oxygen.[4] The adult brain has a blood supply of 50 ml/100 g/min, an oxygen consumption of 45 ml/min and a $CMRO_2$ of 3.5 ml/100 g tissue/min. Oxygen extraction by the brain is slightly greater than the whole body at rest as reflected by the

higher arterio-venous difference of 5-7 ml% as compared to 3-5 ml% for the whole body. There is differential vulnerability of the various parts of the brain to hypoxia. For example, the neurons are more vulnerable than the glial tissue, the cerebral cortex more than the brain stem and the grey matter more than the white matter.[5] Furthermore, remote anatomical connections through axons and dendrites extend the injury from the primary focus. The closed cranial cavity and the absence of a lymphatic circulation increase the risk of infarction of the ischaemic penumbra and cerebral oedema under conditions of intracranial hypertension.[6] The combination of small glucose reserves and high oxygen extraction limits the ability to sustain brain energy metabolism to about 20 seconds during total cessation of cerebral flow. Thus, even brief interruptions in cerebral blood flow produce disruption of cerebral autoregulation and EEG and biochemical changes within the brain.[7-9] The end result of ischaemia and cerebral hypoxaemia is lactic acidosis, ionic changes within the cell, excitotoxicity and free radical formation.

## FACTORS DETERMINING CEREBRAL OXYGENATION

The three main factors determining cerebral oxygenation are cerebral blood flow (CBF), arterial oxygen content ($CaO_2$) and cerebral metabolic rate of oxygen consumption ($CMRO_2$). A variety of monitoring modalities are available and are listed in Table 1. They include:

1. Systems providing a global measurement of cerebral oxygenation;
2. Systems providing a regional measurement of cerebral oxygenation; and,
3. Systems measuring cerebral metabolism.

**Table 1**
Classification of cerebral oxygenation monitors

| |
|---|
| **A. Global cerebral oxygenation monitors** |
| 1. Intracranial pressure (ICP) and cerebral perfusion pressure (CPP) |
| 2. Measurement of cerebral blood flow (CBF) |
| 3. Transcranial Doppler ultrasonography |
| 4. Jugular venous oximetry |
| 5. Cerebrospinal fluid gas tensions |
| **B. Regional cerebral oxygenation monitor** |
| 1. Near infra red spectroscopy (NIRS) |
| 2. Intraparenchymal probes |
| 3. Laser Doppler flowmetry |
| **C. Monitors of cerebral metabolism** |
| 1. Cerebral microdialysis |
| 2. Cerebral bioenergetics |

### Global cerebral oxygenation monitors

*Intracranial pressure (ICP) and cerebral perfusion pressure (CPP)*

Cerebral blood flow is an important determinant of cerebral oxygen delivery. Its measurement with nitrous oxide or the Xenon 133 radiotracer is cumbersome and not practical in the ICU. The cerebral perfusion pressure is frequently used as an alternative for CBF measurements. CPP is calculated as the difference between the mean arterial pressure (MAP) minus the ICP. Reliable measurements of ICP are required to calculate CPP. To measure ICP reliably, a fluid filled intraventricular catheter or a solid state fiberoptic device is required. Fluid filled systems provide

accurate data, facilitate CSF drainage to decompress the ventricular system and are considered the "gold standard" to measure ICP. However, they have limited frequency response and present a risk of infection.[10, 11] Solid state fiberoptic devices produce highly accurate data initially, are less invasive and easier to position, but do not allow CSF drainage and are more expensive.[12] The additional benefits of measuring ICP include the ability to analyze the arterial and ICP waveforms to evaluate the shift in amplitude and frequency and to determine intracranial compliance. Alterations in the ICP amplitude and pressure index have been shown to correlate with neurological outcome.[13]

With decreased intracranial compliance, small changes in intracranial volume can induce gross changes in ICP and therefore CPP. The relationship between intracranial pressure and volume is given by the pressure-volume index (PVI), which is defined as the change in intracranial volume that produces a ten-fold rise in ICP.[14] The normal value is about 26 ml, but may be lower in patients with head injury. Thus, small reductions in intracranial volume may result in significant decrease in ICP. This may be achieved by mannitol or hypertonic saline which reduce cerebral swelling. Manoeuvres such as hyperventilation reduce cerebral blood volume, producing an acute reduction in intracranial volume and thus ICP.

The attraction of using CPP as an index of cerebral oxygenation is its linear relationship to CBF. Based on available data (which are predominantly Class II or Class III), an optimal CPP of 60-70 mmHg has been proposed in a number of studies with an increased likelihood of poor outcome if it is not maintained above this threshold.[15]

*Caveats when interpreting ICP and CPP*

The use of CPP as an index of CBF does not take into account the heterogeneity of CBF in head injury and the variability of cerebrovascular resistance. A "target CPP" may not guarantee adequate low CBF in the setting of vasospasm or other intracranial or carotid artery disease. Therapeutic measures designed to maintain an optimum CPP may by themselves induce cerebral ischaemia or produce serious systemic side-effects such as ARDS of high dose inotropes.[16] The presence of patient dependent variations in the level of the optimum ICP should also be considered. ICP waveform analysis is a complex process and remains largely a research tool. Also, as CPP and CBF are haemodynamic variables, ideally they are interpreted in conjunction with arterio-venous differences for oxygen. The relationship between CBF and $CMRO_2$ is given by the cerebral extraction of oxygen. This forms the basis for jugular venous oximetry, which is discussed in detail below.

*Transcranial Doppler (TCD)*

TCD allows the measurement of blood flow velocity in the basal cerebral arteries using three naturally occurring acoustic windows — transtemporal, transorbital and transforaminal.[17] The major cerebral arteries can be insonated using the transtemporal approach. The systolic, diastolic and the mean flow velocity are measured, while the pulsatility index is a derived variable, calculated as the difference between the systolic and the diastolic flow velocity divided by the mean flow velocity. A range of flow velocities has been described, depending on the artery and the age of the patient. A raised velocity, particularly in the setting of traumatic subarachnoid haemorrhage, may be indicative of vasospasm. Other useful information provided by TCD includes:

(1) cerebrovascular reactivity utilizing the response of flow velocity to changes in $CO_2$ concentration;
(2) spasm and hyperaemia, based upon the comparison between extracranial and intracranial flow velocity;
(3) an assessment of CPP using the pulsatility index; and,
(4) detection of brain death.[18]

The limitations of TCD include a variable relationship between velocity and flow,[19] problems of long term fixation of the ultrasound probe, inter- and intra-observer variability in the measurement process,[20] and marked moment to moment variability of velocity patterns in both volunteers and in patients. The latter may limit the usefulness of intermittent transcranial doppler.

*Jugular venous oximetry*

Jugular bulb oxyhaemoglobin saturation monitoring ($SjO_2$) has been widely purported to reliably assess the adequacy of the cerebral blood flow. The jugular venous saturation is directly proportional to CBF and arterial oxygen saturation ($SaO_2$) and inversely proportional to cerebral metabolic rate of oxygen ($CMRO_2$). It is measured using a catheter inserted into the dominant jugular bulb. The correct position of the catheter can be assessed by a lateral skull X-ray. The technique of insertion and calibration has been described by Andrews et al.[21] $SjO_2$ can be monitored continuously by a fiberoptic catheter. If $SaO_2$ remains constant and $CMRO_2$ is assumed to be fixed, then changes in $SjO_2$ are proportional to changes in cerebral blood flow. Normal values range from 55-75%.

Despite the availability of data suggesting a relationship between jugular desaturations and poor outcome,[22] there is no evidence to suggest that prevention and treatment of jugular venous desaturation improves outcome. Consequently, there has been a waning of enthusiasm for this mode of monitoring the last few years. Several other factors have also contributed to this loss of support. Investigations by Latronico et al. and Stocchetti et al. have demonstrated significant differences in measurement between the two sides,[23, 24] thus raising the question of which side should be monitored. The invasive nature of the technique, the potential for erroneous readings resulting from catheter malposition, impaction and thrombus formation, and the likelihood of extracerebral contamination of jugular venous blood flow have also mitigated against its use.[25-27] The $SjO_2$ data reflect global cerebral oxygenation and may miss important regional changes, as has been shown in comparison with brain tissue oxygen tension ($P_bO_2$).[28, 29] Changes in $SjO_2$ are an index of CBF only if $CMRO_2$ is assumed to be constant. This is not the case, however, in patients with neurotrauma, as they often have fever or seizures which will influence the $CMRO_2$.

*Measurement of cerebrospinal fluid (CSF) gas tensions and pH*

CSF is produced by the choroid plexus of the cerebral ventricles, circulates though the lateral, third and the fourth ventricles and enters the subarachnoid space through the Foramina of Luschka and Magendie. It circulates on the cortical surface of the brain in the subarachnoid space and is absorbed into the venous sinuses through the Pachionian granulations. Given its production in the "core" of the brain and its circulation through various compartments of the cranial cavity, it would seem logical that measurement of its gas tensions would reflect those of the brain. Venkatesh et al have demonstrated that aspiration of CSF and measurement of its gas tensions in a

blood gas analyzer produces inaccurate results.[30] The only reliable way of measuring CSF gas tensions accurately is to use a continuous gas sensor. Venkatesh et al also demonstrated the feasibility of measuring CSF gas tensions continuously using such a gas sensor inserted through an intraventricular drain in patients with neurotrauma.[31] However, they reported the potential for inaccurate results resulting from a clot or devitalized tissue present in the drain tip. The technique provides only a global index of brain oxygenation and necessitates the positioning of an external ventricular drain.

**Regional cerebral oxygenation monitors**

*Near infrared spectroscopy (NIRS)*

NIRS is a non-invasive technique developed for the continuous measurement of oxyhaemoglobin, reduced haemoglobin and cytochrome aa3 in brain tissue.[32] It is based on the principle of absorption or transmission spectroscopy of infrared light. The attenuation of the incident beam of light can be attributed almost solely to absorption by oxyhaemoglobin, reduced haemoglobin and cytochrome aa3. The extent of light absorption is governed by the principles stated in the Beer & Lambert law.

Owing to the small size of the cranium and the thin vault, the technique has found more application in neonates than in adults.[33] In adults, because of a longer path length of light, there is greater scattering of light, preventing adequate transmission to the opposite side of the skull.[34] This necessitates sampling of scattered light by a probe placed ipsilateral to the source probe, resulting in a loss of resolution. The contribution of scalp blood volume to the overall absorption of light needs to be excluded from the final equation. The proportion of the various vascular compartments contributing to light absorption is only an estimate, based on previous studies and not an accurate measure.

Whilst studies have demonstrated the usefulness of NIRS in assessing the impact of respiration, changes in $PCO_2$ and carotid cross clamping on cerebral oxyhaemoglobin and reduced haemoglobin, its place in the monitoring of patients with head injury remains to be established.[33, 35] Studies comparing the sensitivity of NIRS with jugular venous oximetry to detect early cerebral hypoxia have produced conflicting results. NIRS requires further refinement and validation before recommending its routine use in clinical practice.

*Brain tissue pH, $PCO_2$ and $PO_2$ measurement*

The miniaturization of the Clark electrode and fiberoptic systems (optodes) has permitted the measurement of gas tensions and pH in tissues. These probes can be placed in the brain surgically either at the time of craniotomy or through a burr hole. For the first time, this technology has enabled us to measure tissue gas tensions in clinical practice. At the present time two systems are available for measurement of brain oxygen tension. Paratrend 7 sensor is a hybrid electrode-optode system and is capable of measuring pH, $PCO_2$, $PO_2$ and temperature. It has been validated for use in bench studies, experimental animal models and in critically ill patients.[31, 36] The Licox sensor is a Clark electrode with a thermocouple to facilitate temperature compensation.[37] The range for normal brain $PO_2$ has been reported to be 10-20 torr. In addition to measuring the absolute values of tissue gas tension, these systems also allow the derivation of "oxygen and carbon dioxide reactivity", i.e. the change in tissue $PO_2$ and $PCO_2$ for a given change in arterial oxygenation or ventilation.[38] This is used as a measure of cerebral autoregulatory status and a loss of reactivity correlates with

poor neurological outcome. A number of studies to date demonstrate that brain $PO_2$ measurement is a sensitive method of tracking cerebral oxygenation in neurotrauma.[2] Furthermore, few adverse events related to these invasive probes have been reported.

There are a number of unanswered questions with these devices with regard to tissue oxygenation monitoring. Firstly, their insertion may be associated with tissue trauma. It has been previously demonstrated that the presence of clot and devitalized tissue may interfere with sensor measurements.[31, 39] Secondly, in vivo calibration is not possible as the "true" tissue gas tension is not known. Thirdly, the scanning area for these sensors is variable, reported to be 20 $mm^2$ for the Licox, which may not be large enough to detect ischaemia in the entire ischaemic penumbra. It has been shown that, on the cortical surface, brain $PO_2$ varies markedly over a distance of a few millimetres.[40] The issues of infection risk and costs also need to be considered.

*Laser Doppler flowmetry (LDF)*

LDF provides a continuous measure of the surface blood flow over the cerebral cortex. The principle behind this technique is the conversion of the Doppler shift present in reflected light to provide an index of microcirculatory flux (the product of red cell concentration and red cell velocity). This method has been validated for measurement of regional cerebral blood flow.[41] Besides real time measurement of cerebral microvascular blood flow, the influence of physiological and pharmacological stimuli on microcirculatory flow can be observed in detail. The limitations of the technique are the small area of tissue scanned, semiquantitative assessment of CBF, the need for surgical placement of cortical probes, the tendency for movement artifacts to influence signal quality and the limited experience with head injured patients.

**Monitors of cerebral metabolism**

*Cerebral Microdialysis*

This relatively new technique enables prolonged measurements of extracellular fluid metabolites in the brain tissue.[42] One or more microdialysis catheters can be inserted through the same burr hole as an intraventricular catheter. The method uses an internally perfused semipermeable membrane probe, which allows water-soluble substances such as lactate, glucose, amino acids and electrolytes to be collected for analysis outside the brain. Normal saline is used as the perfusate. The results of the microdialysis can be influenced mainly by the length of the catheter and by the perfusion fluid and its rate. In combination with a measure of cerebral blood flow, the method provides a novel approach to studying the relationship between cerebral perfusion and metabolism.[43] Fluctuations in concentrations of extracellular fluid metabolites may provide evidence of ischaemic damage[43] (glucose, lactate, pyruvate, K, ADP) or excitotoxicity (glutamate, aspartate),[44] or alteration in the phospholipids membrane caused by oxygen radicals.

The limitations of the microdialysis technique are its invasiveness, its ability to sample a small volume of cortical tissue and the potential for introduction of infection.[45] Until data demonstrating an improvement in outcome using microdialysis generated endpoints become available, it can only considered a research tool.

*Monitoring cellular energetics*

With development in technology, emphasis is shifting from monitoring global parameters to measurement of substrate utilization at the level of the cell. The end

product of carbohydrate metabolism is the generation of ATP (38 molecules of ATP per molecule of glucose metabolized via the Krebs cycle). In the brain and muscle, a portion of this ATP is converted to creatine phosphate (CP). High-energy phosphate levels can be measured using freeze clamping, NIR spectroscopy and NMR spectroscopy.[46, 47]

Freeze clamping involves invasive tissue sampling using a chilled instrument, rapid cooling and measurement of metabolites in the specimen. The technique is invasive, time consuming and does not lend itself to continuous measurement. NIR spectroscopy facilitates the measurement of tissue oxidized haemoglobin and cytochrome aa3. This technique has been described in detail in a previous section. $P^{31}$ NMR spectroscopy allows the non-invasive measurement of high-energy phosphates within the cell.[47] Refinement in this technology has permitted the measurement of brain metabolism in animals and humans.

Although they represent exciting developments in the area of cellular energetics, owing to their bulk, cost and potential electromagnetic interference with other ICU monitoring systems, the above techniques remain largely research tools with limited clinical application.

## UNANSWERED ISSUES IN CEREBRAL OXYGENATION MONITORING

Despite a number of studies demonstrating potential usefulness of all the monitoring modalities discussed, there are still a number of unanswered questions with regard to the information generated. What is an appropriate level of brain $PO_2$ or $SjO_2$ to be maintained? Recently, data have emerged that apoptosis is a significant pathological process following neurotrauma and one which may have an impact on outcome. Critical $PO_2$ thresholds exist at which apoptosis may be triggered but these need to be identified in humans.[48] If there is a decline in $SjO_2$ from 75% to 65%, is there a need to intervene even though this is a decrement within the normal range? What is the relationship between changes in $SjO_2$ and brain $PO_2$? Is it the white matter or the gray matter or the gray white interface $PO_2$ that is most important? What is the safety record of the newer devices? For example, no single trial has been large enough to detect a complication rate of 1 in 1000. How long can tissue oxygen devices be left in place? One of the other problems is the lack of Class I data on the benefits of ICP monitoring, the appropriate ICP treatment threshold or the newer monitoring modalities. The above issues need to be clarified in large randomized controlled trials before drawing conclusions on usefulness, longevity of use and safety.

## MONITORING CEREBRAL OXYGENATION — A REALISTIC OR AN ELUSIVE GOAL?

The ability to monitor cerebral oxygenation, either intermittently or continuously, is now possible. Some of the techniques, such as ICP, CPP, TCD and $SjO_2$ measurement are commonly used. Others are still primarily research tools. The complex and as yet incompletely understood physiology of the injured brain necessitates certain assumptions to be made when these devices are used. In order to overcome some of these limitations, there has been a shift towards multimodal monitoring to provide an increased power of interpretation. Clearly, some of the currently available techniques need further refinement and evaluation before justifying their routine use in clinical practice. The other requirements for the use of these systems are safety and that

management based on the information provided should lead to an improved patient outcome. The first step has certainly been made in that direction with the generation of more accurate data. With improved measurement and monitoring techniques, determination of critical tissue $PO_2$ thresholds for neurological recovery and thus titration of therapy to this level of tissue $PO_2$ becomes a possibility.

## REFERENCES


1. Chesnut RM, Marshall LF, Klauber MR et al. The role of secondary brain injury in determining outcome from severe head injury. *J Trauma* 1993; 34:216-222.
2. van Santbrink H, Maas AI, Avezaat CJ. Continuous monitoring of partial pressure of brain tissue oxygen in patients with severe head injury [see comments]. *Neurosurgery* 1996; 38:21-31.
3. Kelly DF, Martin NA, Kordestani R et al. Cerebral blood flow as a predictor of outcome following traumatic brain injury. *Journal of Neurosurgery* 1997; 86:633-641.
4. Schurr A, Payne R, Levy R, Rigor B. Understanding cerebral energy metabolism: a key to successful neuroprotection. *Clinical Anaesthesiology* 1996; 10:409-425.
5. Graham D. Hypoxia and vascular disorders. In: Adams J, Duchen L, Arnold E, editors. *Greenfield's Textbook of Neuropathology* 1992; p. 156-170.
6. Obrenovitch T. The ischaemic penumbra: twenty years on. Cerebrovascular and Brain Metabolism Reviews 1995; 7:297-323.
7. Kirsch WM, Leitner JW. Glycolytic metabolites and co-factors in human cerebral cortex and white matter during complete ischaemia. *Brain Res* 1967; 4:358-368.
8. Murkin JM. Cerebrovascular reactivity after transient circulatory arrest. *Stroke* 1992; 23 (Suppl 1):3.
9. Rossen R, Kabat H, Anderson JP. Acute arrest of cerebral circulation in man. *Arch Neurol Psychiat* 1943; 50:510-528.
10. Gaab M, Heissler H, Erhardt K. Physical characteristics of various methods of measuring ICP. In: Hoff J, Betz A, editors. Intracranial pressure. Berlin: Springer-Verlag; 1989. p. 16-21.
11. Guyot LL, Dowling C, Diaz FG, Michael DB. Cerebral monitoring devices: analysis of complications. *Acta Neurochir Suppl (Wien)* 1998; 71:47-49.
12. Lang EW et al. Intracranial pressure. Monitoring and management. *Neurosurg Clin N Am* 1994; 5:573-605. Review.
13. Chopp M, Portnoy H. Systems analysis of intracranial pressure. Comparison with volume pressure test and CSF- pulse amplitude analysis. *J Neurosurgery* 1980; 53:516-527.
14. Ursino M, Lodi C, Rossi S, Stocchetti N. Estimation of the main factors affecting ICP dynamics by mathematical analysis of PVI tests. *Acta Neurochir Suppl (Wien)* 1998; 71:306-309.
15. Foundation TBT. Guidelines for cerebral perfusion pressure. *J Neurotrauma* 2000; 17:507-511.
16. Robertson CS, Valadka AB, Hannay HJ et al. Prevention of secondary ischemic insults after severe head injury. *Crit Care Med* 1999; 27:2086-2095.
17. Aaslid R, Huber P, Nornes H. Evaluation of cerebrovascular spasm with transcranial Doppler ultrasound. *Journal of Neurosurgery* 1984; 60:37-41.
18. Babikian VL, Feldmann E, Wechsler LR et al. Transcranial Doppler ultrasonography: year 2000 update. *J Neuroimaging* 2000; 10:101-115.
19. Newell DW. Transcranial Doppler measurements. *New Horiz* 1995; 3:423-430.
20. Shen Q, Stuart J, Venkatesh B. Inter observer variability of the transcranial doppler ultrasound technique. Impact of lack of practice on the accuracy of measurement. *J Clin Monit Comput* 1999; 15:179-184
21. Andrews PJ, Dearden NM, Miller JD. Jugular bulb cannulation: description of a cannulation technique and validation of a new continuous monitor. *Br J Anaesth* 1991; 67:553-558.
22. Gopinath S, Robertson C, Contant C, al e. Jugular venous desaturation and outcome after head injury. *J Neurol Neurosurg Psychiatry* 1994; 57:717-723.
23. Latronico N, Beindorf AE, Rasulo FA et al. Limits of intermittent jugular bulb oxygen saturation monitoring in the management of severe head trauma patients. *Neurosurgery* 2000; 46:1131-1138; discussion 1138-1139.
24. Stocchetti N, Paparella A, Bridelli F, al e. Cerebral venous oxygen saturation studied with bilateral samples in the internal jugular veins. *Neurosurgery* 1994; 34:38-44.
25. Matta BF, Lam AM. The rate of blood withdrawal affects the accuracy of jugular venous bulb. Oxygen saturation measurements. *Anesthesiology* 1997; 86:806-808.

26. Inglis A. Jugular bulb saturation: which side to measure? [letter; comment]. *Anesth Analg* 1995; 81:659-660.
27. Gemma M, Beretta L, De Vitis A et al. Complications of internal jugular vein retrograde catheterization. *Acta Neurochir Suppl (Wien)* 1998; 71:320-323.
28. Kiening KL, Unterberg AW, Bardt TF, Schneider GH, Lanksch WR. Monitoring of cerebral oxygenation in patients with severe head injuries: brain tissue PO2 versus jugular vein oxygen saturation. *J Neurosurg* 1996; 85:751-757.
29. Gopinath SP, Valadka AB, Uzura M, Robertson CS. Comparison of jugular venous oxygen saturation and brain tissue $PO_2$ as monitors of cerebral ischemia after head injury [see comments]. *Crit Care Med* 1999; 27:2337-2345.
30. Venkatesh B, Boots RJ. Carbon dioxide and oxygen pressure measurements in the cerebrospinal fluid in a conventional blood gas analyser: Analysis of bias and precision. *Journal of the Neurological Sciences* 1997; 147:5-8.
31. Venkatesh B, Boots R, Tomlinson F, Jones RD. The continuous measurement of cerebrospinal fluid gas tensions in critically ill neurosurgical patients: a prospective observational study. *Intensive Care Med* 1999; 25:599-605.
32. Elwell C, Matcher S, Tyszczuk L, al e. Measurement of cerebral venous saturation in adults using near infrared spectroscopy. *Adv Exp Med Biol* 1997; 411:453-460.
33. Buchner K, Meixensberger J, Dings J, Roosen K. Near-infrared spectroscopy — not useful to monitor cerebral oxygenation after severe brain injury. *Zentralbl Neurochir* 2000; 61:69-73.
34. Elwell C, Edwards A, al e. Measurement of cerebral blood flow in humans using near infrared spectroscopy — methodology and possible errors. *Adv Exp Med Biol* 1992; 406:235-244.
35. Pollard V, Prough DS, DeMelo AE et al. The influence of carbon dioxide and body position on near-infrared spectroscopic assessment of cerebral hemoglobin oxygen saturation. *Anesth Analg* 1996; 82:278-287.
36. Venkatesh B, Clutton Brock TH, Hendry SP. A multiparameter sensor for continuous intra-arterial blood gas monitoring: a prospective evaluation [see comments]. *Crit Care Med* 1994; 22:588-594.
37. Meixensberger J, Jager A, Dings J, Baunach S, Roosen K. Multimodal hemodynamic neuromonitoring—quality and consequences for therapy of severely head injured patients. *Acta Neurochir Suppl (Wien)* 1998; 71:260-262.
38. Steiner LA, Czosnyka M, Piechnik SK et al. Continuous monitoring of cerebrovascular pressure reactivity allows determination of optimal cerebral perfusion pressure in patients with traumatic brain injury. *Crit Care Med* 2002; 30:733-738.
39. Maas AI, Fleckenstein W, de Jong DA, van Santbrink H. Monitoring cerebral oxygenation: experimental studies and preliminary clinical results of continuous monitoring of cerebrospinal fluid and brain tissue oxygen tension. *Acta Neurochir Suppl (Wien)* 1993; 59:50-57.
40. Leniger-Pollart E, Lubbers DW, Wrabotz W. Regulation of local tissue $PO_2$ of the brain cortex at different $O_2$ partial pressures. *Pflugers Arch* 1975; 359:81-95.
41. Kaiser M, During M. Combining laser Doppler flowmetry with microdialysis: a novel approach to investigate the coupling of regional cerebral blood flow to neuronal activity. *J Neuroscience Methods* 1995; 60:165-173.
42. Hutchinson PJ, O'Connell MT, Al-Rawi PG et al. Clinical cerebral microdialysis: a methodological study. *J Neurosurg* 2000; 93:37-43.
43. Langemann H, Mendelowitsch A, Landolt H, Alessandri B, Gratzl O. Experimental and clinical monitoring of glucose by microdialysis. *Clinical Neurology & Neurosurgery* 1995; 97:149-155.
44. Matsumoto K, Lo E, Pierce A, Halpem E, Newcomb R. Secondary elevation of extracellular neurotransmitter amino acids in the reperfusion phase following focal cerebral ischaemia. *J Cereb Blood Flow & Metab* 1996; 16:114-124.
45. Alvarez del Castillo M. Monitoring neurologic patients in intensive care. *Curr Opin Crit Care* 2001; 7:49-60.
46. Ackerman JJH GT, Wong GG et al. Mapping of metabolites in whole animals by 31P NMR using surface coils. *Nature* 1980; 167-170.
47. Cady EB CA, Dawson MJ et al. Non invasive investigation of cerebral metabolism in newborn infants by phosphorus nuclear magnetic resonance spectroscopy. *Lancet* 1983; 1059-1062.
48. Venkatesh B, Gobe G, Morgan T. Critical tissue oxygen thresholds for the induction of apoptosis in critical illness: A review. In: JL V, editor. *2002 Yearbook of Intensive Care and Emergency Medicine*. Berlin Heidelberg: Springer-Verlag; 2003. p. Accepted for publication.

# Transcranial Doppler Ultrasound

DAVID WILLIAMS, MB, ChB, FANZCA
Neuroanaesthesia Fellow, Royal Adelaide Hospital

GUY LUDBROOK, MB, BS, FANZCA, PhD
Professor and Head, Departments of Anaesthesia. University of Adelaide and Royal Adelaide Hospital

*David Williams completed his training at the Royal Adelaide Hospital. His provisional fellowship incorporated neuroanaesthesia with clinical research into the cerebral pharmacokinetics of intravenous agents. Transcranial Doppler was used both clinically and as a research tool. He plans to return to Waikato Hospital, Hamilton, NZ in 2004 as a Staff Specialist.*

*Guy Ludbrook was the Douglas Joseph Professor of the College for 2001-2002. His interests include neurologic monitoring and experimental neuropharmacology. His clinical activities include trauma and neuroanaesthesia. He is actively involved in clinical teaching and is a college examiner.*

## INTRODUCTION

Transcranial Doppler ultrasonography (TCD) was introduced into clinical practice in 1982.[1] It provides a relatively cheap, portable, non-invasive and continuous estimate of cerebral blood flow (CBF) based on blood velocity.[2] TCD is particularly useful for demonstrating relative changes in flow, although technical limitations must be considered when interpreting results. Due to increasing expertise, an expanding body of literature and evolving technology, the clinical role of TCD is becoming more clearly defined.

## PRINCIPLES

An understanding of the underlying principles provides a foundation for the appropriate application and interpretation of TCD. Range gated, pulsed wave, low frequency ultrasound is used to calculate flow velocity within selected intracranial and extracranial arteries. Typically, two MHz ultrasound is used to penetrate the bony vault via specific acoustic "windows".[3] Whilst blood velocity can be readily recorded, it should be recognised that anatomical images have relatively poor resolution. The probe emits pulsed wave ultrasound to an area of enquiry 10-15 mm in diameter, referred to as the sample volume. The sample volume is a known and adjustable distance from the probe, giving rise to the term "range-gating". The frequency shift measured in reflected ultrasound allows calculation of flow velocity in this window using the Doppler principle.

Data undergoes fast Fourier transformation and is displayed as a spectral waveform with height proportional to velocity. Thus, the displayed image looks similar to an arterial waveform. The preferred units are cm/s rather than Hertz (referring to change in frequency). Considered more physiologic, cm/s allows measurements from

instruments using different emission frequencies to be compared.[3] The depth, sample volume, power, gain and velocities are usually numerically displayed. A number of parameters (e.g. Gosling pulsatility index) have been derived to estimate factors such as downstream resistance.

**Flow velocity**

Under normal circumstances, the calibre of the major (or basal) cerebral arteries insonated by TCD is assumed to remain constant.[4, 5] Therefore, the measured velocity is proportional to flow. Clinically, the diameter of insonated vessels is unknown but proportional changes in velocity reflect proportional changes in flow and so real-time estimates of changes in CBF can be made.[4, 6, 7] Accurate interpretation of TCD requires an understanding of when this assumption is valid, as small changes in diameter may also occur in basal arteries and become sources of error.[8]

Cerebral metabolic rate and systemic blood pressure have little impact on this assumption. Metabolic and pressure autoregulation of the cerebral circulation occurs largely at the arteriolar level. The change in calibre involves vessels 400 $\mu$m in diameter or less which are "downstream" from the vessels insonated by TCD.[8]

During steady state anaesthetic conditions, it is acceptable to interpret changes in flow velocity as changes in CBF, although controversy still exists.[4, 5] The intravenous agents do not appear to affect the basal vessels. Most but not all the evidence suggests that inhalational agents have negligible effects on the diameter of the conductance vessels and thus have little impact on the accuracy of TCD measurements.[4] The same can be said for some other vasoactive drugs. Sodium nitroprusside and phenylephrine, for example, do not significantly affect proximal MCA diameter.

There are some pathophysiological circumstances (e.g. cerebral artery vasospasm) in which the calibre of basal arteries is known to change. Some vasoactive drugs, such as GTN, are known to cause basal artery vasodilatation in healthy volunteers.[4, 5] Furthermore, other physiological, pathological and technical factors may also affect measured TCD velocities. These are broadly classified in Table 1.

**Table 1**
Factors affecting TCD flow velocity

| *Physiological* | *Pathological* |
|---|---|
| ICP | |
| **Cerebral blood flow** | Vasospasm |
| Cerebral metabolic rate | Stenosis |
| CO2 levels | |
| Haematocrit | *Pharmacological* |
| Age | GTN |
| Collateral flow patterns | |
| Body temperature | *Technical* |
| | Angle of Insonation |

**Insonation angle**

Unless the ultrasound beam is directed along the axis of flow, a difference exists between the measured and true velocities. The true velocity is proportional to the cosine of the angle of insonation, so the relationship is non-linear. Using the Doppler principle, the measured velocity is always lower than or equal to the real flow velocity.[9] Using an insonation angle of less than 30°, the error is between 0-13.5%. At 60°, the

error is 50%.[3] Angle of insonation must therefore be minimised. It exceeds 30°, however, for approximately 25% of vessels insonated.[8] This highlights the need to insonate vessels at similar angles if serial measurements are to be reliable. Insonation angle limitations, combined with restricted acoustic windows, render many intracranial vessels inaccessible to TCD.

**Examination**

Natural foraminae and relatively thin areas of the cranium are used to access the intracranial vasculature. Specifically, the transtemporal (anterior, middle, posterior cerebral arteries), transorbital (carotid syphon, ophthalmic artery) and suboccipital (basilar, vertrabral arteries) routes are used. The temporal window permits insonation of the proximal segment (M1) of the middle cerebral artery (MCA) in 90% of people.[4] Younger subjects generally have larger windows, making studies technically easier to obtain.[10] Advancing age, female gender and African descent are factors associated with smaller acoustic windows, leading to failure rates as high as 20-30% in these populations.[2, 4, 8, 10] About 75-80% of the carotid artery blood flow passes through the ipsilateral MCA, making the MCA representative of hemispheric blood flow and a key vessel for TCD analysis.[4]

Handheld probes can be used to perform brief studies to scan the cerebral circulation. For prolonged monitoring, and to ensure a constant angle of insonation over time, various devices allow the fixation of probes over the temporal window for monitoring of up to four vessels simultaneously.

**Vessel Identification**

Characteristics such as the acoustic window, vessel depth, direction of flow, relative flow velocities, spatial relationship to the MCA/ACA bifurcation and response to ipsilateral carotid artery compression are all used to discriminate between vessels.[10] Distinct visual waveform contours and audio feedback may assist in localisation of vessels and optimisation of signal.[10] A knowledge of the relevant anatomy is important.

It is clear, therefore, that TCD is not a "set and forget" monitor. It requires skill and expertise to allow mapping of elements of the cerebral circulation and the signals must be considered in the clinical context. Nevertheless, it is a valuable tool. The applications of TCD for clinical practice are discussed below.

**APPLICATIONS OF TCD**

In 1990 the American Academy of Neurology (AAN) reported that TCD was of established value in the assessment of intracranial stenoses, collateral circulation, cerebral vasoconstriction (especially after subarachnoid haemorrhage), arteriovenous malformations and brain death.[11] Numerous other applications have been investigated. In 2000, with the endorsement of the American Society of Neuroimaging and the Neurosonology Research Group of the World Federation of Neurology, a panel of 11 international experts issued an evaluation of the clinical utility of TCD.[8] It followed the AAN "Format for an Assessment" and incorporated personal experience, solicited opinions, extensive Medline searches and literature reviews.

**Ischaemic cerebrovascular disease**

Intracranial atherosclerosis accounts for up to 10% of strokes and transient ischaemic attacks. TCD criteria for the diagnosis of intracranial stenosis includes

circumscribed flow velocity increase, distal signal damping and side to side differences in velocity.[8] Signal absence from an artery in the presence of signals from other vessels insonated through the same window and signs of collateral flow suggest vessel occlusion. The international panel recommended TCD as being of established value for the detection of intracranial atherosclerotic lesions as well as the evaluation of patients with internal carotid artery disease based on class II evidence (provided by one or more well-designed clinical studies).[8]

TCD has been used to identify acute cerebral arterial occlusion, monitor arterial recanalisation and evaluate the risk of haemorrhagic transformation of large volume ischaemic lesions.[4] Some centres use serial TCD to guide thrombolytic therapy.[8, 12] TCD is of established value for the evaluation of patients with acute cerebral infarction, based on class II evidence.[8]

**Embolism**

TCD will detect both gaseous and solid emboli.[2] Formed element emboli produce high intensity transient signals (HITS) which move within the spectral envelope and with the direction of flow. Artifacts are bidirectional and have maximum intensities at low frequencies.[8] Asymptomatic emboli are common in the presence of known carotid artery stenosis and may be an independent predictor of future stroke risk.[2] They may localise an embolic source and provide a surrogate endpoint in the evaluation of novel therapies.[2, 8] There is strong evidence that TCD can detect cerebral microembolisation. Babikian et al concluded that the clinical usefulness of this detection required more study before a definite recommendation could be given.[8]

**Vasospasm and Subarachnoid Haemorrhage**

A variety of central nervous system disorders may induce cerebral vasospasm, or contraction of the intracranial arteries. Cerebral ischaemia may occur, leading to transient or permanent neurological dysfunction. Vasospasm is the leading cause of morbidity and mortality in patients who initially survive an aneurysmal subarachnoid haemorrhage (SAH), even after the aneurysm has been secured surgically or radiologically.[4, 13] It occurs in 70% of all patients presenting with aneurysmal SAH and leads to symptomatic brain ischaemia or infarcts in up to 36% of all patients.[8, 13] Mortality in the first 2 weeks after SAH increases 150-300% in the presence of vasospasm.[13]

Various time courses have been described.[14] Generally, vasospasm requires 4 days to develop and resolves by day 17.[12] The pathogenesis is not precisely known, although it appears to be related to the presence of subarachnoid blood. Prolonged smooth muscle contraction is mediated by extravasated oxyhaemoglobin.[4] The process may be focal or diffuse.[12] Secondary morphological changes such as intimal hyperplasia or sub-endothelial fibrosis can follow. Although once regarded as the major cause of lumen narrowing, these changes are now considered delayed and nonspecific responses to cerebral vasospasm, appearing after its resolution.[13]

The haemodynamic effects of vasospasm are similar to that of a stenosis — an increase in blood flow velocity and a loss of pressure occur through the narrow segment. Initially, cerebral blood flow will be maintained by dilatation of the distal vascular bed but, when the compensatory effects of collateral circulation and distal autoregulation are exhausted, critical reductions in CBF can occur.[15]

Causes of cerebral vasospasm include SAH, trauma, meningitis and pre-eclampsia,

and all can be associated with ischaemic cerebral damage. In the presence of basal artery vasospasm, flow velocity can no longer be used to infer cerebral blood flow. Although the assumption of unchanging vessel calibre is no longer valid, TCD still provides important clinical information regarding the presence and time-course of vasospasm. TCD diagnosis in patients at risk for vasospasm depends on:

1. Absolute velocities (MFV-MCA >120 cm/s)
2. Relative rise (>50 cm/s/d)
3. Hemispheric index or Lindegaard ratio (MCA/ICA velocity ratio >3)

Normal mean flow velocity (MFV) in the MCA is 30-80 cm/s.[15] Mean velocities above 120cm/s are associated with mild to moderate angiographic vasospasm.[12, 13] MFV above 200 cm/s suggest severe vasospasm (more than 50% diameter narrowing), although some patients may remain asymptomatic.[8, 12, 13, 14, 15] Intermediate velocities may have poorer predictive value of angiographic and symptomatic vasospasm than lower (<120 cm/s negative) or higher (>200 cm/s positive) velocities.[13] Early, rapid development of severe spasm is associated with an increased risk of cerebral infarction.[8, 12, 13, 14, 16] An early rise in MFV-MCA (≥110 cm/s on or before day 5 post SAH) is an independent predictor of symptomatic vasospasm. Predictive value is higher (68±8%) when combined into a "symptomatic vasospasm risk index".[16]

Increased MFV due to vasospasm/stenosis can be distinguished from that due to hyperaemia by the MCA/extracranial ICA velocity ratio, which is normally 1.7±0.4. Lindegaard et al demonstated that a ratio >3 suggests vasospasm (ratio >6 indicates severe vasospasm).[14]

TCD is most reliable in detecting MCA vasospasm (sensitivity 75-90%, specificity >90%), as well as that of the vertebral and basilar arteries. Collateral flow patterns render it unreliable in detecting vasospasm of the anterior cerebral arteries (ACA).[8, 12] The TCD effects of haemodynamic therapy and nimodipine are not well studied although higher velocities have been recorded even in the absence of vasospasm.[13, 14]

While conventional and digital subtraction cerebral angiography are the standard techniques used to diagnose vasospasm, they are invasive and associated with significant morbidity.[8] It is clear that TCD is valuable for the non-invasive and early diagnosis of vasospasm, before the onset of ischaemic symptoms, which can allow the early institution of appropriate treatment regimens and thus potentially the prevention of irreversible damage.[4] The Babikian group gave TCD a positive recommendation based on class II evidence.[8] Daily, or alternate day TCD assessment has been advocated.[4, 12] Monitoring is particularly useful between days 4-10 or until spasm diminishes to low moderate levels or less.[8] CBF imaging techniques such as xenon computed tomography (Xe CT) and single-photon emission computed tomography (SPECT) in conjunction with TCD assessment may be helpful in directing therapy.[8, 17]

The treatment of vasospasm includes induced hypervolaemia, hypertension, haemodilution (TripleH therapy) and the administration of cerebral selective calcium channel antagonists. The rationale is that, in areas supplied by affected vessels, cerebral blood flow regulation is impaired and becomes pressure-dependant.[13] These interventions have potentially life threatening complications, such as cardiac failure, electrolyte abnormalities, catheter related problems, rupture of an unsecured aneurysm or exacerbation of cerebral oedema.[13] Despite the lack of robust outcome data, triple H therapy received a qualified recommendation by the Stroke Council of the American Heart Association in 1994.[18] The use of nimodipine was strongly recommended to reduce poor outcome related to vasospasm.[18] Intracranial angioplasty

and intra-arterial papaverine infusions are also used in selected patients, particularly those refractory to medical treatment.

**Cerebral circulatory arrest**

Brain death is a clinical diagnosis which can be supported by TCD evidence of absent cerebral blood flow at all insonation sites.[8] Potential patterns are low diastolic flow velocities, systolic peaks, oscillating blood flow (retrograde diastolic flow), short systolic spikes, and absent TCD signals.[12] These patterns are similar to those observed with severely elevated intracranial pressure and several caveats apply. Monitoring should continue for at least 30 minutes at normal body temperature.[8] Repeated measures improve reliability. Flow velocity differences of more than a few cm/s may indicate a technically inadequate study. Only loss of a previously identifiable waveform can be interpreted as the absence of cerebral flow.[12]

TCD may facilitate shortening the observation period before organ harvest, and provide laboratory confirmation of the clinical diagnosis. There is class II evidence supporting the established value of TCD for evaluating cerebral circulatory arrest associated with brain death.[8]

**Arteriovenous Malformations**

Arteriovenous malformations (AVM) are developmental anomalies with direct communication between arteries and veins. Without vasomotor arterioles and capillaries, these vessels are characterised by high blood velocity, low pulsatility, low perfusion pressure and decreased $CO_2$ reactivity.[4] AVMs may cause intracerebral haemorrhage, seizures, or both.[8] Treatment includes endovascular embolisation, radiotherapy and surgical removal. Differentiation between healthy and AVM "feeder" vessels is important.

The evidence for TCD assessment of AVMs consists of small, non-blinded series, which lack randomised controls (class III).[8] The Babikian group concluded that TCD was unacceptable as a screening test, but is effective in identifying supply arteries to medium or large AVMs. It is effective in assessing staged embolisation or resection of AVMs.[8]

**Perioperative monitoring**

Stroke is both the major indication for carotid endarterectomy and its major complication.[4] The majority of perioperative strokes are embolic, but optimisation of cerebral perfusion during cross clamping is considered desirable.[4, 8] Shunt insertion increases stroke risk and is avoided where possible.[4]

TCD can be used to monitor ipsilateral MCA and thus provide an endpoint for treatment regimes such as optimal perfusion pressure or the need to shunt. Stroke rate is increased with initial fall in MFV to $\leq 15\%$ pre-clamp value and with persistent decrease to $<40\%$ baseline after 5 minutes. A post-clamp reduction in MFV-MCA to less than 40% baseline is a commonly accepted indication for shunt.[4] Postoperatively, HITS monitoring to detect and count emboli can be performed. Furthermore, TCD can detect kinked or thrombosed shunts.[4, 8]

With a strong consensus of class III evidence, the panel gave TCD a positive recommendation for use during and after carotid endarterectomy. It was considered investigational during cardiac surgery, but overall received a "promising" rating for perioperative monitoring.[8]

**Sickle cell disease**

Arterial stenotic lesions and cerebral infarction are common complications of sickle cell disease.[19] Time averaged maximum mean flow velocity of 200 cm/s or greater in MCA or intracranial internal carotid artery is strongly associated with an increased risk of stroke in neurologically asymptomatic children. Prophylactic blood transfusion in these patients reduced relative risk of ischaemic stroke by 90%.[20] The National Heart, Lung and Blood Institute has issued a clinical alert recommending that sickle cell patients between the ages of 2 and 16 receive TCD screening.[19] This is the strongest evidence for the effective clinical application of TCD.

**Autoregulation and vasomotor reactivity**

Cerebrovascular responses can be impaired with some pathological conditions (e.g. closed head injury) and definition of these responses may be critical in determining therapeutic strategies such as deliberate hypertension. The integrity of cerebral autoregulation can be evaluated using TCD. Good temporal resolution reflects changes as they occur. The limits of pressure autoregulation are delineated using dynamic or static techniques. Dynamic recovery after a rapid transient decrease in mean blood pressure is measured after deflation of large thigh cuffs.[4] Static response to vasopressor infusion has also been used.[4]

Vasomotor reactivity (VMR) is tested using changes in flow velocity in response to acetazolamide injection, hyperventilation, or $CO_2$ inhalation.[2, 8] Flow velocities may drop 35% with hyperventilation and increase 50% with hypercapnia in healthy patients.[8] Maximally dilated vessels are refractory to dilator stimuli and exhibit no increase in flow velocity. This indicates that vasomotor reserve is exhausted and a fall in perfusion pressure, for whatever reason, may lead to an ischaemic brain injury.[21] Diminished vasomotor reactivity has been associated with an increased risk of stroke in patients with severe extracranial internal carotid disease.[2, 8] Although the testing of VMR has only limited clinical applications, TCD is considered useful for the evaluation of vasomotor reactivity on the basis of class III evidence.[8]

**Other Applications**

Among the other major applications reviewed, TCD was considered promising for the evaluation of patients with meningeal infection. It received negative ratings for periprocedural monitoring (e.g. angiography), migraine and the assessment of central venous thrombosis, where it was regarded as investigational.[8]

**TCD TRAINING**

TCD is operator dependent. Accurate acquisition and interpretation of data requires knowledge, skill, and experience. An understanding of the clinical context is important. The American Academy of Neurology and the American Society of Neuroimaging have recommended certification of physicians performing or interpreting TCD studies, and endorsed guidelines for training.[8] It has been suggested that about 200 examinations are required to gain familiarity with the technique.[22]

**DEVELOPMENTS**

Echo-contrast enhancing agents have facilitated the combination of Doppler blood flow imaging with B-mode tissue imaging (transcranial colour-coded duplex or TCCD).[2, 22] These microbubble preparations survive cardiopulmonary and capillary

transit and increase the signal to noise ratio by a factor of 1000.[23] The failure rate due inadequate acoustic window is reduced.[23] Spatial resolution is enhanced allowing visual identification of vessels, sampling of defined volumes within the vessel and calculation of angle-corrected flow velocities.[2, 22] The principle application of TCCD is in the evaluation of cerebrovascular status in stroke patients.[23] The use of contrast agents during TCD was deemed investigational by Babikian et al.[8]

## SUMMARY

Transcranial doppler provides a non-invasive estimate of cerebral blood flow based on flow velocity. Its advantages and disadvantages are summarised in Table 2. In the hands of skilled operators, real time investigation of cerebral haemodynamics is possible at the bedside. Despite a number of technical limitations, TCD has an established role in the evaluation of ischaemic cerebrovascular disease, subarachnoid haemorrhage, arteriovenous malformations and cerebral circulatory arrest. It is effective in the evaluation of sickle cell disease and shows promise as a perioperative monitor, particularly for carotid endarterectomy.

**Table 2**
Advantages and disadvantages of TCD

| Advantages | Limitations |
|---|---|
| effective | operator dependant |
| noninvasive | skilled personnel required |
| uses small, portable equipment | patient must lie still |
| relatively inexpensive | technical limitations including acoustic window, angle of insonation |
| repeatable | poor spatial resolution |
| safe | |
| good temporal resolution | |

## REFERENCES


1. Aaslid R, Markwalder TM, Nornes H. Non-invasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries. *J Neurosurg* 1982; 57:769.
2. Markus HS. Transcranial Doppler Ultrasound. *Br Med Bulletin* 2000; 56:378-388.
3. Aaslid R. The Doppler principle applied to measurement of blood flow velocity in cerebral arteries. In: Aaslid R, ed. Transcranial Doppler Sonography. Springer-Verlag, Wein, New York 1986; 22-38.
4. Matta B. Transcranial Doppler Ultrasonography in Anesthesia and Neurosurgery. In: Cottrell JE, Smith DS, eds. Anesthesia in Neurosurgery. St Louis, C.V. Mosby 1994; 219-235.
5. Werner C. Transcranial Doppler sonography: Trend monitor of cerebral hemodynamics? *J Neurosurg Anesth* 1991; 3:73-76.
6. Bishop CCR, Powell S, Rutt D, Browse NL. Transcranial Doppler measurement of middle cerebral artery blood flow velocity: A validation study. *Stroke* 1986; 17:913-915.
7. Markus HS. Transcranial Doppler ultrasound. *J Neurol Neurosurg Psychiatry* 1999; 67:135-137.
8. Babikian VL, Feldmann E, Wechsler LR et al. Transcranial Doppler Ultrasonography: Year 2000 Update. *J Neuroimaging* 2000; 10:101-115.
9. Aaslid R. Developments and principles of transcranial doppler. In: Newell DW, Aaslid R, eds. Transcranial Doppler. Raven Press, New York 1992; 1-8.
10. Fujioka KA, Douville CM. Anatomy and freehand examination techniques. In: Newell DW, Aaslid R. eds. Transcranial Doppler. Raven Press, New York 1992; 9-31.
11. van den Noort S, Caplan L, Dyck P et al. Assessment: Transcranial Doppler. Report of the American Academy of Neurology, Therapeutics and Technology Assessment Subcommittee. *Neurology* 1990; 40:680-681.
12. Manno EM. Transcranial Doppler ultrasonography in the neurocritical care unit. *Crit Care Clin* 1997; 13:79-104.

13. Treggiari-Venzi MM, Suter PM, Romand J. Review of medical prevention of vasospasm after aneurysmal subarachnoid hemorrhage: A problem of neurointensive care. *Neurosurgery* 2001; 48:249-261.
14. Seiler RW, Newell DW. Subarachnoid Hemorrhage and vasospasm. In: Newell DW, Aaslid R, eds. Transcranial Doppler. Raven Press, New York 1992; 101-122.
15. Seiler RW, Aaslid R. Transcranial Doppler for evaluation of cerebral vasospasm. In: Aaslid R, ed. Transcranial Doppler Sonography. Springer-Verlag. Wein, New York 1986; 119-131.
16. Qureshi AI, Sung GY, Razumovsky AY, Lane K, Straw RN, Ulatowski JA. Early identification of patients at risk for symptomatic vasospasm after aneurysmal subarachnoid hemorrhage. *Crit Care Med* 2000; 28(4):984-990.
17. Newell DW, Elliott JP, Eskridge JM, Winn HR. Endovascular therapy for aneurysmal vasospasm. *Crit Care Clin* 1999; 15:685-699.
18. Mayberg MR, Batjer HH, Dacey R et al. Guidelines for the management of aneurysmal subarachnoid haemorrhage: a statement for healthcare professionals from a special writing group of the Stroke Council, American Heart Association. *Circulation* 1994; 90:2592-2605.
19. Clinical Alert From the National Heart, Lung, and Blood Institute. National Institutes of Health. September 18 1997.
20. Adams RJ, McKie VC, Hsu L et al. Prevention of a first stroke by transfusions in children with sickle cell anemia and abnormal results on transcranial Doppler Ulltrasonography. *NEJM* 1998; 339:5-11.
21. Ringelstein EB, Otis SM. Physiological testing of vasomotor reserve. In: Newell DW, Aaslid R, eds. Transcranial Doppler. Raven Press, New York 1992; 83-100.
22. Gahn G, von Kummer R. Ultrasound in acute stroke: a review. *Neuroradiology* 2001; 43:702-711.
23. Zipper SG, Stolz E. Clinical application of transcranial colour-coded duplex sonography — a review. *Eur J Neurol* 2002; 9:1-8.

# Management of Severe Peri-operative Coagulopathy: Role of Recombinant Activated Factor VII

NEVILLE GIBBS, MB, BS, MD, FANZCA
Sir Charles Gairdner Hospital, Nedlands, Western Australia

*Dr Gibbs is Head of the Department of Anaesthesia at the Sir Charles Gairdner Hospital in Perth, Western Australia. His clinical interests include anaesthesia for cardiac surgery and liver transplantation. His research interests include peri-operative coagulation disturbances. This paper is based on "An Update on Coagulation" that he presented at the Cardiothoracic Vascular and Perfusion Special Interest Group Meeting in Uluru, in July 2001.*

**Introduction**

Recombinant activated factor VII (rFVIIa) (eptacog alfa, NovoSeven®, Novo Nordisk, Copenhagen, Denmark) was introduced for the treatment of bleeding episodes in FVIII or FIX deficient patients who do not respond to standard replacement therapy, due to the presence of inhibitory antibodies.[1] However, since its introduction, rFVIIa has been used for a wide range of other coagulation disorders, including peri-operative coagulopathies associated with trauma, major surgery and organ transplantation.[1] Initial reports suggest that it is effective in reducing or arresting bleeding in these situations, leading to improvements in patient outcome and conservation of other coagulation products.[1] At the same time, few adverse effects have been reported. This has led to increased requests for peri-operative use of the drug, prompting concerns that it may be used somewhat indiscriminately, despite its high cost (approximately AUS$6000 per dose). In response, many hospitals have developed guidelines for the peri-operative availability of rFVIIa, restricting its use to those patients who have failed 'conventional treatment' and whose recovery would otherwise be unexpected.

The aim of this paper is to briefly summarise what is considered "conventional treatment" for severe peri-operative coagulopathy, to describe the mechanism of action of rFVIIa, and to discuss its current use and role in the peri-operative period.

**Definition of Severe Coagulopathy**

A coagulopathy implies that clot formation is impaired. In the most severe coagulopathies, clot formation does not occur. In the surgical setting, a severe coagulopathy is associated with diffuse microvascular bleeding manifesting as persistent ooze from wound edges or vascular anastomotic sites, as well as absence of clot formation in spilled blood. Laboratory features may include severe thrombocytopenia and markedly prolonged coagulation times (e.g. prothrombin time, activated partial thromboplastin time), with or without evidence of fibrinolysis.

**Haemostasis vs Coagulation**

The terms haemostasis and coagulation are often used interchangeably, which can lead to confusion in the surgical setting. Strictly, haemostasis is the arrest of bleeding through any mechanism, whereas coagulation refers only to the formation of a fibrin clot. Haemostasis can occur in the absence of coagulation (e.g. surgical ligation of a bleeding vessel in a fully heparinised patient). On the other hand, haemostasis may be incomplete despite normal coagulation (e.g. injury to a major blood vessel). Peri-operatively, the situation may fall between these two extremes. For example, a patient may have a coagulopathy, but it may not be clear whether this is the sole cause of bleeding, or whether surgical haemostasis can be improved. In this scenario, attempts should be made to correct the coagulopathy in the first instance. However, if the coagulopathy cannot be corrected in a timely manner or if bleeding is sufficient to cause haemodynamic compromise, surgical intervention is required despite a persistent coagulopathy.

**Causes of Peri-operative Coagulopathy**

In some circumstances, the cause of a peri-operative coagulopathy is obvious and the treatment can be directed accordingly. However, in many cases, the exact cause is not known, or there is more than one mechanism involved. The causes of most peri-operative coagulopathies fall under one (or more) of the following headings (Table 1):

**Table 1**
Causes of peri-operative coagulopathy

| | |
|---|---|
| Drugs | Unfractionated heparin, low molecular weight heparin, heparinoids, direct thrombin inhibitors, warfarin, aspirin, other COX-1 inhibitors, thienopyridines, GpIIb/IIIa inhibitors |
| Diseases | Pre-existing congenital or acquired deficiencies of platelets or other coagulation factors; liver disease; uraemia |
| Dilution | Dilution of coagulation factors caused by blood loss and replacement with factor deficient fluids |
| Destruction | Destruction or reduction of platelet receptors during cardiopulmonary bypass |
| Drop in temperature | Hypothermia <34°C |
| DIC | Disseminated intravascular coagulation leading to consumption of coagulation factors and secondary fibrinolysis; also consider primary fibrinolysis |

*Drugs*

Unfractionated heparin is often given during vascular or cardiac surgery. After protamine reversal, some heparin effect may persist due to inadequate dosage or heparin rebound.[2] Protamine does not fully reverse the anti-Xa effects of low molecular weight heparin or heparinoids.[2] Even if heparin is fully reversed, excess protamine has anticoagulant effects of its own.[2,3] Heparin has other effects that inhibit platelet aggregation.[4,5] In rare instances, heparin may be given inadvertently (e.g. flushing of heparin present in a central venous catheter). Other drugs to consider are oral anticoagulants (i.e. warfarin), antiplatelet agents, and novel direct thrombin inhibitors.[2,6] Recent ingestion of aspirin or other COX-1 inhibitors alone will not cause a severe coagulopathy, but may complicate other coagulation disturbances.[6] In contrast, recent clopidogrel may impair platelet function sufficiently to be the primary cause of a coagulopathy.[6] Glycoprotein IIb/IIIa antagonists such as abciximab are

potent inhibitors of platelet aggregation and increase blood loss if given within a few hours of surgery.[6] Tirofiban has a shorter half-life than abciximab and is of lesser concern.[6] Fibrinolytic agents such as streptokinase or tissue plasminogen activator will lyse any clot that is formed, but the effects are short lived.[7] Direct thrombin inhibitors (eg hirudin, argatroban) may have been used in patients who have recently undergone cardiological interventional procedures.[2,3]

*Diseases*

Patients may have a pre-existing coagulation factor deficiency (e.g. factor VIII, factor IX, von Willebrand factor), thrombocytopenia or a platelet function disorder. In emergency situations, this information may not be available. Rarely, the condition may have been previously undiagnosed. Patients with severe liver disease have deficient clotting factor production with varying degrees of thrombocytopenia.[8] Patients with obstructive jaundice may have ineffective vitamin K dependent factor production (II, VII, IX, X, protein C). Uraemic patients may have impaired platelet function.[8]

*Dilution*

Patients may develop or exacerbate a coagulopathy peri-operatively if persistent blood loss is replaced with fluids that contain no coagulation factors (or only low levels of them). Platelets in cold-stored whole blood are inactive within a few hours.[9] The heat labile factors V and VIII become less active in cold-stored whole blood within days. Packed red cells have minimal plasma volume, making them essentially deficient of all coagulation factors. With ongoing blood loss, the decrease in platelet count is the most important defect, although the fall is often less than expected, due to the release of stored platelets from the spleen and other reticulo-endothelial sites.[9]

In any event, abnormal surgical bleeding is unlikely if the platelet count is $>75{,}000/\mu L$ (unless platelet function is abnormal). Below a platelet count of $50{,}000/\mu L$, abnormal surgical bleeding can be expected. Dilution of coagulation factors rarely impairs coagulation unless their levels fall to $<30\%$ of normal.[9] This degree of dilution would require blood loss greater than one blood volume.

Severe anaemia itself may contribute to a bleeding tendency.[10] During normal laminar flow, platelets are concentrated peripherally near the vessel wall. Severe anaemia promotes turbulent flow, which mixes platelets throughout the blood flow stream, causing an apparent decrease in platelet count at the vessel wall.

*Drop in Temperature*

Below 34°C there is progressive impairment of coagulation, particularly if there are other concurrent coagulation disturbances.[11] Hypothermia slows the enzymatic reactions involved in coagulation, and the platelet count falls due to sequestration in the liver and spleen. Hypothermia-induced coagulation disturbances appear to self-correct on return to normothermia.

*Destruction*

Cardiopulmonary bypass (CPB) exposes blood to artificial surfaces leading to activation of coagulation pathways and platelets.[12] Large doses of heparin prevent fibrin formation, but do not prevent the activation of platelets and other coagulation factors. Activated platelets degranulate during CPB, leaving them relatively ineffective

in the post-CPB period. The density of platelet adhesion receptors (e.g. GpIb, GpIIb/IIIa) also falls during CPB, which may result in impaired aggregation subsequently. If fibrinolysis occurs, platelet receptors may be lysed by plasmin. Heparin-bonded circuits (e.g. Carmeda AB, Medtronic Inc, Minneapolis, MN, USA) may reduce activation of coagulation during CPB.[3]

*Disseminated Intravascular Coagulation/Fibrinolysis*

Disseminated intravascular coagulation (DIC) may be a complication of severe trauma, septicaemia, fat or amniotic fluid embolism or shock from any cause.[8] The mechanism involves massive exposure to tissue factor, either through extensive tissue injury or activation of blood components. The result is extensive microthrombi formation, particularly in small vessels, with rapid consumption and depletion of coagulation factors. In response, tissue plasminogen activator is released resulting in a secondary fibrinolysis. This complicates the underlying consumptive coagulopathy. Primary fibrinolysis is less common, but may occur in patients with severe liver disease (eg. during liver transplantation), or during urological surgery, as a result of urokinase release. Activation of fibrinolysis may also occur during prolonged CPB.

**Prevention**

The development of a peri-operative coagulopathy should be prevented where possible by correction of pre-existing deficiencies, cessation of anticoagulant or antithrombotic drugs an appropriate interval pre-operatively, core temperature maintenance >35°C by warming IV fluids and using warm air blankets, and maintaining cardiovascular stability and tissue perfusion by prompt replacement of blood loss (Table 2). During prolonged CPB or deep hypothermic circulatory arrest, activation of coagulation can be minimised by using a Carmeda (or similar) heparin-bonded circuit. In complex cardiac or vascular cases, or major organ transplantation, prophylactic antifibrinolytic therapy should be considered (see below).

**Table 2**
Prevention of peri-operative coagulopathy

| | |
|---|---|
| Cease | Cease anticoagulant and antiplatelet drugs an appropriate interval pre-operatively |
| Correct | Correct pre-existing or current deficiencies before they manifest clinically |
| Core temperature | Maintain core temperature >34°C |
| Cardiovascular stability | Maintain cardiovascular stability and tissue perfusion to avoid shock and release of tissue thromboplastin |
| Carmeda (or similar) circuit | Consider heparin-bonded circuit during complex cases involving cardiopulmonary bypass |

**Monitoring**

Coagulation tests are of limited use in patients with a severe coagulopathy, particularly if multiple mechanisms are involved.[9] If there is no evidence of clot formation, most standard coagulation tests are likely to be abnormal. Screening tests such as the prothrombin time (PT) and activated partial thromboplastin time (aPTT) are unable to discriminate between intrinsic and extrinsic pathway defects if there is a severe common pathway defect. Nevertheless, they may provide an index of the severity of a coagulopathy. A heparin effect can be detected by protamine titration against the

aPTT or TCT ex vivo. The fibrinogen concentration is useful if DIC or fibrinolysis is suspected.[8,9] A platelet count provides accurate information about platelet numbers, but not function.

Laboratory tests often involve a long turnaround time, which limits their value in rapidly changing situations. Point of care measurements reduce the delay, but may not be as accurate. Thrombelastography and Sonoclot analysis are useful for mild to moderate coagulopathies, but not for severe coagulopathies when blood remains fluid. The activated clotting time is used for monitoring heparin effect, but is less specific if there are other coagulation disturbances.[3] Elevated D-dimer levels confirm fibrinolysis, but this test is rarely available in a timely manner. For the above reasons, it is often necessary to treat severe coagulopathy on an empirical basis.

**Conventional Treatment**

There are few controlled trials or meta-analyses available to guide management. Most of the available information has been abstracted from studies examining blood loss reduction strategies in elective or semi-elective cases. In any event, there are few therapeutic options. These can be considered under the following headings (Table 3):

**Table 3**
Treatment of peri-operative coagulopathy

| | |
|---|---|
| Protamine | Reverse effects of heparin if present |
| Platelets | Transfuse platelets to maintain platelet count >75,000/μL (where possible); If platelet function is impaired, transfuse platelets irrespective of platelet count |
| Plasma or plasma fractions | Replace factor deficiencies with plasma or plasma fractions |
| aProtinin | Use aprotinin or other anti-fibrinolytic drug to reduce or prevent fibrinolysis |
| ddavP | For mild FVIII or von Willebrand factor deficiency only |

*Protamine*

The only role for protamine in the treatment of a coagulopathy is to reverse the effects of heparin. If there is no heparin present, there is no point in giving protamine. However, if there is heparin present, there is no point in giving anything else until the heparin has been reversed. Therefore, a possible heparin effect should be considered before commencing other treatment. A heparin effect can be confirmed by protamine titration ex vivo, either in the laboratory, or at the bedside (e.g. Hepcon HMS, Medtronic, Minneapolis, MN, USA). However, free protamine has an anticoagulant effect and excessive dosage must be avoided.[3]

Newer agents that are being developed as alternatives to protamine include heparinase and recombinant platelet factor.[4]

*Platelets*

In the presence of continued bleeding, a platelet count of >75,000/μL should be maintained where possible. Transfusion of platelet concentrates may be necessary.[9] One unit of platelet concentrate increases the platelet count by 5000-10,000/μL (in a 70 kg adult). If platelet function is impaired (e.g. drugs, destruction of receptors) the platelet count is less relevant, and they should be administered empirically.

The function of allogeneic platelets depends on their duration of storage. As the severity of a coagulopathy increases, the requirement for fresher platelets increases.

However, even fresh platelets will be ineffective if the cause of the platelet dysfunction is still present (e.g. antiplatelet drugs or other inhibitors in the plasma). Ultrafresh whole blood contains the most effective platelets, but is rarely considered, due to its associated risks.[13]

*Plasma or Plasma Fractions*

Fresh frozen plasma (FFP) contains all the normal coagulation factors and will correct most factor deficiencies irrespective of their cause.[2,8]

Warfarin effect should be reversed pre-operatively, if possible, by the administration of vitamin K. The amount of FFP required to reverse a warfarin effect can be monitored using the PT. The amount of FFP required to correct a dilutional coagulopathy is more difficult to assess, but should match the number of units of packed red cell equivalents, particularly after the loss of one blood volume. Low levels of fibrinogen, such as occurs with DIC or fibrinolysis may require additional cryoprecipitate, which contains higher concentrations of fibrinogen, factor VIIIC, and vWF.[9] For patients with known FVIII or FIX deficiencies, highly purified or recombinant specific factor concentrates are available. Recombinant FVIIa is available for FVII deficiencies, and for patients with inhibitory antibodies to FVIII or FIX.[1] Recombinant FVIIa is effective in correcting many other severe coagulopathies (see below).

*Aprotinin*

If fibrinolysis is present, an anti-fibrinolytic agent such as aprotinin is indicated.[14] Aprotinin is a serine protease inhibitor that forms reversible complexes with trypsin, plasmin, and kallikrein.[14] By binding to plasmin, it inhibits the breakdown of fibrinogen and fibrin. High-dose aprotinin (2 million KIU IV followed by 500,000KIU/hour) inhibits fibrinolysis and reduces peri-operative blood loss in a range of surgical procedures. Its efficacy is reduced at lower doses.

The alternative to aprotinin is epsilon-aminocaproic acid, a synthetic lysine analogue that acts by blocking the active site of plasmin.[14] Epsilon-aminocaproic acid (150 mg/kg IV, followed by 15 mg/kg/hour) significantly reduces blood loss, but appears to be less effective than aprotinin. Tranexamic acid is another synthetic lysine analogue that acts in a similar manner to epsilon-aminocaproic acid.[14] The recommended dose is 10 mg/kg/ IV plus 1 mg/kg/hour IV. Tranexamic acid has fewer side effects, and is cheaper than epsilon-aminocaproic acid, but is not available for IV use in Australia at the present time.

All anti-fibrinolytic drugs theoretically increase the likelihood of thrombosis, although this is less likely in the presence of a severe coagulopathy.[15]

*DdavP*

Desmopressin (1-deamino-8-d-arginine vasopressin), an analogue of arginine vasopressin, temporarily increases the plasma levels of FVIII and vWF.[14] This is achieved by stimulating the release of these factors from endothelial sites. Desmopressin (0.3 mcg/kg over 20 min) is effective for the temporary management of mild cases of haemophilia A, and some types of von Willebrand's disease. There is little or no evidence that it is effective for the reduction of blood loss in other conditions, particularly if FVIII and vWF are being concurrently administered as plasma products.

*Packed Cells*

Severe anaemia (Hb<7g/dL) should be avoided by the administration of packed red cells, whole blood or washed autologous cell-saved blood where available. Correction of severe anaemia is required to maintain oxygen delivery, but also promotes laminar flow allowing platelets to concentrate peripherally near endothelial surfaces.[10]

**Failure of Conventional Treatment**

It is difficult to define when conventional treatment of a severe peri-operative coagulopathy has failed, so each hospital should establish its own criteria. For example, one set of criteria could include the loss and replacement of at least one blood volume, with no significant improvement in bleeding or coagulation despite maximal treatment of all treatable causes, for a minimum period of at least 6 hours. The actual criteria will depend on the type of patients treated and the resources available. In any event, once conventional treatment has "failed", few therapeutic options remain. It may be possible to continue with supportive therapy until the situation either improves, deteriorates, or allogeneic products are exhausted. Another option is the use of ultrafresh allogeneic whole blood from on-site donors. This is usually considered only as a last resort if recovery would be unprecedented otherwise.[13] Recombinant FVIIa is the new alternative when conventional therapy has failed.[1]

**Recombinant Activated Factor VII**

*Mechanism of Action*

Factor VIIa has a fundamental role in the initiation of coagulation.[1, 16] Under normal circumstances, only a small proportion of circulating FVII is in the activated form. Following vascular injury, FVIIa binds to tissue factor (TF) expressed on the membranes of extravascular cells (Figure 1). The TF:VIIa complex activates FIX and FX. Positive feedback occurs when local FVIIa, FIXa, and FXa activate additional FVII. Factor Xa and its co-factor FVa, bind to exposed phospholipid on the membrane of activated platelets. This complex has "prothrombinase" activity, and converts prothrombin to thrombin. Thrombin, in addition to its effects on fibrinogen, activates FV, FVIII, and platelets. The process is rapidly controlled by local inhibitors unless further FXa is produced. Under normal circumstances, the generation of additional FXa by the FIXa:FVIIIa complex is required to sustain thrombin production (Figure 1).

Recombinant VIIa is identical in structure and function to human FVIIa. However, in addition to its TF-dependent activity, rFVIIa has TF-independent activity.[1] Recombinant FVIIa directly activates FX on platelet surfaces in a dose-dependent manner. The concentration of FVIIa required for this effect is much greater than the levels normally present in plasma. The ability to directly activate FX explains how FVIIa promotes prothrombinase production in the absence of FVIIIa and FIXa. The ability to administer rFVIIa in high doses explains its efficacy in a range of situations.

*Clinical Experience with rFVIIa*

Recombinant VIIa has been approved for the control of bleeding in patients with haemophilia with inhibitors in many countries since 1996.[1] Before that time, it was available as part of a compassionate use programme only. Although it has been approved only for a specific indication, it has been shown to improve coagulation and reduce bleeding complications in a wide range of other conditions. These include

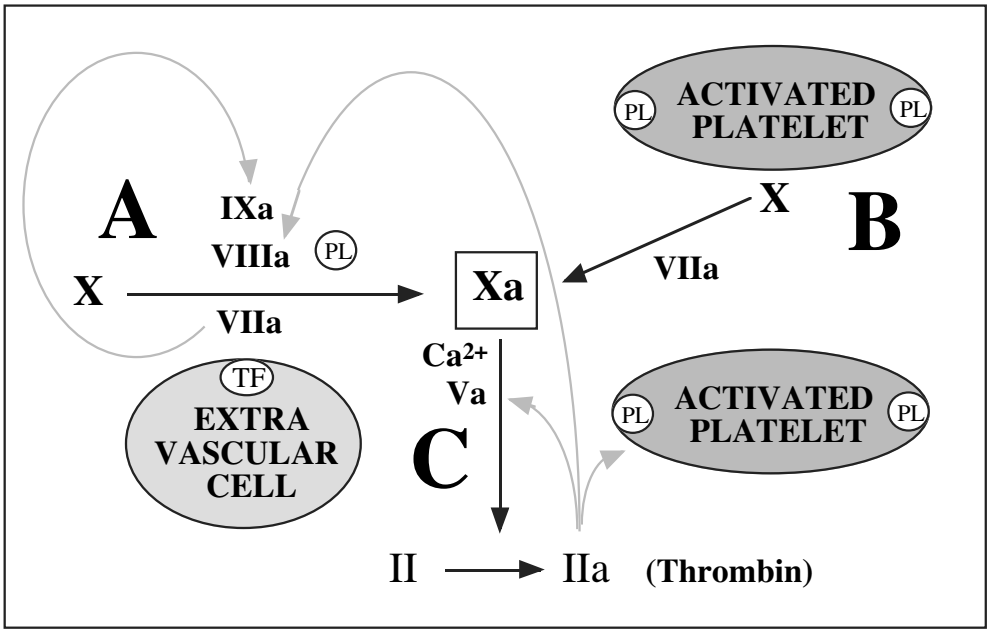


**Figure 1.** Role of factor VII in coagulation.

**A. Tissue Factor Dependent Pathway.** The small amount of activated FVII in plasma binds to tissue factor (TF) on extra vascular cells exposed at sites of vascular injury. The FVIIa:TF complex activates FX and FIX. The FIXa:VIIIa complex activates further FX, thereby amplifying FXa production.

**B. Tissue Factor Independent Pathway.** High concentrations of FVIIa directly activate FX bound to the surface of activated platelets.

**C. Common Pathway.** Factor Xa with its cofactor, FVa, calcium ions, and platelet phospholipid (PL) converts prothrombin (II) to thrombin (IIa).

thrombocytopenia, warfarin toxicity, liver disease (including fulminant hepatic failure), bleeding in neonates, surgical bleeding and trauma.[1, 17, 18, 19]

Most of the information on the use of rFVIIa in patients with peri-operative bleeding is in the form of case reports. The first report was by Kenet et al in 1999.[20] They described a patient with intractable bleeding due to a gunshot wound. This patient did not respond to maximal conventional therapy, but improved following two doses of rFVIIa, 60 mcg/kg iv. The same group later reported a series of nine trauma patients who failed to respond to maximal transfusion therapy, but who responded to rFVIIa.[21] None developed clinical signs of thrombosis.

Since 1999, there have been numerous case reports or case series on the efficacy of rFVIIa in surgical patients with severe coagulopathies and intractable bleeding. These have included abdominal surgery,[22] urological surgery,[23] orthopaedic surgery,[24] cardiac surgery,[25, 26, 27] obstetric bleeding[28] and transplantation.[29, 30] Initially, most of these cases involved the use of rFVIIa as a last resort in life-threatening bleeding when conventional management had failed. However, the encouraging results have prompted the use of rFVIIa in less severe bleeding. More recent reports suggest that rFVIIa is effective in reducing blood loss in patients with normal coagulation.[31, 32] Nevertheless, there have been no controlled trials on the efficacy or safety of rFVIIa, other than in patients with haemophilia.

*Safety of rFVIIa*

No human proteins or blood products are used in the preparation of NovoSeven®.[33] By using recombinant technology, the risk of virus transmission is avoided. Since July 1996 over 250,000 standard doses of rFVIIa have been administered to patients.[33] Few adverse effects have been reported. Non-serious adverse effects that have been attributed to rFVIIa are fever, headache, vomiting and skin hypersensitivity reactions. These are rare and self-limiting. There is no data confirming the safety of rFVIIa in pregnancy.

The incidence of thromboembolic complications is <1%, despite its use in many patients with significant co-morbidities.[1] The low incidence of thrombotic complications may relate to the localisation of rFVIIa effects to sites of vascular injury. Under physiological conditions, rFVIIa requires TF, which is normally present only on extravascular cells. Higher doses of rFVIIa directly activate FX, but only on the surface of activated platelets at sites of vascular injury.

*Dosage and Administration*

Recombinant FVIIa is available for intravenous bolus use only. It is active within minutes after injection. The half-life of the drug in adults is approximately 2.5 h. This is not affected by liver or renal dysfunction. The recommended dose in patients with haemophilia A or B with inhibitors and active surgical bleeding is 100 $\mu$g/kg given every 2 hours until haemostasis is achieved.[1] For a 70 kg patient the dose would be 100 $\mu$g $\times$ 70 = 7.0 mg. In view of the cost per ampoule, it is recommended that the dose is rounded off to the nearest 1.2 mg. A similar dose is recommended for patients with a severe coagulopathy associated with intractable bleeding. In these patients the dose should be repeated every 2-4 hours as necessary.

*Cost of rFVIIa*

The cost of rFVIIa (NovoSeven®) in Australia is currently AUS$1050.00 for a 1.2 mg ampoule.[33] A 5 mg ampoule is available at a proportionally higher cost. The approximate cost for a 70 kg patient is AUS$6000 per dose.

*Availability*

Recombinant FVIIa was approved for use in haemophilia patients with inhibitors by the Australian Therapeutic Goods Administration (TGA) in late 1998, and has been marketed by NovoSeven® in Australia since that time. It has been available in New Zealand since 1997. Since it is a marketed drug, it is available on prescription by a physician. However, if it is used for indications that do not have TGA approval, the physician must take responsibility for any adverse effects encountered.

NovoSeven® can be purchased through Novo Nordisk Pharmaceuticals (Copenhagen, Denmark). It has a shelf-life of 2 years. Due to its high cost, most hospitals have a protocol for its use, and will not release the drug until certain criteria are met. For peri-operative use, the criteria are usually designed to ensure that rFVIIa is used only for life-threatening situations when conventional treatment of coagulopathy and bleeding have failed. Obtaining the drug usually requires a request at a consultant level, with the involvement of a haematologist.

*Current Indications for Peri-operative Use of rFVIIa*

At present, the peri-operative use of rFVIIa should be limited to those patients with

severe coagulopathy and life-threatening bleeding who have not responded to maximal conventional therapy and appropriate surgical intervention. Recombinant FVIIa should not be used as a first line treatment for bleeding until its efficacy and safety in the treatment of peri-operative blood loss have been conclusively demonstrated. An exception may be made for patients undergoing procedures with massive blood loss, who do not accept allogeneic blood products for religious reasons. Recombinant FVIIa may be used as a first line treatment only for patients with haemophilia A or B with inhibitors, or for patients with pre-existing FVII deficiency.

*Monitoring the Effect of rFVIIa*

Although there is a near linear relationship between rFVIIa activity and its plasma level, it is rarely possible to measure rFVIIa activity in the peri-operative period. The recommended dose is based on a target concentration >30 U/mL.[33] Standard laboratory tests are unable to guide therapy.[33] The PT will normalise at about 5 U/mL, well below its peak effect. In contrast the aPTT may not normalise, even with peak therapeutic concentrations. Therefore, monitoring rFVIIa activity is mainly clinical.

*The Future*

Current and future clinical trials will determine the efficacy, safety and cost-effectiveness of rFVIIa in patients with peri-operative bleeding.[34] It is likely that the indications will widen, particularly if there is reduced availability or increased cost of allogeneic blood products. A reduction in the cost of rFVIIa will also influence its use. In the longer term, rFVIIa may be used as an alternative to conventional therapy, rather than as an adjunct. It may also be used for the control of bleeding in non-coagulopathic patients. At present, the relative safety of rFVIIa in coagulopathic patients is of secondary concern, because its use is limited to patients who might not otherwise survive. However, if its use widens, safety issues will be of greater concern. If it fulfils its potential, rFVIIa will be the most significant advance in transfusion medicine for several decades.

**References**

1. Hedner U, Erhardtsen E. Potential role for rFVIIa in transfusion medicine. *Transfusion* 2002; 42:114-124.
2. Warkenstein TE, Crowther MA. Reversing anticoagulants both old and new. *Can J Anaesth* 2002; 49:S11-S25.
3. Despotis GJ, Joist JH. Anticoagulation and anticoagulation reversal with cardiac surgery involving cardiopulmonary bypass: an update. *J Cardiothor Vasc Anesth* 1999; 13(Suppl 1):18-29.
4. Walker CPR, Royston D. Thrombin generation and its inhibition: a review of the scientific basis and mechanism of action of anticoagulant therapies. *Br J Anaesth* 2002; 88:848-863.
5. Murithi EW, Belcher PR, Day SP et al. Heparin-induced platelet dysfunction and cardiopulmonary bypass. *Ann Thorac Surg* 2000; 69:1827-1832.
6. Samama CM, Bastien O, Forestier F et al. Antiplatelet agents in the peri-operative period: expert recommendations of the French Society of Anesthesiology and Intensive Care (SFAR) 2001 — a summary statement. *Can J Anaesth* 2002; 49:S26-S35.
7. Marder VJ, Sherry S. Thrombolytic therapy: current status. *N Eng J Med* 1988; 318:1512-1520.
8. Martlew VJ. Peri-operative management of patients with coagulation disorders. *Br J Anaesth* 2000; 84:446-455.
9. Miller RD. Transfusion therapy. In: Miller RD, ed. Anesthesia, 5th Ed. Churchill Livingstone, Philadelphia, 2000; 1613-1644.
10. Hellem AJ, Borchgrevink CF, Ames SB. The role of red cells in hemostasis: the relationship between hematocrit, bleeding time, and platelet adhesiveness. *Br J Haematol* 1961; 7:42-50.

11. Sessler DI. Complications and treatment of mild hypothermia. *Anesthesiology* 2001; 95:531-543.
12. Spiess BD. Maintenance of homeostasis during cardiopulmonary bypass. *J Cardiothor Vasc Anesth* 1999; 13(Suppl 1):2-5.
13. Erber WN, Tan J, Grey D, Lown LA. Use of unrefrigerated fresh whole blood in massive blood transfusion. *Med J Aust* 1996; 165:11-13.
14. Mannucci PM. Drug therapy: hemostatic drugs. *N Eng J Med* 1998; 339:245-253.
15. Levy J. Hemostatic agents and their safety. *J Cardiothor Vasc Anesth* 1999; 13(Suppl 1):6-11.
16. Hoffman M, Monroe DM III. A cell-based model of hemostasis. *Thromb Haemost* 2001; 85:958-965.
17. Kessler C. Haemorrhagic complications of thrombocytopenia and oral anticoagulation: is there a role for recombinant activate factor VII? *Intensive Care Medicine* 2002; 28:S228-S234.
18. Veldman A, Fischer D, Voight B et al. Life-threatening hemorrhage in neonates: management with recombinant factor VII. *Intensive Care Medicine* 2002; 28:1635-1637.
19. Lynn M, Jeroukhimov I, Klein Y, Martinowitz U. Updates in the management of severe coagulopathy in trauma patients. *Intensive Care Medicine* 2002; 28:S241-S247.
20. Kenet G, Walden R, Elad A, Martinowitz U. Treatment of traumatic bleeding with recombinant factor VIIa. *Lancet* 1999; 354:1879.
21. Martinowitz U, Kenet G, Segal E et al. Recombinant activated factor VII for adjunctive hemorrhage control in trauma. *J Trauma* 2001; 51:431-439.
22. Moscardo F, Perez F, de la Rubia J et al. Successful treatment of severe intra-abdominal bleeding associated with disseminated intravascular coagulation using recombinant activated factor VII. *Br J Haematol* 2001; 113:174-176.
23. Friederich PW, Geerdink MGF, Spatoro M et al. The effect of the administration of recombinant activated factor VII (NovoSeven®) on perioperative blood loss in patients undergoing transabdominal retropubic prostatectomy: the PROSE study. *Blood Coag Fibrin* 2000; 11(Suppl 1):S129-S132.
24. Tobias JD. Synthetic VIIa to treat dilutional coagulopathy during posterior spinal fusion in two children. *Anesthesiology* 2002; 96:1522-1525.
25. Hendriks HG, van der Maaten, de Wolf J et al. An effective treatment of severe intractable bleeding after valve repair by one single dose of activated recombinant factor VII. *Anesth Analg* 2001; 93:287-289.
26. Al Douri M, Shafi T, Al Khudairi E et al. Effect of the administration of recombinant activated factor VII (rFVIIa, NovoSeven®) in the management of severe uncontrolled bleeding in patients undergoing heart valve replacement surgery. *Blood Coag Fibrin* 2000; 11(Suppl 1):S121-S127.
27. Von Heymann C, Hotz H, Konertz W, Kox WJ, Spies C. Successful treatment of refractory bleeding with recombinant factor VIIa after redo coronary artery bypass graft surgery. *J Cardiothor Vasc Anesth* 2002; 16:615-616.
28. Breborowicz GH, Sobieszcyk S, Szymankiewicz M. Efficacy of recombinant activated factor VII (rFVIIa, NovoSeven®) in prenatal medicine. *Arch Prenatal Med* 2002; 8:21-27.
29. Kallcinski P, Kaminski A, Drewniak H et al. Quick correction of hemostasis in two patients with fulminant liver failure undergoing liver transplantation by recombinant activated factor VII. *Transplantation Proceedings* 1999; 31:375-379.
30. Shami VM, Caldwell SH, Hespenheide EE. Recombinant activated factor VII for coagulopathy in fulminant hepatic failure compared with conventional therapy. *Liver Transplantation* 2003; 9:138-143.
31. O'Neill PA, Bluth M, Gloster ES et al. Successful use of recombinant activated factor VII for trauma-associated hemorrhage in a patient without a pre-existing coagulopathy. *J Trauma* 2002; 52:400-405.
32. Mayer SA. Intracerebral hemorrhage: natural history and rationale of ultra-early hemostatic therapy. *Intensive Care Medicine* 2002; 28:S235-S240.
33. Product information. Novo Nordisk Pharmaceuticals, Copenhagen, Denmark.
34. Erhardtsen E. Ongoing NovoSeven® trials. Intensive Care Medicine 2002;28:S248-S255.

# Intraaortic Balloon Counterpulsation: Principles and Review of Clinical Outcomes

SHANE C. TOWNSEND, MB, BS, FANZCA, FJFICM
Staff Intensive Care Specialist, Royal Melbourne Hospital Intensive Care Unit, Grattan St, Parkville, Melbourne, Victoria

*Dr Townsend graduated with Honours from the University of Queensland in 1989. During his anaesthesia training in Brisbane, he was awarded both the Renton and Cecil Grey Prizes. While a staff specialist anaesthetist at Prince Charles Hospital in Brisbane from 1999 to 2001 he gained experience working with cardiac surgical patients and aspects of managing shock states with circulatory assist devices. He is now a full time intensivist at The Royal Melbourne Hospital, where he is undertaking a Doctorate of Medicine through the University of Melbourne.*

## Introduction

Patients with end-stage heart failure consume significant health resources throughout the world. Poorer surgical candidates are undergoing cardiac and non-cardiac surgical procedures. Cardiogenic shock remains a very challenging condition to treat with a high associated mortality. There have been significant advances in the pharmacotherapy of heart failure and ischaemic heart disease in recent years with improved outcomes and greater patient expectations from treatment. Invasive and expensive modalities of managing acute circulatory failure deserve a review of the evidence for their efficacy. It is against this backdrop that a reappraisal of the experience with intraaortic balloon counterpulsation (IABC) is timely.

Balloon counterpulsation as a means of supporting the circulation was first proposed in 1962 and applied clinically later that decade by Kantrowitz.[1] The initial experience was mixed; while resulting in better haemodynamics, mortality failed to significantly improve. However, interest was reignited by a report in 1972 demonstrating the utility of the balloon pump in separating cardiac surgical patients from bypass.[2]

## Principles and Equipment

The principle goal of IABC is a reduction of cardiac afterload, permitting myocardial rest and a decrease in inotropic drug doses. The peak diastolic blood pressure is augmented above the assisted systolic pressure, theoretically improving coronary perfusion. The diastolic pressure-time index is increased while the systolic pressure-time index is reduced. The combined result is reduced myocardial ischaemia. A modest increase in cardiac output is achieved, but this may be as low as 10 to 15%. Mitral valve regurgitation associated with a low cardiac output state and pulmonary congestion is likely to benefit significantly from counterpulsation by an improvement in forward flow and decreased end-diastolic filling pressures.

The intraaortic balloon pump (IABP) consists of an external console and pneumatic

pump connected to a specialised arterial catheter. The catheter possesses a lumen for transduction of the aortic pressure wave and a lumen to conduct helium gas to and from the distal balloon. Inflation is triggered by the patient ECG, pacing spike or aortic pressure wave signal. The device can be triggered internally. Timing and augmentation can be adjusted in relation to the displayed aortic pressure wave, by examining the contour of assisted and unassisted systole. A pressure trace from the balloon is displayed and can be interrogated for signs of incorrect inflation or gas leaks. Alarms detect inadequate augmentation, gas leaks indicative of balloon failure and inability to detect the patient's ECG or cardiac pacing spike.

The balloon catheter can be placed into the aorta transthoracically at the time of surgery, occasionally a necessity in severe vasculopaths. Most commonly it is placed percutaneously via the femoral artery under radiological guidance, into the descending aorta just distal to the origin of the left subclavian artery. Surgical cut-down to the femoral artery is also an option. Correct positioning may be confirmed by transoesophageal echocardiography in the operating theatre or ICU. Placement of the balloon counterpulsator into the pulmonary artery to assist right ventricular dysfunction has also been described. The approach to insertion has traditionally incorporated a vascular sheath and sleeve, analogous to a pulmonary artery flotation catheter. This necessitates a larger aperture in the femoral artery, which may be related to subsequent complications. Newer, finer 8 French gauge catheters inserted without a sheath have improved the ease of insertion and helped to minimise vascular complications in arteriopaths.

Paediatric application imposes several equipment design constraints. A volume limiting chamber has been employed for balloon volumes less than 20 ml. Paediatric balloons matched to the aortic diameter have tended to be too long, threatening renal and mesenteric vascular occlusion. Selection of a balloon volume estimated to be 1/2 the recipient's stroke volume is recommended. The narrow 4 or 5 French femoral catheters necessary in children and high physiological heart rates mandate rapidly responsive pneumatics. A sinusoidal pressure waveform has been shown to be clinically effective and mechanically easier to achieve.

Optimal timing of balloon counterpulsation requires reference to the aortic pressure wave trace. In a dog model of counterpulsation, balloon deflation during the isovolumic contraction phase was found to improve ventriculoarterial coupling and mechanical efficiency.[3] Early and late inflation or deflation of the balloon will adversely affect ventricular afterload, augmentation of diastolic pressure and the balance between myocardial oxygen supply and demand. Arrhythmias may interfere with counterpulsation timing owing to the irregularity of systolic and diastolic periods. Atrial fibrillation may require extending the interval of balloon inflation for optimal efficiency. Correct placement of the balloon catheter immediately distal to the origin of the left subclavian artery and appropriate selection of catheter volume relative to patient height and weight are essential to achieve ideal haemodynamic performance.

Anticoagulation is necessary to prevent arterial thrombosis and embolism and degradation of catheter performance. Intravenous heparin infusion to achieve therapeutic systemic anticoagulation has been standard. Because clot formation is related to stasis of blood elements, maintenance of IABP catheter pulsation while in-vivo is important and initially may suffice, particularly in coagulopathic patients and immediately post cardiac surgery. Low molecular weight heparin and antiplatelet drugs may be alternative strategies in suitable patients. Strict monitoring and control

of anticoagulation is required to avoid haemorrhagic complications. Heparin pharmacokinetics during balloon counterpulsation are comparable with other clinical situations, with a measured elimination $t_{1/2}$ of 2.4±0.08 hr after discontinuation and a continuous infusion requirement of 16±2.5 U/kg/hr in one study.[4]

**Indications**

Traditional indications for balloon insertion have encompassed shock, myocardial ischaemia, failure to separate from cardiopulmonary bypass and severe acute mitral regurgitation. More recently, balloon counterpulsation has been suggested to prevent coronary re-occlusion following angioplasty and has been combined with thrombolysis in acute myocardial infarction and cardiogenic shock. A role in stabilising patients for transfer to a tertiary referral centre with cardiac surgical and interventional cardiological services is gaining acceptance. Although not suitable for bridging long-term to cardiac transplantation, the intraaortic balloon pump has been useful in "bridging to partial recovery" of myocardial function after ischaemic insults, facilitating later cardiac transplantation as an elective procedure. As well, counterpulsation has been applied successfully in the stabilisation of patients suffering ventricular arrhythmias refractory to medical therapy. There may be a role for the balloon pump in supporting poor-risk cardiac patients undergoing non-cardiac surgery. Studies have also employed the IABP to augment cerebral blood flow in subarachnoid haemorrhage. Application of mechanical circulatory support has been suggested in septic shock.

Severe aortic regurgitation is a contraindication to insertion of the device.

**Published Clinical Audits**

Three large clinical series summarising the accumulated experience with over 6000 intraaortic balloon pump insertions, have been published. Creswell's group from Washington University[5] published their experience with 669 patients over a five year period. In their series nearly 10% of cardiac surgical procedures were attended by balloon insertion. Thirty day mortality was significantly lower with preoperative insertion of the balloon pump, although most devices being inserted either intra- or post-operatively. The St. Louis University[6] group reported that cardiac surgical patients receiving a balloon pump had an overall operative mortality of 44%. Interestingly, they noted that the majority of survivors were in a favourable New York Heart Association functional class and survived long-term. The largest series reported to date is from the Massachusetts General Hospital.[7] In their practice, the majority of balloon pumps were inserted preoperatively and for control of myocardial ischaemia. In a retrospective analysis of independent predictors of mortality for this series, insertion of the IABP for ongoing ischaemia, isolated CABG or angioplasty carried a better prognosis. The authors argued the improved mortality in these subgroups warrants earlier insertion of the device. Mortality increases sharply with later insertion of the balloon pump in cardiac surgical patients. This may reflect the benefits of stabilising ischaemia preoperatively versus attempts at salvage once cardiogenic shock is established. Selection bias and the retrospective nature of the data preclude firm conclusions.

Trends in balloon pump usage since 1968 reveal a growing number of patients receiving counterpulsation for control of myocardial ischaemia.[7] The number of patients receiving counterpulsation for haemodynamic indications has remained relatively unchanged, while overall mortality has improved.

**Myocardial Infarction and Cardiogenic Shock**

In spite of advances in treatment of acute myocardial infarction during the interventional era, patients who present with cardiogenic shock continue to suffer a high mortality. Practices relating to counterpulsation have changed over the last 25 years,[7] with myocardial ischaemia now a common indication, while rates of IABP insertion post-cardiotomy have remained largely unchanged.

Use of the IABP remains uncommon among patients who develop cardiogenic shock complicating acute myocardial infarction,[8] perhaps as a consequence of the perceived potential for complications as opposed to the lack of definitive prospective evidence for its efficacy. A number of studies have reported lower mortalities for cardiogenic shock treated with intraaortic balloon counterpulsation. Survival in these studies is most strongly correlated with coronary revascularisation; the balloon pump playing an adjunctive role in stabilising the patient until angioplasty or surgery could be scheduled.

Results of two randomized prospective trials[9, 10] conducted in the pre-thrombolytic era did not show a survival advantage in cardiogenic shock patients treated with counterpulsation without undergoing revascularisation. More recent data, suggesting a survival advantage for patients who achieve infarct-related artery patency, has renewed interest in the role of the IABP in opening occluded coronary arteries and preventing reocclusion in low coronary flow states. Several retrospective studies[8, 11, 12] have supported the combination of thrombolysis with intraaortic balloon counterpulsation in patients developing shock as a complication of acute myocardial infarction.

Using a canine model, Gurbel et al[13] demonstrated more rapid clot lysis and coronary reperfusion with counterpulsation after administration of intravenous rTPA. A critical left anterior descending stenosis was created by an occluder with injection of a blood and thrombin mixture into the artery. Dogs treated with rTPA infusion and concomitant counterpulsation demonstrated reperfusion occurring three times faster than animals treated with rTPA alone ($P=0.02$). They did not show an increase in mean coronary blood flow during IABC and suggested that reperfusion was enhanced by pressure effects, possibly disrupting the thrombus and increasing surface area exposed to thrombolytic agent. In another animal study, Prewitt et al[14] injected radioactive blood clot into the left anterior descending artery of dogs, which subsequently underwent phlebotomy to a systolic blood pressure of approximately 90 mmHg. In this model of cardiogenic shock, IABC significantly increased the rate of thrombolysis with rTPA and exhibited a trend to increased peak diastolic coronary blood flow.

In the GUSTO-1 Study,[8] 315 patients presented with cardiogenic shock. IABC was instituted within 24 hours in 62 patients (20%). Compared to the late insertion or no-IABP group, those patients receiving early insertion of an IABP showed a trend toward lower 30-day mortality, which persisted at one year. However, the early IABP group recorded a greater number of adverse events, perhaps reflecting survival of a sicker cohort of patients. There were also more frequent episodes of moderate bleeding. Interestingly, 32% of shocked patients received IABC in the United States, whereas the corresponding figure for other participating countries was only 7.2%.

Historically, thrombolysis was a relative contraindication to insertion of a balloon pump. A retrospective study conducted by Stomel et al[11] examined survival in acute myocardial infarction and cardiogenic shock patients treated with thrombolytic therapy and IABC in a community hospital. Those patients treated with combined

thrombolysis and insertion of an IABP were more often stabilised and transferred to a tertiary referral centre for revascularisation and had a significantly greater survival (68%) compared to those treated with thrombolysis alone (23%) or IABC alone (28%) $P=0.0049$. A further study by Kovack et al,[12] also in the setting of two community hospitals, examined retrospectively 46 patients who presented in cardiogenic shock and who received thrombolysis within 12 hours of acute infarction. Twenty-seven of these patients had an IABP inserted. The IABP group had a significantly higher survival in the community hospital (93% vs 37% $P=0.0002$) and a greater proportion were transferred (85% vs 37%). Overall hospital and one year survival was also greater in the IABP group (67% vs. 32% $P=0.019$). Results from these retrospective studies must be interpreted cautiously, a major weakness being potential selection bias favouring those patients chosen for IABP insertion. The contribution of IABC to survival is difficult to distinguish in those patients subsequently undergoing transfer and revascularisation at a second institution.

The SHOCK Trial Registry[15] randomised 302 patients with cardiogenic shock due to acute infarction to receive either initial medical therapy or early revascularisation. There was a non-significant trend to a lower 30-day mortality in the revascularisation group and a 12% reduction in mortality at 6 months. Interestingly, the majority of controls received combined thrombolysis and counterpulsation with an overall mortality of 56% at 30 days for medical therapy. This result is superior to other published reports and may provide indirect evidence for employing counterpulsation and thrombolysis when cardiac catheterisation is not readily available.

**Cardiac Surgery**

Post-cardiotomy, the balloon pump has proven useful to assist in separation from cardiopulmonary bypass. Unfortunately, mortality rates of approximately 30% are reported for this indication and rise even higher if the balloon pump is inserted later in the intensive care unit in an attempt to salvage the patient. Certain groups also fare less well with insertion of a balloon pump, notably valvular surgery and graft failure after cardiac transplantation.

Aortic counterpulsation has steadily increased as an adjunct to management of cardiac surgical patients, with the largest series published by the Massachusetts General Hospital[7] reporting IABP insertion in 12.3% of all cardiac surgical procedures. IABC is applied perioperatively with the aim of preoperative stabilisation and control of refractory ischaemia, management of low cardiac output states and to facilitate separation from cardiopulmonary bypass. Mortality remains high in most series reporting IABP insertion in cardiac surgical patients, ranging from 16.3% to 48.4%.[5, 6, 7, 16, 17] It is likely that this variability results from differing philosophies in selecting patients for IABP insertion.

Examining 193 intraoperative balloon placements, Torchiana et al[7] identified age, mitral valve replacement, prolonged cardiopulmonary bypass, urgent surgery, preoperative renal dysfunction, and complex ventricular ectopy or right ventricular failure after cross clamp removal as independent predictors of mortality. However, many of these factors are general predictors of operative mortality and do not discriminate between those patients who will benefit from IABP support and those requiring a ventricular assist device. Timing of IABP insertion was shown to be significant, with a worse outcome after placement to wean from bypass or after emergency reinstitution of bypass. Torchiana et al[7] drew attention to the better

outcome observed with a bias favouring preoperative balloon insertion (70.6% of devices inserted), reporting a 16.3% overall perioperative mortality. This data is retrospective and subject to selection bias; however, the authors contend that liberal use of the IABP to stabilise myocardial ischaemia in the preoperative phase may have an operative survival advantage. Creswell et al[5] found similar results with a 19.6% mortality for patients receiving an IABP preoperatively versus 32.3% intraoperatively and 40.5% postoperatively ($P<0.001$).

Cardiac transplant patients receiving an IABP were also noted to fare better if counterpulsation was used as a preoperative bridge to transplantation. However, intra-operative insertion carried a 40% mortality, perhaps reflecting poor donor organ preservation or the pre-transplant condition of the recipient.

Long-term survival after IABP support in cardiac surgical patients was also related to specific procedures. Cardiac transplant and isolated CABG patients enjoyed the highest long-term survival, whereas CABG and MVR, CABG and AVR, CABG and left ventricular aneurysm repair, isolated MVR and isolated AVR were associated with poorer survival.[5]

Naunheim et al[6] studied 580 cardiac surgical patients who received a perioperative IABP. A multivariate analysis identified six independent prognostic indicators; preoperative NYHA class, transthoracic route of insertion, preoperative nitroglycerin infusion, age, female gender and preoperative balloon insertion. However, the $r^2$ value in the linear regression model was 0.128, permitting only a small proportion of deaths to be attributed solely to these variables. Survivors enjoyed good long-term cardiac function; 81% of patients at follow-up were classified as NYHA class I or II. Non-survivors of cardiac surgery despite IABP support, have been studied by Baldwin et al.[16] They identified four variables which significantly predicted mortality in a logistic regression model; complete heart block, advanced age, preoperative BUN, and female gender. The derived equation was then prospectively validated against a further 330 patients at another institution.

The value of defining such predictors of mortality among patients otherwise eligible for IABC, lies in triaging patients to alternative forms of circulatory support. The IABP may achieve only modest increments in cardiac output, in the order of 10%. Non-survivors despite IABC may have received inadequate haemodynamic support and are potential candidates for ventricular assist devices, provided that they can be identified early. A further study[17] of 129 consecutive patients receiving IABP support found that acute myocardial infarction with severe ventricular dysfunction (LVEF <30%) carried a high associated mortality. Fifty per cent of the 64 non-survivors died from cardiac failure, reinforcing the concept that the limited haemodynamic support achievable with the IABP may be inadequate faced with extensive infarction and severe ventricular failure. Aggressive measures including ventricular assist devices may be more appropriate.

Christenson et al[18] performed a randomised controlled trial on 60 high-risk patients undergoing CABG. Those patients randomised to pre-operative insertion of an intra-aortic balloon pump had a significantly lower mortality, shorter bypass time, higher post-operative cardiac index and shorter length of stay in the ICU and hospital. The benefit was realised with insertion just 2 hours prior to surgery. Earlier studies[19, 20, 21, 22] also support the elective use of counterpulsation to improve outcome in high-risk patients undergoing coronary surgery. Preoperative insertion of an IABP in 101 high-risk elective cardiac surgical cases was found to significantly reduce risk-adjusted

mortality versus predicted mortality (5.7% vs 20% $P=0.0014$) in a recent Australian study.[23]

**Myocardial Ischaemia and Coronary Angioplasty**

There are no randomised controlled trials for when unstable angina is the indication to insert an IABP. Nevertheless, this clinical approach enjoys a long tradition with a relatively low mortality. The emphasis in recent years has shifted from managing haemodynamic decompensation to control of ischaemia as an indication for IABP insertion, with 51.6% of devices placed for ischaemia in the Massachusetts General Hospital series.[7] This series revealed an 11.9% mortality for balloon placement to control ischaemia, whereas placement for haemodynamic instability had a higher (38.2%) mortality. Angioplasty recipients supported by IABC had an overall mortality of 14.7%.

Myocardial function does not recover immediately with reperfusion and coronary blood flow is often impaired in the early reperfusion period. IABC may support myocardial function by afterload reduction, augment coronary blood flow and enhance the efficacy of chemical and mechanical methods of reperfusion. A prospective analysis of 810 consecutive patients entered into the TAMI trials,[25] treated with thrombolytic therapy for acute myocardial infarction, found 85 patients who received an IABP. Patients presenting in established cardiogenic shock were excluded. The IABP group were characterised by greater haemodynamic instability, more prevalent multivessel and left anterior descending coronary artery disease, were older and more likely to suffer from diabetes and hypertension. The hospital mortality for the IABP group was significantly higher (32% vs 4%), possibly reflecting the selection of patients with higher acuity. No death was attributable directly to insertion of the IABP. Bleeding complications necessitating transfusion were greater in the IABP group, although only 6% required surgical repair for vascular injury. Recovery of global and non-infarct zone left ventricular function was greater among patients receiving IABP support. Non-infarct zone myocardial function has been identified as a prognostic factor.[26] The IABP may preserve non-infarct zone function by reducing myocardial oxygen consumption, increasing flow across critical stenoses, or by improving collateral flow through the subendocardium. Reinfarction or reocclusion of the infarct-related artery was not observed in the IABP group. The IABP may confer a survival advantage in high-risk patients, especially the "rescue angioplasty" group who appear particularly at risk from reocclusion.

Survival following acute myocardial infarction is independently related to both patency of the infarct-related artery and left ventricular function. In a study undertaken by Ishihara et al,[27] 114 patients with anterior myocardial infarcts underwent emergency percutaneous transluminal coronary angioplasty (PTCA). Forty-eight of these subsequently had an IABP inserted. Reinfarction was not observed in the IABP cohort and reocclusion was significantly lower (2.4% vs 17.7% $P<0.05$). There was a marginally significant improvement in left ventricular ejection fraction in the IABP-treated group ($9.2\pm13.0\%$ vs $4.5\pm12.2\%$ $P=0.08$). There was no significant difference in hospital mortality. Patients were not randomised and changes in angioplasty technique over time may have confounded the results.

A prospective randomised multicentre trial of prophylactic IABC for 48 hours following emergency PTCA in acute myocardial infarction, was conducted by Ohman and co-workers.[28] One hundred and eighty-two patients were enrolled, with 96 assigned

to IABC and 86 to standard management. Patients randomised to the IABP arm had a significantly lower rate of reocclusion of the infarct-related artery at follow-up angiography (8% vs 21% $P<0.03$). IABC also achieved a reduction in adverse events, with a composite clinical end-point of death, stroke, reinfarction, urgent angioplasty or surgical revascularisation, and recurrent ischaemia (13% vs 24% $P<0.04$). Patients presenting in established cardiogenic shock were excluded, however, on the grounds that it would have been unethical to withhold IABC from these individuals

The adjunctive role of IABC in coronary angioplasty will need to be re-evaluated in the light of newer pharmacological therapies, specifically the platelet aggregation inhibitors abciximab and clopidogrel.

**Paediatrics**

The intraaortic balloon pump was slow to find widespread acceptance in paediatric practice during the late 1970s and early 80s. An early report by Pollock et al[29] suggested that it was not a successful strategy in children under five years of age and overall survival was significantly less than 50%. Technical challenges confronting enthusiasts for counterpulsation included difficulties in balloon insertion, a limited range of balloon volumes and lengths, pumping consoles restricted to delivering volumes no less than 20 ml at lower rates, and poor diastolic pressure augmentation in young children with a highly distensible aorta. Right ventricular failure and pulmonary hypertension may contribute to low cardiac output states in children and may portend a poor response to the balloon pump. However, a range of balloons from 2.5 to 15 ml is now readily available and consoles using helium instead of $CO_2$ can deliver smaller volumes at higher heart rates.[30] Good outcomes and associated clinical improvements in haemodynamics and renal function have been noted, despite inability to achieve suprasystolic diastolic augmentation.[31]

Webster and Veasy[32] reported a series of 18 children who underwent counter-pulsation. Of these, 50% survived to wean from the IABP, with 86% of these survivors weighing less than 15 kg. In this series, 16 patients had undergone corrective cardiac surgery or coronary bypass grafting and two suffered cardiogenic shock, one from viral myocarditis and the other Haemophilus influenzae sepsis. These authors suggested indications for IABP should include selected congenital heart lesions, myocarditis and sepsis if myocardial recovery is anticipated. However, there were only five long-term survivors among their cohort. A further nine children undergoing balloon counter-pulsation for refractory ventricular failure, inability to wean from cardiopulmonary bypass and myocardial ischaemia were reported by Park et al.[30] Four survived weaning from the IABP after support extending to 96 hours. In another study of 14 children with a median age of 3 years, a 71% rate of successful weaning from IABP support was observed leading to a 57% long-term survival following cardiac surgery.[33]

Traditionally ECMO has been used for postcardiotomy support of children, with reports of 44% successfully weaning.[33] The association between congenital heart defects, right ventricular failure and abnormal pulmonary vascular resistance often compounds the left ventricular dysfunction ensuing from complicated surgery with long cross clamp times. ECMO may be expected to achieve greater success under these circumstances. The balloon pump assists the left ventricular ejection phase while augmenting diastolic coronary perfusion, improving the endocardial viability ratio. Thus, IABC is ideally suited to conditions characterised by myocardial ischaemia and permits a reduction in inotrope dosages while preserving organ perfusion. Patients

undergoing the Fontan Procedure might be expected to benefit from IABP assisted reduction of left ventricular afterload and filling pressures, but clinical experience has been mixed. Long-term survival of children requiring counterpulsation after Fontan has varied from 11% to 28% in small published series.[33, 34] Many of the deaths reported in the literature have occurred in cardiac surgical patients subsequently shown at autopsy to have severe residual cardiac lesions incompatible with long-term survival, rather than as a direct consequence of balloon pump complications or a failure to favourably alter haemodynamics and end-organ perfusion.[32, 33]

Complications from balloon counterpulsation in children have been comparable to adults. They range from mild limb ischaemia[32] and loss of distal pulses without associated tissue necrosis[30] to fatal mesenteric ischaemia.[33] Premature removal of the IABP mandated by life- or limb-threatening complications appears uncommon.

**New Directions**

*Subarachnoid haemorrhage*

New applications for the balloon pump have been reported in the literature. Counterpulsation has been proposed to support patients with cerebral vasospasm and ventricular dysfunction after subarachnoid haemorrhage.[35,36] It may assist in sustaining cerebral perfusion in conjunction with hypervolaemic resuscitation, haemodilution and systemic hypertension. While an attractive approach to minimising vasopressor doses and improving haemodynamic stability among recipients of nimodipine, it necessitates anticoagulation in patients who are at risk of further cerebral haemorrhage.

*High-risk cardiac patients undergoing general surgery*

These patients may benefit from perioperative circulatory support using the balloon pump,[37] although questions regarding risk-benefit, patient selection and postoperative care need to be addressed.

*Refractory arrhythmias*

A recent report suggested that patients experiencing ventricular arrhythmias, refractory to conventional medical therapy, may respond to intraaortic counterpulsation.[38] The authors speculate that ventricular unloading (mechano-electrical feedback) and reduced catecholamine levels, rather than protection against ischaemia, may explain the beneficial effect on cardiac rhythm.

*Septic shock*

Advanced septic shock may deteriorate to a hypodynamic state as a consequence of decreased coronary perfusion and cytokine mediated myocardial depression. An early laboratory study by Roberts et al[39] employing a dog model of Klebsiella-induced sepsis tested the hypothesis that intraaortic counterpulsation would favourably affect cardiorespiratory, metabolic and haematological responses. They found that IABP supported animals at 24 hours after bacterial infusion exhibited better preserved left ventricular stroke work and myocardial performance, less metabolic acidosis and recovery of circulating leukocyte count. Application in clinical septic shock has been studied infrequently. One early report notes the survival of two patients with coronary artery disease complicated by severe septic shock managed by IABC.[40]

*Right heart failure*

Right ventricular infarction and consequent failure may lead to overdistension and compromised left ventricular filling despite volume resuscitation owing to ventricular interdependence. Right ventricular ischaemia will be exacerbated by high end-diastolic filling pressures, which increase oxygen demand while imperilling coronary supply. Intraaortic counterpulsation may play a role in treating right heart failure refractory to inotropic support.[41] Massive pulmonary embolism accompanied by right ventricular overload and ischaemia has similarly been proposed as an indication amenable to IABC.[42]

To date, these proposed indications for the balloon pump remain novel therapies with only anecdotal or small retrospective case series for support. Further research is necessary before any recommendations can be made.

**Weaning and Removal**

Discontinuation of balloon pump support remains a difficult judgement based on the initial indication for counterpulsation, the degree of myocardial recovery as demonstrated by echocardiography and invasive haemodynamic monitoring, the level of inotropic support and the natural history of the disease process. Catheters are readily maintained in-situ up to seven days and beyond, however risks of infection and vascular injury or thromboembolic complications assume greater importance over time.

The nature of the circulatory assistance offered by the IABP is subtle, being the nett interaction of afterload reduction, relief from ischaemia, enhanced end organ perfusion, and, perhaps, protection against serious arrhythmias. Except for severe refractory ventricular failure, placing the IABP console into stand-by mode will be unlikely to precipitate discernible change in haemodynamics initially. Over the ensuing hours a persistent low cardiac output state may manifest in progressive metabolic acidosis, oliguric renal failure and respiratory failure. One approach is to wean the augmentation or ratio of IABP support over several hours, while observing signs of end-organ hypoperfusion associated with an inadequate cardiac index. Myocardial recovery may be ascertained by echocardiography prior to catheter removal. Having achieved a reduction in inotropic support as a beneficial consequence of cardiac rest and recovery offers scope to offset the withdrawal of mechanical circulatory assistance by adjusting inotropic doses upwards.

Choosing the sequence for discontinuing IABP and ventilatory support depends on the condition of the individual patient. Factors to be considered include mobilisation of a supine patient tethered to a femoral catheter following extubation, the propensity for pulmonary congestion after weaning from mechanical ventilation and withdrawal of circulatory support, and relative priorities in avoiding respiratory and vascular complications.

**Complications**

Morbidity attributable to balloon pumps has included balloon rupture, aortic dissection and perforation, haemorrhage, limb ischaemia, mesenteric infarction, renal ischaemia and infection. However, with the advent of the percutaneous approach, narrow gauge catheters and sheathless insertion technique and, perhaps, greater vigilance and early removal in the event of complications, serious adverse outcomes and deaths have declined sharply. In the largest series published by the Massachusetts

General Hospital,[7] catastrophic vascular complications including iliac or aortic perforation, mesenteric infarction and amputation were less than 1%. Naunheim et al[6] also found that balloon pump related complications did not adversely affect survival. The renewed interest in combining IABC with thrombolysis in the treatment of acute myocardial infarction has raised concern over possible bleeding. Serious morbidity occurred in 10 to 15% of patients, although there were no deaths as a consequence of thrombolysis and IABC in two studies conducted in community hospitals.[11, 12] The GUSTO-1 Study[8] recorded bleeding outcomes in 62 patients who underwent IABC within 24 hours following thrombolysis. Patients receiving early IABC had significantly more moderate bleeding complications and required greater transfusion of packed cells ($P=0.0001$). However, the incidence of severe or life-threatening complications was not significantly different between the IABP and no-IABP groups ($P=0.16$).

The duration the IABP catheter is maintained in-situ has been identified as an independent risk factor for sepsis, carrying an odds ratio of 1.5 for each pump day.[43] Recently a prospective study of 60 cardiac surgical patients managed with an IABP catheter found a 15% incidence of true bacteraemia and a 12% incidence of sepsis, raising the question of whether antibiotic prophylaxis would be appropriate.[44]

Iatrogenic vascular injury from balloon pump catheter insertion is commonly reported with 25 out of 90 patients suffering vascular complications in one retrospective analysis over a 10-year period.[45] Female gender and peripheral vascular disease were identified as significant risk factors, but the site and difficulty of insertion, duration of counterpulsation, anticoagulation and antibiotics were not predictive of vascular complications.[45] The sheathless insertion technique has been demonstrated to reduce the incidence of vascular complication (8.8% vs 25.9% $P<0.01$) and lower limb ischaemia in particular.[46] In a study of 509 patients over a 15-year period, major vascular complications were identified in 8%, with a notable reduction in the incidence of complications over the last five years of the analysis.[47] The authors attributed the improvement in iatrogenic injury to the use of catheters with smaller diameters. They noted that the presence of vascular complications per se was not a significant independent predictor of mortality or other serious morbiditiy.[47] Lower limb ischaemia associated with balloon pumps in the absence of peripheral vascular disease may be amenable to pharmacological therapy with $PGE_1$ infusion via the intra-aortic catheter.[48]

**Conclusion**

In light of current knowledge, are we using the balloon pump appropriately?

As estimated from the GUSTO-1 study,[8] 32% of shocked patients underwent insertion of an intraaortic balloon in the United States. The corresponding figure for the other participating countries was only 7.2%, and this figure is likely to represent local practice. International data on usage from the Benchmark Registry recorded 20.6% for haemodynamic support associated with catheterisation, 18.8% cardiogenic shock, 16.1% for weaning from cardiopulmonary bypass, 13% preoperative insertion in high-risk patients and 12.3% refractory unstable angina as indications for counterpulsation.[49] These figures suggest a broader application of the technique than occurs locally. Perhaps our fear of complications, lack of availability and unfamiliarity with counterpulsation results in fewer patients benefiting from the device than might otherwise be the case. The literature clearly supports early insertion rather than

salvage therapy. In contrast, clinical practice in many centres has seen the majority of balloons inserted "after the fact" in theatre or, even later, in the ICU, possibly disadvantaging our patients in terms of morbidity and survival.

At this stage, combining lysis with counterpulsation in cardiogenic shock remains controversial, although promising. If proven to be of benefit, the implications particularly for regional hospitals without a cardiac surgical service would be significant. Acquisition of the equipment and training of personnel in regional centres would pose substantive logistical challenges.

A study of the major determinants of survival with counterpulsation, published in the *American Heart Journal* in 1995,[17] found that myocardial infarction with a LVEF <30% carried a high associated mortality. The limited haemodynamic support offered by the balloon pump may be inadequate in the face of severe shock. Identifying these patients early and triaging to a ventricular assist device may be more appropriate.

The decision to wean from counterpulsation remains a difficult balance between the potential for further recovery and the likelihood of complications should we delay. Upon withdrawal of the balloon pump, deterioration may be slow, manifesting hours later. Many questions regarding timing and the order of weaning from cardiorespiratory support remain unanswered by available research.

In conclusion, the intraaortic balloon pump has gained in popularity as a means of controlling refractory myocardial ischaemia and may further offer protection against coronary reocclusion and reinfarction after angioplasty. A potential new role in treating cardiogenic shock in combination with thrombolysis may offer a survival advantage, especially to those patients presenting to peripheral centres denied immediate access to revascularisation procedures. The use of IABP in cardiac surgery has remained largely unchanged, however retrospective data suggests that earlier preoperative insertion of the IABP may confer a survival advantage by stabilising shocked patients at risk of further ischaemia and infarction. Prospective randomised trials are in progress to answer questions regarding the efficacy and safety of these approaches to balloon counterpulsation in modern practice.

**References**

1. Kantrowitz A, Tjonneland S, Freed PS, Phillips SJ, Butner AN, Sherman JL. Initial clinical experience with intraaortic balloon pumping in cardiogenic shock. *JAMA* 1968; 203:135-140.
2. Buckley MJ, Craven JM, Gold HK. IABP assist for cardiogenic shock after cardiopulmonary bypass. *Circulation* 1972; 46(Suppl):II76.
3. Sakamoto T, Suzuki A, Kazama S, Komatsu S, Sasaki S, Shoji Y. Effects of timing on ventriculoarterial coupling and mechanical efficiency during intraaortic balloon pumping. *ASAIO J* 1995; 41:M580-583.
4. Jacob AS, Steingart RH, Schweger MJ. Heparin elimination following continuous infusion during intraaortic balloon counterpulsation. *Cathet Cardiovasc Diagn* 1985; 11:389-392.
5. Creswell LL, Rosenbloom M, Cox JL et al. Intraaortic balloon counterpulsation: patterns of usage and outcome in cardiac surgery patients. *Ann Thorac Surg* 1992; 54:11-20.
6. Naunheim KS, Swartz MT, Pennington DG et al. Intraaortic balloon pumping in patients requiring cardiac operations: risk analysis and long-term follow-up. *J Thorac Cardiovas Surg* 1992; 104:1654-1661.
7. Torchiana DF, Hirsch G, Buckley MJ et al. Intraaortic balloon pumping for cardiac support: trends in practice and outcome, 1968-1995. *J Thorac Cardiovasc Surg* 1997; 113:758-769.
8. Anderson RD, Ohman EM, Holmes DR et al. Use of intraaortic balloon counterpulsation in patients presenting with cardiogenic shock: observations from the GUSTO-1 study. *J Am Coll Cardiol* 1997; 30:708-715.
9. O'Rourke MF, Norris RM, Campbell TJ, Chang VP, Sammel NL. Randomized controlled trial of intraaortic balloon counterpulsation in early myocardial infarction with acute heart failure. *Am J Cardiol* 1981; 47:815-820.

10. Flaherty JT, Becker LC, Weiss JL et al. Results of a randomized prospective trial of intraaortic balloon counterpulsation and intravenous nitroglycerin in patients with acute myocardial infarction. *J Am Coll* Cardiol 1985; 6:434-446.
11. Stomel RJ, Rasak M, Bates ER. Treatment strategies for acute myocardial infarction complicated by cardiogenic shock in a community hospital. *Chest* 1994; 105:997-1002.
12. Kovack PJ, Rasak MA, Bates ER, Ohman EM, Stomel RJ. Thrombolysis plus aortic counterpulsation: improved survival in patients who present to community hospitals with cardiogenic shock. *J Am Coll Cardiol* 1997; 29:1454-1458.
13. Gurbel PA, Anderson RD, MacCord CS et al. Arterial diastolic pressure augmentation by intraaortic balloon counterpulsation enhances the onset of coronary artery reperfusion by thrombolytic therapy. *Circulation* 1994; 89:361-365.
14. Prewitt RM, Gu S, Schick U, Ducas J. Intraaortic balloon counterpulsation enhances coronary thrombolysis induced by intravenous administration of a thrombolytic agent. *J Am Coll Cardiol* 1994; 23:794-798.
15. Hochman JS, Sleeper LA, Webb JG et al. Early revascularization in acute myocardial infarction complicated by cardiogenic shock. *NEJM* 1999; 341:625-634.
16. Baldwin RT, Slogoff S, Noon GP, Sekela M, Frazier OH, Edelman SK, Vaughn WK. A model to predict survival at time of postcardiotomy intraaortic balloon pump insertion. *Ann Thorac Surg* 1993; 55:908-913.
17. Pi K, Block PC, Warner MG, Diethrich EB. Major determinants of survival and nonsurvival of intraaortic balloon pumping. *Am Heart J* 1995; 130:849-853.
18. Christenson JT, Simonet F, Badel P, Schmuziger M. Optimal timing of preoperative intraaortic balloon pump support in high-risk coronary patients. *Ann Thorac Surg* 1999; 68:934-939.
19. Gunstensen J, Goldman BS, Scully HE, Huckell VF, Adelman AG. Evolving indications for preoperative intraaortic balloon pump assistance. *Ann Thorac Surg* 1976; 22:535-546.
20. Christakis GT, Weisel RD, Fremes SE et al. Coronary artery bypass grafting in patients with poor ventricular function. *J Thorac Cardiovasc Surg* 1992; 103:1083-1092.
21. Dietl CA, Berkheimer MD, Woods EL, Gilbert CL, Pharr WF, Benoit CH. Efficacy and cost-effectiveness of preoperative IABP in patients with ejection fraction of 0.25 or less. *Ann Thorac Surg* 1996; 62:401-409.
22. Christenson JT, Simonet F, Badel P, Schmuziger M. Evaluation of preoperative intraaortic balloon pump support in high-risk coronary patients. *Eur J Cardiothorac Surg* 1997; 11:1097-1103.
23. Kang N, Edwards M, Larbalestier R. Preoperative intraaortic balloon pumps in high-risk patients undergoing open heart surgery. *Ann Thorac Surg* 2001; 72:54-57.
24. Tung C, Sauri D, Fintel D, Gheorghiade M. Management of high-risk subsets in unstable angina. *Cardiology Clinics* 1999; 17:415-437.
25. Ohman EM, Califf RM, George BS et al. The use of intraaortic balloon pumping as an adjunct to reperfusion therapy in acute myocardial infarction. *Am Heart J* 1991; 121:895-901.
26. Grines CL, Topol EJ, Califf RM et al. Prognostic implications and predictors of enhanced regional wall motion of non-infarct zone after thrombolysis and angioplasty therapy of acute myocardial infarction. The TAMI study groups. *Circulation* 1989; 80:245-253.
27. Ishihara M, Sato H, Tateishi H, Uchida T, Dote K. Intraaortic balloon pumping as the postangioplasty strategy in acute myocardial infarction. *Am Heart J* 1991; 122:385-389.
28. Ohman EM, George BS, White CJ et al. Use of aortic counterpulsation to improve sustained coronary artery patency during acute myocardial infarction. Results of a randomized trial. *Circulation* 1994; 90:792-799.
29. Pollock JC, Charlton MD, Williams WG, Edmonds JF, Trusler GA. Intraaortic balloon pumping in children. *Ann Thorac Surg* 1980; 29:522.
30. Park JK, Hsu DT, Gersony WM. Intraaortic balloon pump management of refractory congestive heart failure in children. *Pediatr Cardiol* 1993; 14:19-22.
31. Christensen DW, Veasy LG, McGough EC, Dean JM. Intraaortic balloon counterpulsation in children: a review of 29 patients. *Crit Care Med* 1991; 19:S75.
32. Webster H, Veasy LG. Intraaortic balloon pumping in children. *Heart and Lung* 1985; 14:548-555.
33. Akomea-Agyin C, Kejriwal NK, Franks R, Driscoll Booker P, Pozzi M. Intraaortic balloon pumping in children. *Ann Thorac Surg* 1999; 67:1415-1420.
34. Del Nido PJ, Swan PR, Benson LN. Successful use of intraaortic balloon pumping in a 2-kg infant. *Ann Thorac Surg* 1988; 46:574-576.

35. Apostolides PJ, Greene KA, Zabramski JM, Fitzgerald JW, Spetzler RF. Intraaortic balloon pump counterpulsation in the management of concomitant cerebral vasospasm and cardiac failure after subarachnoid haemorrhage: technical case report. *Neurosurgery* 1996; 38:1056-1059.
36. Nussbaum ES, Sebring LA, Ganz WF, Madison MT. Intraaortic balloon counterpulsation augments cerebral blood flow in the patient with cerebral vasospasm: a xenon-enhanced computed tomography study. *Neurosurgery* 1998; 42:206-213.
37. Masaki E, Takinami M, Kurata Y, Kagaya S, Ahmed A. Anaesthetic management of high-risk cardiac patients undergoing noncardiac surgery under the support of intraaortic balloon pump. *J Clin Anesth* 1999; 11:342-345.
38. Fotopoulos GD, Mason MJ, Walker S et al. Stabilisation of medically refractory ventricular arrhythmia by intraaortic balloon counterpulsation. *Heart* 1999; 82:96-100.
39. Roberts AJ, Hoover EL, Alonso DR, Combes JR, Dineen P, Gay WA, Subramanian VA. Prolonged intraaortic balloon pumping in klebsiella-induced hypodynamic shock: cardiopulmonary, hematological, metabolic, and pathological observations. *Ann Thorac Surg* 1979; 28:73-86.
40. Foster ED, Subramanian VA, Vito L et al. Response to intra-aortic balloon pumping. *Am J Surg* 1975; 129:464.
41. Hollenberg SM, Kavinsky CJ, Parrillo JE. Cardiogenic Shock. *Ann Intern Med* 1999; 131:47-59.
42. Weber KT, Janicki JS. Intraaortic balloon counterpulsation. A review of physiological principles, clinical results, and device safety. *Ann Thorac Surg* 1974; 17:602-636.
43. Aksnes J, Abdelnoor M, Berge V, Fjeld NB. Risk factors of septicaemia and perioperative myocardial infarction in a cohort of patients supported with intraaortic balloon pump (IABP) in the course of open heart surgery. *Eur J Cardiothorac Surg* 1993; 7:153-157.
44. Crystal E, Borer A, Gilad J et al. Incidence and clinical significance of bacteraemia and sepsis among cardiac patients treated with intraaortic balloon counterpulsation pump. *Am J Cardiol* 2000; 86:1281-1284.
45. Alle KM, White GH, Harris JP, May J, Baird D. Iatrogenic vascular trauma associated with intraaortic balloon pumping: identification of risk factors. *Am Surg* 1993; 59:813-817.
46. Tatar H, Cicek S, Demirkilic U, Ozal E, Suer H, Aslan M, Ozturk OY. Vascular complications of intraaortic balloon pumping: unsheathed versus sheathed insertion. *Ann Thorac Surg* 1993; 55:1518-1521.
47. Arafa OE, Pedersen TH, Svennevig JL, Fosse E, Geiran OR. Vascular complications of the intraaortic balloon pump in patients undergoing open heart operations: 15-year experience. *Ann Thorac Surg* 1999; 67:645-651.
48. Nakano T, Tominaga R, Shiraishi K, Yasui H. Prostaglandin $E_1$ from the tip of an intraaortic balloon catheter for lower limb ischaemia. *Ann Thorac Surg* 1998; 65:1158-1160.
49. Ferguson JJ, Cohen M, Freedman RJ, Stone GW, Miller MF, Joseph DL, Ohman EM. The current practice of intraaortic balloon counterpulsation: results from the benchmark registry. *J Am Coll Cardiol* 2001; 38:1456-1462.

# Pulse Contour Analysis and Transpulmonary Thermodilution

MARK LENNON, FCARCSI, FANZCA

*Dr Lennon is a Senior Registrar in Anaesthesia at Sir Charles Gairdner Hospital in Western Australia. He is also completing specialist training in Intensive Care.*

"The pulse ranks first among our guides; no surgeon can despise its counsel, no physician shut its ears to its appeal. Since then, the information the pulse affords is of so great importance ... surely it must be to our advantage to appreciate fully what it tells us, and to draw from it all that it is capable of imparting."

— (*Dr Mohammed. Turn of the century)*

## INTRODUCTION

The measurement of cardiac output and assessment of a patient's volume status is clearly a useful aim for many critically ill patients and for a number of patients undergoing major surgery, particularly cardiac procedures. While bedside clinical methods of assessing organ perfusion will not be replaced, in certain situations more advanced monitoring is nevertheless required to assess the effects of potent vasoactive agents and monitor the progression of serious illnesses.

While debate over the benefits and risks of the pulmonary artery catheter (PAC) remains unresolved after 30 years, a new generation of hemodynamic monitors have evolved and are ready to compete for the attention of sceptical clinicians throughout the world. Expenditure on these devices in the US has increased from $6.3 million in 1999 to $32.9 million in 2002. By 2006, this is expected to reach $61.5 million. Concerns about the safety and accuracy of the PAC and perhaps the unfair pressure of expectation placed on the monitor to alter outcome has led to its gradual fall from grace. This article will look at the alternative technique of pulse contour analysis combined with transpulmonary thermodilution (TPTD) using the PiCCO device (Pulsion Medical Systems, Munich, Germany). The basic principles involved will be explained and its application in clinical practice both in anaesthesia and intensive care will be assessed.

## HISTORICAL PERSPECTIVE

Since the discovery of circulation by Harvey in 1616, theorists, mathematicians, physicists, physiologists and, of course, doctors have been developing methods to measure cardiac output and have yet to reach a consensus. Adolf Fick first proposed a method to measure cardiac output in 1870, the principles of which still apply today. In 1897, Stewart introduced indicator dilution theory but the first attempt at analysing the *pulse* to yield information on cardiac output was ascribed to a German physicist-mathematician turned physiologist, Otto Frank in 1899. A few years later in 1904,

Erlanger published a study on the relationship between blood pressure and pulse pressure. While work on pulse contour analysis lost momentum thereafter, indicator dilution techniques were developed by Hamilton and colleagues over the next 25 years. Thermodilution was subsequently introduced as a method of measuring blood flow by Fegler in 1954. At the time the idea that heat could serve as an indicator was received with scepticism. In the same year Isaac Starr performed experiments on cadavers simulating systole and studied the relationship between pulse pressure and stroke volume.[1] A simple formula was derived whereby the stroke volume could be calculated from the blood pressure and age of the patient. Comparisons with other contemporary methods were surprisingly good. This was one of the earlier demonstrations of the non-linearity of aortic compliance with age.

It was not until the 1960s, however, that practical results from pulse contour analysis began to appear. Formulae derived by Herd, Warner, Kouchoukos and Remington were studied in animals initially, and usually compared to a direct form of cardiac output monitoring using electromagnetic flow measurement.[2] While reasonable correlation figures were obtained, they lacked precision. A factor common to these earlier formulae was a lack of correction for individual aortic impedance. Subsequent work by Wesseling and colleagues in the 1970s set about correcting for this and has continued until the present day.[3]

## THE THEORY OF PULSE CONTOUR ANALYSIS

The basic algorithm for the determination of cardiac output from the pulse contour was developed by Wesseling. In a complex 37 page document, published in 1983, paradoxically entitled "A *simple* device for continuous measurement of cardiac output", Wesseling describes the mathematical principles involved in great detail.[4]

Put more simply, the arterial pressure waveform is a result of the interaction between the heart and the vascular tree. Left ventricular ejection creates a wave that travels through the arteries (~6-10 m/s) and is reflected from the periphery. This reflection of the arterial pressure wave, known as the Wetterer hypothesis, is a summation of reflected waves arising at bifurcations near the heart and reflected waves from the peripheral part of the system.[5] The forward wave in turn may be divided into a wave that is set up by the heart (the primary wave), and a wave that is the backward wave reflected at the heart. Theoretically if these reflected waves were absent then the pressure and flow contours would be identical, the size of the wave being related to aortic impedance. Differences between pressure and flow contours in the aorta occur because the reflected waves cause an increase in the pressure but a decrease in flow. A simple analogy is a wave reflected from a cliff causing incoming waves to increase in height but lose some power.

According to Wesseling's algorithm, the left ventricular stroke volume (SV) is computed by measuring the area under the systolic portion of the arterial pressure waveform, which is then divided by the aortic impedance (Figure 1). The SV is multiplied by heart rate to give the cardiac output. To calculate for individual aortic impedance, an arterial thermodilution measurement is performed simultaneously to calibrate the system.

The calculation as utilised by the PiCCO system (Pulsion Medical Systems, Munich, Germany) for continuous cardiac output incorporates a patient specific calibration factor, the heart rate, the area under the pressure curve, and finally the compliance and shape of the pressure curve (Figure 2).

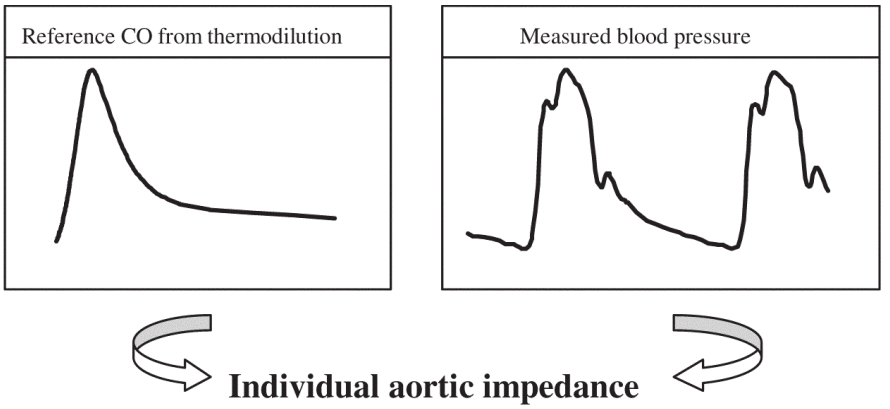


**Figure 1**. Typical thermodilution and arterial blood pressure curves from which individual aortic impedance is derived (see text for explanation).

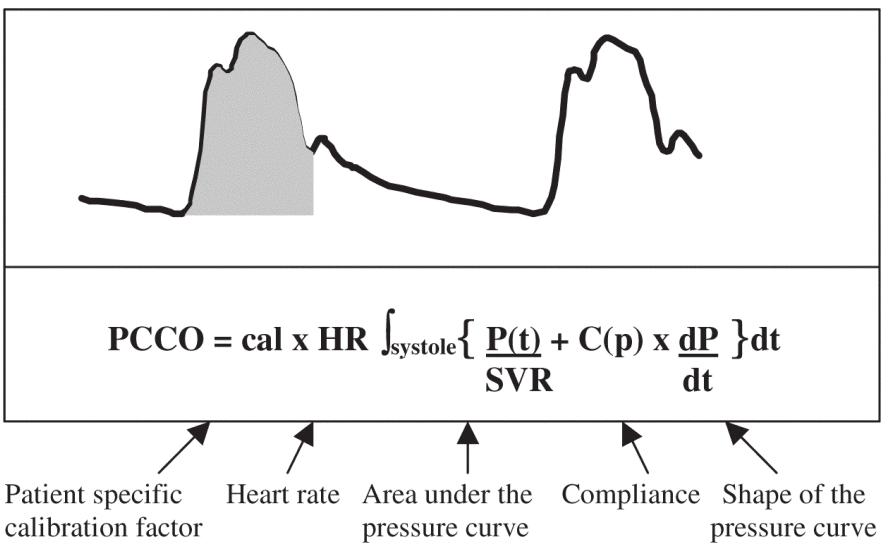


**Figure 2**. Calculation used by the PiCCO for measurement of continuous cardiac output.

## TRANSPULMONARY THERMODILUTION (TPTD)

The thermodilution technique of the PiCCO system requires a standard central venous catheter (CVC) and a 4F thermistor tipped femoral (or axillary) artery catheter. Based on the same principle of pulmonary thermodilution, a cold indicator is injected into the CVC and the temperature change is detected at the femoral artery catheter. The cardiac output is calculated in the usual manner using the Stewart-Hamilton algorithm. This technique differs to pulmonary thermodilution in some respects. The longer distance to travel results in a longer, flatter curve and is thus more sensitive to baseline thermal variations. To reduce the signal to noise ratio, the injectate is cooled to ~4°C and the volume injected can be increased depending on the

size of the patient. The reading however is less affected by variations in respiration, slowing of the heart rate, and tricuspid regurgitation, which have been shown to affect the PAC readings. It is also noted that en route to the femoral artery thermistor the indicator will have simultaneously flowed into other vascular beds (cerebral, upper limb, gut, kidneys). This does not alter the validity of the result, however, as the ratio of local blood flow (in the distal aorta) and local recovery of indicator is equal at all sites of detection, as is the area under the thermodilution curve.[6]

## ASSESSMENT OF INTRAVASCULAR VOLUME BY INDICATOR DILUTION

In addition to the cardiac output result, a measure of the volume status of the patient can be obtained by analysing characteristic transit times of the indicator. After injection of the indicator, the particles mix thoroughly with the flowing blood but do not reach the point of detection (femoral artery) simultaneously. Rather, each particle has its own particular transit time. **The Mean Transit Time (MTt)** is taken as the time for the first particle to be detected plus the mean time difference between the detection of the first particle and all the following particles[7] (Figure 3).

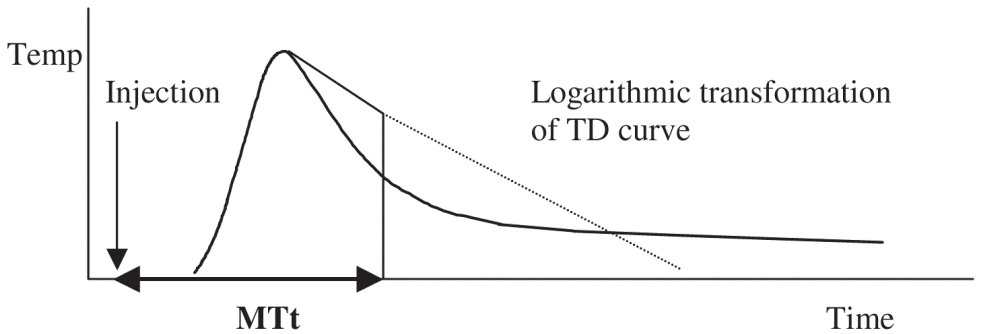


**Figure 3.** Estimation of mean transit time (MTt) using logarithmic transformation of the thermodilution (TD) curve.

Given that we have measured the flow between the two points, the cardiac output (CO), which represents volume per unit time, and we have also determined the time taken for the indicator to travel the distance, we can estimate the volume by multiplying the CO by the transit time.[8]

$$\mathbf{CO \times MTt = Volume}$$

Where CO=cardiac output and MTt=mean transit time

For an indicator that remains in the intravascular space, such as indocyanine green (ICG) this represents the blood volume from the point of injection to the point of detection and is termed the **Intrathoracic Blood Volume (ITBV)**.[9]

thus for ICG: $\mathbf{CO \times MTt = ITBV}$

Note: The volume between the point of injection and the point of detection is not a closed circuit and clearly some indicator has flowed elsewhere (into the cerebral, upper limb and splanchnic circulation), which has not been detected by the thermistor. While this has not been clearly explained by most investigators, it has to be assumed that the transit time we are measuring relates only to the indicator particles which have taken

the direct route to the femoral artery and thus, this is the volume which it reflects. By the same logic we can place the catheter in the axillary artery and get a result, which is essentially the same but does not include the descending aortic volume.

We know that a thermal indicator, however, rapidly equilibrates between the intra and extravascular space (~×100 times faster than molecular diffusion) so the total volume measured in this way is termed the **Intrathoracic Thermal Volume (ITTV)**[10] (Figure 4).

for a thermal indicator: **CO × MTt=ITTV**

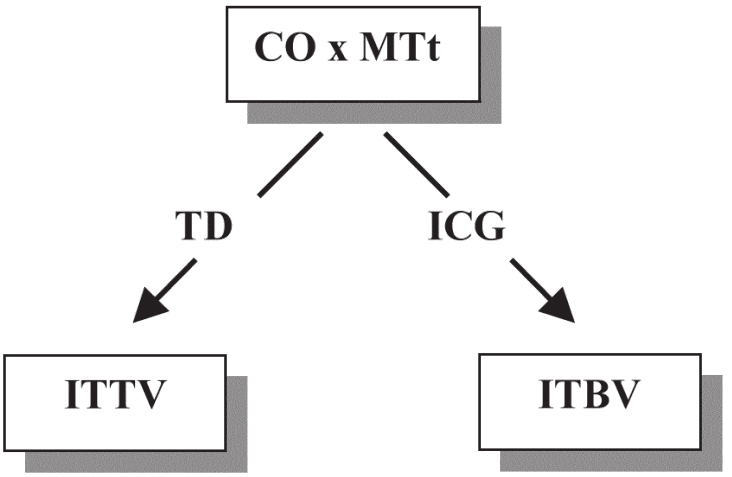


**Figure 4.** Depending on the indicator used for the estimation of MTt (ie thermodilution, TD, or indocyanine green, ICG) the derived volume will be either intrathoracic thermal volume (ITTV) or intrathoracic blood volume (ITBV).

The **ITBV** is composed of the volume of blood in the heart, **or Global End-Diastolic Volume (GEDV)**, the **Pulmonary Blood Volume (PBV)**, and the volume of blood in the aorta to the point of detection.

**ITBV=GEDV+PBV**

Where ITBV=intrathoracic blood volume, GEDV=global end-diastolic volume, and PBV=pulmonary blood volume.

The **ITTV** is composed of the ITBV (as above) plus the extravascular compartment, termed **Extravascular Lung Water (EVLW)** (Figure 5).

**ITTV=GEDV+PBV+EVLW**

Where ITTV=intrathoracic thermal volume, GEDV=global end-diastolic volume, PBV=pulmonary blood volume, and EVLW=extravascular lung water.

## MEASUREMENT OF EXTRAVASCULAR LUNG WATER

Depending on the type of indicator used either the ITBV or the ITTV can be measured. If both indicators are used simultaneously, as previous monitors used to do (COLD medical systems, Munich, Germany), then simply by subtracting the ITBV from the ITTV, the resulting volume is the EVLW.

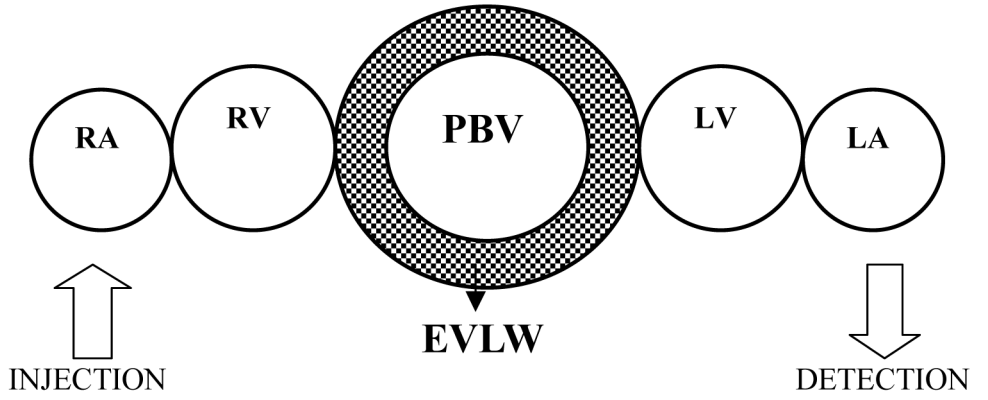


**Figure 5.** Diagrammatic representation of extravascular lung water (EVLW). Global end-diastolic volume (GEDV)=sum of right atrial (RA), right ventricular (RV), left ventricular (LV), and left atrial (LA) end-diastolic volumes. Intrathoracic thermal volume (ITTV)=GEDV+pulmonary blood volume (PBV)+ EVLW.

**EVLW=ITTV-ITBV**

As the ICG techniques were time consuming and more expensive, the PiCCO monitor was developed to use only thermodilution. How then does it give us all of the information? As explained above it can tell us the ITTV by multiplying the CO by the MTt. Using one more characteristic transit time, termed the **Exponential Downslope Time (DSt)**, the **Pulmonary Thermal Volume (PTV)** can be measured. PTV is the sum of PBV and the EVLW.

**PTV=PBV+EVLW**

This is explained by Newman's theory that an indicator will always mix evenly, and in a series of chambers the dilution curve of the indicator will depend on the largest mixing chamber which will be reflected by an exponential decay curve.[11] Although the pulmonary thermal volume is not actually the *largest* chamber by volume, for the purposes of thermal exchange, it is the most significant given the surface area involved. Hence by the same principles as above, by multiplying the CO by the DSt, we can measure the PTV.

**CO × MTt=ITTV**

**CO × DSt=PTV**

By subtracting the PTV from the ITTV, we are left with the GEDV.

**GEDV=ITTV-PTV**

In the final calculation, by using a structural regression analysis, the mathematical relationship between GEDV and ITBV has been established in patient population studies.[12] This regression equation is used to estimate ITBV from GEDV.

**ITBV=1.25 × GEDV**

Using the estimated ITBV, an estimated EVLW can be calculated also.

**EVLW=ITTV-ITBV**

There are additional derived variables the PiCCO monitor also displays, the **systemic vascular resistance (SVR)** and the **stroke volume variation (SVV)**. The latter has been used as a measure of volume responsiveness, which may be of particular use in patients who are on fully controlled mechanical ventilation.

## DISCUSSION

Pulse contour analysis provides a **continuous measurement of cardiac output**. A continuous rather than an intermittent measure of cardiac output may be more useful in patients who experience sudden changes in haemodynamic state that you would wish to recognise and treat without delay, such as during "off-pump" cardiac revascularisation, aortic cross clamping for aneurysm repair, and other major vascular procedures in patients with cardiac disease. In the intensive care unit, the unstable patient with septic shock is the most obvious example especially as more potent vasoactive agents such as vasopressin, are increasingly being used. There are many other patients who might also benefit from this form of monitoring including patients with severe burns, cardiogenic shock post-myocardial infarction, and labile cardio-thoracic surgery patients in the early post operative period. Critically ill paediatric patients who are either too small or in whom it is too difficult to place a PAC may also benefit from this form of monitoring.

The **Intrathoracic Blood Volume (ITBV)** is a measure of **cardiac preload**. This can be indexed against body surface area and is usually about one third of the blood volume, with a normal range of 850-1000 ml/m$^2$. This has been compared to CVP and pulmonary capillary wedge pressure (PCWP) as an indicator of preload in animal experiments by Pfeiffer.[13] A quote from the author "*...The results of this study demonstrate that the interpretation of PCWP as an indicator of circulating volume could be complete nonsense.*" Similar studies performed by Bindels[14] and Sakka[15] in humans comparing ITBVI to CVP and PCWP in critically ill ICU patients revealed good correlation between ITBV and aortic stroke volume but poor correlation between PCWP and stroke volume. Hinder assessed volume status by transoesophageal echocardiography and indicator dilution techniques in 15 patients undergoing cardiac surgery.[16] He concluded that although CVP and PCWP correlated with each other, they did not correlate with either left ventricular end-diastolic area (LVEDA) nor ITBVI, while the latter two, given certain conditions were essentially interchangeable. Several other studies looking at ITBV as a marker of preload seem to suggest reliability.

**Extravascular Lung Water (EVLW)** can be used as a measure of **fluid overload**. In patients with acute respiratory distress syndrome, and many forms of respiratory failure, the EVLW indices may prove very useful. The normal range is ~5-7 ml/kg. Greater than 10 ml/kg may predict fluid overload. In general terms this may be more useful to intensivists than anaesthetists as the former group often has to rescue near drowning victims from the latter! Clearly, this is one area of intensive care where we can be more scientific. Beyond our clinical assessment and fluid balance indices, other common markers of fluid overload such as worsening pulmonary compliance and oxygenation are perhaps "too little, too late". Although there are scoring systems for grading pulmonary oedema on a CXR, the usual comments on the morning X-ray rounds, such as ..."*it looks a little wet*" seems to lack scientific rigour. A number of studies have looked at the relevance and accuracy of EVLW measurements. The

**Table 1**
Studies comparing pulmonary artery thermodilution cardiac output (COTDpa) vs transpulmonary thermodilution cardiac output (COTDa)

| Author | patients (n)/ observations | COTDa-COTDpa (bias±SD, l/min/m$^2$) | Correlation (r) |
|---|---|---|---|
| McLuckie et al, 1996[22] | 9/162 | 0.19±0.21 | — |
| Goedje et al, 1998[23] | 30/150 | 0.16±0.31 | 0.96 |
| Goedje et al, 1998[24] | 30/810 | 0.26±0.7 | 0.96 |
| Goedje et al, 1999[25] | 24/216 | –0.29±0.66 | 0.93 |
| Sakka et al, 1999[26] | 37/449 | 0.68±0.62 | 0.97 |
| Sakka et al, 2000[27] | 12/51 | 0.73±0.38 | 0.98 |
| Zöllner et al, 2000[28] | 19/76 | 0.21±0.73 | 0.96 |
| Bindels et al, 2000[29] | 45/283 | 0.49±0.45 | 0.95 |

original paper by Lewis in 1979 pointed out potential for error with asymmetric lung disease and macroscopic pulmonary embolisation.[17] Also pointed out by Bock was the issue that EVLW measurements have tended to slightly overestimate lung water.[10] This could be due to loss of thermal indicator, although this was an inconsistent finding in other studies,[18] or more likely is due to heat exchange with non-pulmonary extravascular structures, pulmonary perfusion defects or recirculation of indicator. On the last issue, recirculation of indicator is generally assumed to be accounted for by monoexponential extrapolation of indicator dilution curves. More importantly, Mitchell looked at how strategies aimed at managing EVLW compared to wedge pressure driven fluid management was associated with improved outcome in terms of ICU stay and ventilation days in a prospective randomised trial involving 101 patients.[19] Other data from Sturm has shown a positive correlation between mortality and increased EVLW.[20] This is an interesting area of research which certainly warrants further appraisal.

Is this technology accurate? I have not attempted to do a meta-analysis of the published data but there appears to be reasonable evidence to suggest that this technology is accurate within the limits of clinical acceptability. The results of some of the studies comparing this form of CO monitoring to traditional PAC derived CO are displayed in Table One. Mostly, the PiCCo monitor has been compared to the PAC in terms of cardiac output data. As pointed out by Bland and Altman, good correlation does not necessarily imply good agreement between the two variable monitoring techniques.[21] This is particularly the case when you consider that the clinical gold standard to which the PiCCO has been compared (PAC) has many inherent inaccuracies itself. Incidentally, the manufacturers of the **lithium dilution monitor** (LIDCo, Cambridge, UK) claim that their monitor is three times more accurate than thermodilution methods. It involves injecting a dose of lithium (~240-fold lower than the standard dose for bipolar affective disorder) into a central venous line and the dilution curve is measured by means of an ion-sensitive electrode sensor in the radial arterial line.

Is this technology safe? It is less invasive than a PAC but still requires central venous access and large bore arterial access, which are not without risk. Femoral artery catheterisation resulting in injury necessitating surgical repair has been estimated at 0.28% in one large series.[30] Is it expensive? The femoral artery catheter costs about

$250.00 (Australian) and the additional hardware costs are approximately $17,500 (Australian).

## CONCLUSION

Pulse contour analysis is an exciting new form of haemodynamic monitoring, which shows considerable promise, and which has the potential to extend the amount of physiological information available to guide therapy in critically ill patients. There are already over 70 such monitors in use in Australasia. Having a clearer understanding of how it works is likely to further increase its acceptance into clinical practice. However, further studies are required to examine its cost-benefit and risk-benefit ratios, and to demonstrate its influence, if any, on patient outcome.

## REFERENCES

1. Starr I. Clinical tests of the simple method of estimating cardiac stroke volume from blood pressure and age. *Circulation* 1954; 9:664-681.
2. Thomas B. Monitoring of cardiac output by pulse contour method. *Acta Anaesth Belg* 1978; 29:259-270.
3. Wesseling KH, Jansen JRC, Settels JJ, Schreuder JJ. Computation of aortic flow from pressure in humans using a nonlinear, three element model. *J Appl Physiol* 1993; 74:2566-2573.
4. Wesseling KH, De Wit B, Weber JAP, Smith NT. A simple device for continuous measurement of cardiac output. *Adv Cardiovasc Physiol* 1983; 5:16-52.
5. Westerhof N, Sipkema P, Van Den Bos GC, Elzinga G. Forward and backward waves in the arterial system. *Cardiovasc Res* 1972; 6:648-656.
6. Buhre W, Weyland A, Kazmaier S et al. Comparison of cardiac output assessed by pulse contour analysis and thermodilution in patients undergoing minimally invasive direct coronary artery bypass grafting. *J Cardiothorac Vasc Anesth* 1999; 13:437-440.
7. Effros RM. Lung water measurements with the mean transit time approach. J *Appl Physiol* 1985; 59:673-683.
8. Meier P, Zierler KL. On the theory of indicator dilution method for the measurement of blood flow and volume. *J Appl Physiol* 1954; 6:731-744.
9. Carlile PV, Beckett RC, Gray BA. Relationship between CO and transit times for dye and thermal indicators in central circulation. *J Appl Physiol* 1986; 60:1363-1372.
10. Bock J, Lewis FR. Clinical relevance of lung water measurement with the thermal-dye dilution technique. Practical applications of fibreoptics in critical care monitoring. Lewis FR, Pfeiffer UJ (Eds). Springer-Verlag, New York, 1990; pp. 164-178.
11. Pfeiffer UJ. A fibreoptics based system for integrated monitoring of cardiac output, intrathoracic blood volume, extravascular lung water, oxygen saturation, and a-v differences. In Lewis FR, Pfeiffer UJ (Eds). Practical applications of fibreoptics in critical care monitoring. Springer-Verlag, New York, 1990; 114-127.
12. Sakka SG, Ruhl CC, Pfeiffer UJ et al. Assessment of cardiac preload and extravascular lung water by single transpulmonary thermodilution. *Intens Care Med* 2000; 26:180-187.
13. Pfeiffer UJ, Perker M, Zeravik J, Zimmerman G. Sensitivity of central venous pressure, pulmonary capillary wedge pressure, and intrathoracic blood volume as indicators for acute and chronic hypovolaemia. Practical applications of fibreoptics in critical care monitoring. Lewis FR, Pfeiffer UJ (Eds). Springer-Verlag, New York, 1990; pp. 25-33.
14. Bindels AJ, van der Hoeven JG, Graafland AD, de Koinig J, Meinders AE. Relationship between volume and pressure measurements and stroke volume in critically ill patients. *Crit Care* 2000; 4:193-199.
15. Sakka SG, Bredle DL, Reinhart K, Meier-Hellmann A. Comparison between intrathoracic blood volume and cardiac filling pressures in the early phase of haemodynamic instability of patients with sepsis or septic shock. *J Crit Care* 1999; 14:78-83.
16. Hinder F, Poelaert JI, Schmidt C et al. Assessment of cardiovascular volume status by transoesophageal echocardiography and dye dilution during cardiac surgery. Eur J Anaesthesiol 1998; 15:633-640.
17. Lewis FR, Elings VB, Sturm JA. Bedside measurement of lung water. J Surg Res 1979; 27:250-261.
18. Arfors KE, Malmberg Pavek K. Conservation of thermal indicator in lung circulation. *Cardiovasc Res* 1971; 5:530-534.
19. Mitchell JP, Schuller D, Calandrino FS, Schuster DP. Improved outcome based on fluid management in critically ill patients requiring pulmonary artery catheterization. *Am Rev Respir Dis* 1992; 145:990-998.

20. Sturm JA. Development and significance of lung water measurement in clinical and experimental practice. In Lewis FR, Pfeiffer UJ, Eds. Practical applications of fibreoptics in critical care monitoring. Springer-Verlag, New York, 1990; 129-140.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 8476:307-310.
22. McLuckie A, Murdoch JA, Marsh MJ, Anderson D. A comparison of pulmonary and femoral artery thermodilution cardiac indices in paediatric intensive care patients. *Acta Paediatr* 1996; 85:336-339.
23. Goedje O, Peyerl M, Seebauer T, Dewald O, Reichart B. Reproducibility of double indicator dilution measurements of intrathoracic blood volume compartments, extravascular lung water and liver function. *Chest* 1998; 113:1070-1077.
24. Goedje O, Hoeke K, Lamm P, Schmitz C, Thiel C, Weinert M, Reichart B. Continuous less invasive haemodynamic monitoring in intensive care after cardiac surgery. *Thorac Cardiovasc Surgeon* 1998; 46:242-250.
25. Goedje O, Hoeke K, Lichtwark-Aschoff M, Faltschauser A, Lamm P, Reichart B. Continuous cardiac output by femoral artery thermodilution calibrated pulse contour analysis: comparison with pulmonary artery thermodilution. *Crit Care Med* 1999; 27:2407-2412.
26. Sakka SG, Reinhart K, Meier-Hellmann A. Comparison of pulmonary artery and arterial thermodilution cardiac output in critically ill patients. *Intens Care Med* 1999; 25:843-846.
27. Sakka SG, Reinhart K, Wegscheider K, Meier-Hellmann A. Is the placement of a pulmonary artery cathter still justified solely for the measurement of cardiac output? *J Cardiothorac Vasc Anesth* 2000; 14:119-124.
28. Zollner C, Haller M, Weis M et al. Beat to beat measurement of cardiac output by intravascular pulse contour analysis: A prospective criterion standard study in patients after cardiac surgery. *J Cardiothorac Vasc Anesth* 2000; 14:125-129.
29. Bindels AJ, Van Der Hoeven JG, Graafland AD, de Koinig J, Meinders AE. Relationship between volume and pressure measurements and stroke volume in critically ill patients. *Crit Care* 2000; 4:193-199.
30. Groome J, Vohra R, Cuschieri RJ, Gilmour DG. Vascular injury after catheterization. *Postgrad Med J* 1989; 65:86-88.

# Bringing Nutritional Support on the ICU into the New Millennium

MICHAEL O'LEARY, FRCA, FJFICM
Staff Specialist Intensive Care,
The St George Hospital, Sydney

*Dr O'Leary is a Staff Specialist in Intensive Care at St George Hospital. His research interests include modulation of the catabolic response to sepsis, regulation of muscle protein turnover in critical illness and use of nutrition support in critical illness.*

Anorexia and reduced food intake are characteristic of severe illness. In hospitalised patients, further limitation of nutritional intake occurs because of traditional "nil-by-mouth' practices prior to procedures requiring anaesthesia or sedation and, in many cases, a lack of emphasis on nutrition as part of medical management. For many patients admitted to the intensive care unit (ICU) it is not uncommon for nutritional intake to have been already compromised for some days. Once in the ICU, patients are generally unable to ensure the adequacy of their own nutritional intake. Intuitively, therefore, food must be provided to ICU patients at some point, to stave off the effects of prolonged starvation, but it is not yet defined at which point in illness reduced food intake becomes a problem. Although it is established teaching that nutrition be provided "early" and "enterally" to ICU patients, two recent systematic reviews have come to differing conclusions as to whether early enteral nutrition is of proven benefit.[1,2]

Nutrition is also required in critical illness because, in addition to anorexia and the effects of iatrogenic starvation, there is a pronounced catabolic response that can lead to massive loss of structural body proteins, predominantly from skeletal muscle. The metabolic and physiological responses to starvation and to critical illness are quite different. In the former, both fat and protein are lost, but nitrogen loss is modified until the very late stages by mobilisation of fat, with enhanced fat oxidation being the principal source of calories. These changes are readily reversed by provision of nutrition. Critical illness is characterised by the breakdown of lean tissue to constituent amino acids, which provide substrate for acute-phase protein synthesis and for gluconeogenesis. In this setting, apparently adequate nutritional support can only attenuate the process and breakdown of lean tissue continues. For ICU patients, it is incorrect to consider the effects of the catabolic state and of starvation in isolation, however, as for most patients elements of both processes occur simultaneously.

That starvation is a problem for many hospitalised patients is under-recognised. Ronald Koretz, in 1995, wrote that "... the risk of true starvation will only arise after several weeks of a disease process. Given the expense and risks of the therapy, it is a reasonable policy to wait those weeks out. ..."[3] We know, however, from hunger strikes by Irish Republican prisoners in 1981 that previously healthy young males will die approximately 60 days after ceasing food intake.[4] In the setting of catabolic illness, the effects of starvation on body protein stores may become rapidly apparent, so that if

nutritional support is not provided both early and in adequate amounts loss of body protein may occur.

Robert Graves, in the mid-nineteenth century, was possibly the first physician to appreciate the link between nutritional support and outcome from illness. He showed that mortality from typhus fever was reduced by giving food and drink to patients, overturning the established dogma that treatment of fever consisted of bleeding, purging and starvation! Unfortunately, there are no controlled trials that demonstrate a benefit from nutrition against no nutrition in illness, and, for obvious reasons there are never likely to be. In this chapter, I address the question of how most recent evidence from clinical trials can guide us to best practice in providing nutritional support to patients with critical illness.

**Do feeding targets matter?**

The past 25 years have seen a significant shift in the amount of calories and nitrogen recommended for critically ill patients. In the 1960s and 70s, it was not uncommon for patients to be prescribed up to 40 kcal/kg/day in non-protein energy with up to 3g protein/kg/day. These estimated requirements were predicated on the belief that "negative nitrogen balance" represented the result of massive increases in energy and substrate demand (hypermetabolism), based on results of inappropriate over-estimation of normal resting energy expenditure (REE) and excessive estimates of the impact of illness on REE. Work by Kinney measuring REE and nitrogen balance in surgical and trauma patients demonstrated that, after major surgery, REE increased by only up to 10% and, even following major trauma, only reached a 25% increase.[5] From work by Hill and others, it became clear that in sepsis (for example) providing protein in excess of 1.5g/kg/day does not improve nitrogen balance,[6] and the optimal non-protein energy requirement is between 20-35 kcal/kg/day.[7] Excessive administration of calories and nitrogen leads to hyperglycaemia, hypertriglyceridaemia and hepatic steatosis. Zaloga has even argued that "permissive under-nutrition" should be the standard of care.[8]

Is there any evidence that meeting nutritional targets affects outcome in critically ill patients? There are few published studies directly addressing this issue, and much of the literature on nutrition in critical illness comes from trauma or elective surgical populations which may not be characteristic of most Australian ICUs. Taylor reported the effect of an "enhanced" enteral feeding protocol on outcome in head injury patients.[9] Protocol patients were prescribed feed to a nutritional target from day 1 of ICU admission, whereas control patients had their feeding escalated "as tolerated". Protocol patients received significantly more feed in the first 7 days of ICU stay, and this was associated with reduced infections and general complications, as well as improved neurological recovery at 3 months. There is also some suggestive evidence of a relationship between target feed delivery and outcome from the published enteral immunonutrition trials.[10] In these studies, benefit was only seen in patients who absorbed (or at least received) more than predetermined quantities of feed, or in the setting of assiduously achieved feed delivery targets.

The best evidence to date of an effect of nutrition on outcome in ICU patients may come from a Canadian study currently published as an abstract only.[11] When hospitals were randomised to develop and implement evidence-based nutritional guidelines versus continue with current practice, a significant reduction in mortality was demonstrated in guideline hospitals. The main effects of guideline implementation were

earlier delivery of enteral feeds and an increase in the total use of enteral feeds and of both enteral or parenteral feeds.

Thus, there appears to be a positive relationship between feed delivery and outcome, yet current feeding practice in many ICUs is suboptimal. In a survey of five ICUs in the UK, Adam and Batson found that only 76% of prescribed feed quantity was delivered, and the prescription varied between 76 and 103% of estimated energy requirement.[12] Similarly, in the USA, McClave found that physicians prescribed a daily mean volume of feed which was 65% of estimated requirements, with only 78% actually delivered.[13] There are several reasons for failing to achieve target requirements with enteral feeding. Problems with gastric motility are probably the most common patient-related reason. However, stopping feeds, often inappropriately, for procedures is clearly of considerable importance. Montejo analysed 3778 feeding days in 400 patients from 37 Spanish ICUs over a single month. Gastrointestinal complications related to feeding occurred in 63% of patients and resulted in withdrawal of feeding by the enteral route in 15%.[14] Of interest, patients with gastrointestinal complications had a lower ratio of volume of feed delivered to that prescribed. They suffered a longer length of stay and a higher mortality when compared with patients with similar APACHE II scores who did not have gastrointestinal complications, again suggesting a link between effective feeding and good outcome.

**Strategies to achieve feeding target goals in the critically ill**

The Canadian guidelines study suggests that implementation of evidence-based practice can improve feeding target compliance.[11] Evidence-based guidelines would include advice on appropriate use of prokinetic agents and the timing and methods to use for passing post-pyloric feeding tubes, which constitute the mainstay of approaches to counter early failure of nasogastric feeding in the critically ill. Interval measurement of gastric residual volume is the usual method of assessing gastroparesis, with high aspirates generally being the trigger for ceasing feeds, prescription of prokinetics or placement of post-pyloric tubes. Whether there is a good relationship between measured gastric residual volume and the presence of gastroparesis is unclear however. McClave found that although critically ill patients were more likely than normal volunteers to have gastric residual volumes of greater than 150 ml eight hours after commencement of nasogastric feeding, a significant number of both groups with low residuals had abnormal gastric emptying on radiology.[15] However, no patient had clinical feed intolerance (i.e. vomiting, aspiration). Thus, the level of gastric residual volume that indicates feed intolerance is not clear.

There is evidence that the use of prokinetic drugs can increase gastric emptying in the critically ill. Boivin has shown that use of erythromycin in nasogastrically fed patients is as effective as transpyloric tube placement in terms of meeting nutritional goals.[16] MacLaren found that both metoclopramide and cisapride, when compared with erythromycin or placebo, resulted in more rapid gastric emptying as assessed by paracetamol absorption.[17] Gastric residual volumes were not different between agents in this study, but in a subsequent study use of metoclopramide was associated with significantly lower gastric residual volumes than was cisapride.[18] Concerns regarding the pro-arrhythmic potential of cisapride and the use of an antibiotic (erythromycin) for a non-antimicrobial indicaton may limit the acceptability of these agents in clinical practice, suggesting that metoclopramide is probably the prokinetic agent of choice.

Where nasogastric feeding, with use of prokinetic agents as indicated, fails, passage

of post-pyloric tubes is recommended. A recent randomised controlled trial found that uncomplicated first-time insertion of a nasojejunal tube by endoscope is possible in 98% of patients.[19] In comparison with nasogastric feeding, nasojejunal feeding was associated with reduced gastric residual volumes and a strong trend to improved tolerance of enteral nutrition. Access to endoscopy services may be difficult to organise in many ICUs, but alternative methods for tube insertion have been described with claimed high success rates.[20]

For a significant number of patients without gastrointestinal failure, even with these manoeuvres, achievement of feeding targets by the enteral route is not possible in the short term. Daily pursuit of enteral feeding targets commits these patients to de facto malnutrition. In many ICUs, the use of parenteral nutrition (PN) is vigorously avoided in patients without gastrointestinal failure, because of the perceived risk of increased complications associated with its use.

**Supplemental parenteral nutrition**

Is there evidence against the use of PN as a supplement to enteral nutrition in these situations?

**Table 1**
Infections and parenteral feeding: Randomised trials comparing parenteral with enteral nutrition in critically ill patients

| Author, year | No. of Subjects | Population | Outcome |
|---|---|---|---|
| Adams 1986[21] | 46 | Trauma laparotomy | No difference in complications |
| Moore 1989[34] | 59 | Major abdo. trauma | ↑ infections & septic morbidity with PN |
| Kudsk 1992[35] | 98 | Major abdo trauma, — blunt & penetrating | ↓ pneumonia, intra-abdo abscesses & line sepsis with EN |
| Borzotta 1994[36] | 48 | Head injury | No difference in infections |
| McClave 1997[37] | 30 | Pancreatitis | ? earlier resolution of stress response with EN |
| Gianotti 1997[38] | 260 | GI cancer; EN vs PN vs immuno-EN | No difference between EN & PN. ↓ infections with immuno-EN |
| Bozzetti 2001[39] | 257 | Upper GI cancer | No difference in nutritional, immunologic & inflammatory goals, complications, length of stay or mortality |

The principal adverse effect attributed to PN is increased rates of infectious complications. This observation is based predominantly on studies comparing enteral or no nutritional support with PN in elective surgical and trauma populations (Table 1). A common failing of these studies is that they do not compare equivalent groups. It is intuitive that, if a patient can be adequately fed enterally, then PN has no role and thus it is inappropriate to randomise to PN patients who can be fed adequately enterally. On the other hand, if patients are "allocated" to PN only if enteral feeding fails, it is inappropriate to compare outcomes between this group and patients who have tolerated enteral feeding. A common characteristic of studies comparing enteral and parenteral feeding is that PN-fed patients receive significantly more calories and nitrogen than enterally fed groups. In many of the classically quoted papers, the total caloric intake of parenterally fed patients reached what today would be considered

hyperalimentation. In a study in which caloric intakes were standardised, no difference between parenteral and enteral feeds could be detected.[21]

There are, however, good theoretical reasons why PN might predispose to increased infective complications. First is the suggestion that lipid emulsions have immune-suppressing properties.[22, 23] Commercial lipid emulsions contain predominately long chain fatty acids of the ω-6 variety. These tend to favour the formation of dienoic prostaglandins, in particular prostaglandin $E_2$ ($PGE_2$). This has a direct suppressant effect on leucocyte function *in vitro*. In the laboratory, fat emulsions have been shown to depress the reticuloendothelial system and pulmonary function. Whether lipid emulsions modify immune responses in vivo in man has not been definitively proven. New lipid formulations are expected in the near future which will have theoretical advantages over traditional formulations in terms of immune modulation.[24] A more important explanation for why use of PN might increase infectious complications is the relationship between use of PN and hyperglycaemia. There is now good evidence that allowing hyperglycaemia to persist in critically ill patients is associated with poor outcome. Van den Berghe and colleagues randomised patients to have blood sugars controlled between 4 and 6 mmol/l or allowed a more liberal target of less than 12 mmol/l.[25] Although the study was complicated by an unusual feeding regimen employing an infusion of 50% glucose on the first day, there was a 50% reduction in ICU mortality in those patients with aggressive control of blood sugar. It remains unclear whether the benefit in this study was due to control of hyperglycaemia or administration of insulin per se; nonetheless, the study suggests that hyperglycaemia should be controlled in the critically ill.

The third major concern with use of PN is promotion of gut mucosal atrophy and consequent increase in gut permeability. Rodents develop villous atrophy within three days of commencement of total PN. While such changes are commonly reported in studies involving small animals, they are poorly characterised in man.[26] Human reports to date have involved long term total PN use and/or total PN administration in children. Increased gut permeability and bacterial translocation have not been convincingly demonstrated to occur in man, and have not been correlated with use of PN. It is unlikely that short term use of PN in critically ill patients without gastro-intestinal failure will have deleterious effects on the gut, especially if given as a supplement to some enteral feed and if, as discussed below, glutamine is included.

The use of supplemental PN in patients with critical illness has been reported by Bauer and colleagues.[27] These authors studied 120 patients from two tertiary ICUs. Patients were randomised to receive either enteral combined with PN or enteral nutrition combined with saline as control. The patients received the same amount of enteral feed in each group and, indeed, the same amount of "theoretical" PN, with an increase in enteral and decrease in PN from days 1 to 7. The total caloric intake was significantly different between the groups, however, with the supplemental group more rapidly and consistently achieving targets. There was no effect on mortality in this small study but, equally, no adverse effect of PN was observed. The supplemental group demonstrated a small but significant reduction in hospital length of stay and improvement in the nutritional markers prealbumin and retinol binding protein.

**Glutamine and immunonutrients**

Glutamine is the principal metabolic fuel for gut mucosal cells. In health, gut mucosa obtains most of its glutamine requirement from luminal contents. In illness,

however, body glutamine requirements appear to be increased. Glutamine is mobilised from skeletal muscle and carried to the splanchnic region where it is utilised in acute-phase protein synthesis, synthesis of anti-oxidants such as glutathione and as a substrate for immune cell replication. In this situation, despite increased splanchnic glutamine supply, gut requirements may not be met without increased dietary intake. Unfortunately, most enteral and parenteral nutrition solutions do not contain glutamine. Therefore, it is probable that supplemental glutamine should be provided to critically ill patients receiving PN, and possible that this is necessary in all critically ill patients. Despite extensive literature on the advantages of glutamine enriched feeds in a number of clinical scenarios, there are few adequate clinical trials in the critically ill. Those published to date are all probably under-powered (Table 2). Nonetheless, from the amassed literature, there does seem to be a signal to suggest that glutamine enriched parenteral feeds are of benefit.[28]

Several other nutrients shown to affect immunological and inflammatory responses have been studied in randomised controlled trials. These include arginine, ω-3 fatty acids and nucleotides. Unfortunately, in vivo studies have to date involved the enteral administration of commercially provided formulations, termed "immunonutrition", and there is no information on the effects of the individual substances. The studies of enteral immunonutrition formulations have recently been the subject of a systematic review.[29] Overall, it appears that immunonutrition is associated with a reduction in infectious complications and this may be associated with reduced hospital length of stay, but no mortality advantage. There is significant heterogeneity between the studies limiting the strength of these results. Of concern, the authors of this review found that those studies with the best methodological quality appeared to show increased mortality with immunonutrition in ICU patients. Pending further large scale studies, it would seem that their conclusion that immunonutrition cannot be recommended

**Table 2**
Studies evaluating glutamine-enriched nutrition in critically ill patients

| Author, year | Population & No subjects | Outcome |
|---|---|---|
| *Parenteral nutrition* | | |
| Tremel, 1994[40] | n=12 | Improved intestinal absorption capacity. |
| Weingartmann 1996[41] | Polytrauma, n=16 | Required high dose to sustain plasma gln levels. |
| Palmer, 1996[42] | n=38, muscle | Failed to influence muscle glutamine biopsies in 16 with five days of feeding |
| Griffiths, 1997[33] | n=84 | Improved survival at six months |
| de Beaux, 1998[43] | Severe pancreatitis, n=14 | Reduced IL-8, trend to improved lymphocyte proliferation |
| Powell-Tuck, 1999[44] | n=168, clinically indicated to receive TPN (not all ICU) | Reduced hospital length of stay & isonitrogenous |
| *Enteral nutrition* | | |
| Jensen, 1996[45] | ICU patients, n=28 | Blunting of hyper-aminoacidarmic response to injury |
| Long, 1996[46] | Trauma, n=30 | Ineffective at maintaining plasma gln level |
| Houdijk, 1998[47] | Trauma, n=72 | Reduction in sepsis, bacteraemia and pneumonia. Soluble TNF receptors reduced. |
| Jones, 1999[48] | ICU patients, n=78 | ? ineffective. Reduced costs using complex analysis. |

for critically ill patients is prudent. A formulation containing eicosapentaenoic acid, γ-linolenic acid and antioxidants has been evaluated in patients with acute respiratory disease syndrome (ARDS).[30] 146 patients were randomly allocated the trial feed or an isonitrogenous, isocaloric control. Use of the trial feed was associated with significant improvements in oxygenation and ventilator variables, which were in turn associated with reduced days of mechanical ventilation and a shorter ICU length of stay. The trial was potentially flawed, however, as the study groups were not comparable in gender distribution and the ventilation management was not standardised. The findings have yet to be replicated by other workers.

**Table 3**
Optimal management of nutrition support in the critically ill

| |
|---|
| Feed early[31] |
| Develop evidence-based guidelines[11] |
| Assess likelihood of successful enteral feeding from day 1:[32] |
| — consider metoclopramide as prokinetic if required[18] |
| — consider post-pyloric tube if nasogastric feeding results in difficulties[19] |
| — any doubt on achieving targets? — add supplemental PN[27] |
| Manage hyperglycaemia aggressively[25] |
| Use glutamine-enriched PN for critically ill patients unable to tolerate EN[33] |
| Assess feeding targets daily |

## Summary

Suggested principles of optimal management of nutrition support of critically ill patients at the start of the new millennium are shown in Table 3. Nutrition research is bedevilled by under powered studies, so unfortunately, few of these principles have been adequately submitted to evaluation by a randomised, controlled trial. Nutritional and metabolic support of the critically ill patient remains in its infancy and there is much work to be done.

## References

1. Marik PE, Zaloga GP. Early enteral nutrition in acutely ill patients: A systematic review. *Crit Care Med* 2001; 29:2264-2270.
2. Moran JL, Peake SL. Early enteral nutrition support improves outcome: hypothesis or fact? *Critical Care and Resuscitation* 2002; 4:15-17.
3. Koretz RL. Nutritional supplementation in the ICU. How critical is nutrition for the critically ill? *Am J Respir Crit Care Med* 1995; 151:570-573.
4. Beresford D. Ten Men Dead: The Story of the 1981 Irish Hunger Strike. (New York: Atlantic Monthly Press) 1987.
5. Kinney JM, Long CL, Duke JH. Carbohydrate and nitrogen metabolism after injury. In: Energy metabolism in trauma. *Ciba found Symp* 1970; 103-126.
6. Ishibashi N, Plank LD, Sando K, Hill G. Optimal protein requirements during the first 2 weeks after the onset of critical illness. *Crit Care Med* 1998; 26:1529-1535.
7. Uehara M, Plank LD, Hill GL. Components of energy expenditure in patients with severe sepsis and major trauma: A basis for clinical care. *Crit Care Med* 1999; 27:1295-1302.
8. Zaloga GP, Roberts P. Permissive underfeeding. *New Horizons* 1994; 2:257-263.
9. Taylor SJ, Fettes SB, Jewkes C, Nelson RJ. Prospective randomised trial to determine the effect of early enhanced enteral nutrition on clinical outcome in mechanically ventilated patients suffering head injury. *Crit Care Med* 1999; 27:2525-2531.
10. Zaloga GP. Immune-enhancing diets: Where's the beef? *Crit Care Med* 1998; 26:1143-1146.
11. Doig CS, Martin CM, Sibbald WJS and the Critical Care Research Network. Implementing evidence-based guidelines for ICU nutrition improves hospital outcome: a cluster randomization trial. *Am J Respir Crit Care Med* 2000; 161:A93.

12. Adam S, Batson S. A study of problems associated with the delivery of enteral feed in critically ill patients in five ICUs in the UK. *Crit Care Med* 1997; 23:261-266.
13. McClave SA, Sexton LK, Spain DA, Adams NA et al. Enteral tube feeding in the intensive care unit: Factors impeding adequate delivery. *Crit Care Med* 1999; 27:1252-1256.
14. Montejo JC. Enteral nutrition-related gastrointestinal complications in critically ill patients: A multicenter study. *Crit Care Med* 1999; 27:1447-1453.
15. McClave SA, Snider HL, Lowen CC, McLaughlin AJ et al. Use of residual volume as a marker for enteral feeding intolerance: prospective blinded comparison with physical examination and radiographic findings. *JPEN J Parenter Enteral Nutr* 1992; 16:99-105.
16. Boivin MA, Levy H. Gastric feeding with erythromycin is equivalent to transpyloric feeding in the critically ill. *Crit Care Med* 2001; 29:1616-1619.
17. McLaren R, Kuhl DA, Gervasio JM, Brown RO et al. Sequential doses of cisapride, erythromycin, and metoclopramide in critically ill patients intolerant to enteral nutrition: a randomised, placebo-controlled, crossover study. *Crit Care Med* 2000; 28:438-444.
18. McLaren R, Patrick WD, Hall RI, Rocker GM et al. Comparison of cisapride and metoclopramide for facilitating gastric emptying and improving intolerance to intragastric enteral nutrition in critically ill, mechanically ventilated adults. *Clin Therapeut* 2001; 23:1855-1866.
19. Davies AR, Froomes PRA, French CJ, Bellomo R et al. Randomized comparison of nasojejunal and nasogastric feeding in critically ill patients. *Crit Care Med* 2002; 30:586-590.
20. Salsadis R, Fleiszer T, Johnston R. Air insufflation technique of enteral tube insertion: a randomized, controlled trial. *Crit Care Med* 1998; 26:1036-1039.
21. Adams S, Dellinger EP, Wertz MJ, Oreskovich MR, Simonowitz D, Johansen K. Enteral versus parenteral nutritional support following laparotomy for trauma: A randomised prospective trial. *J Trauma* 1986; 26:882-891.
22. Pomposelli JJ, Bistrian BR. Is total parenteral nutrition immunosuppressive? *New Horizons* 1994; 2:224-229.
23. Gogos CA, Kalfarentzos F. Total parenteral nutrition and immune system activity: A review. *Nutrition* 1995; 11:339-346.
24. Garnacho-Montero J, Ortiz-Leyba C, Jime'nez-Jime'nez FJ, Garci-Garmendia JL et al. Clinical and metabolic effects of two lipid emulsions on the parenteral nutrition of septic patients. *Nutrition* 2002; 18:134-138.
25. Van den Berghe G, Wouters P, Weekers F, Verwaest C et al. Intensive insulin therapy in critically ill patients. *N Engl J Med* 2001; 345:1359-1367.
26. Jeejeebhoy KN. Total parenteral nutrition: potion or poison? *Am J Clin Nutr* 2001; 73:975-983.
27. Bauer P, Charpentier C, Bouchet C, Nace L, Raffy F, Gaconnet N. Parenteral with enteral nutrition in the critically ill. *Intens Care Med* 2000; 26:293-300.
28. Novak F, Heyland DK, Avenell A, Drover JW, Su X. Glutamine supplementation in serious illness: a systematic review of the evidence. *Crit Care Med* 2002; 30:2022-2029.
29. Heyland DK, Novak F, Drover JW, Jain M, Su X, Suchner U. Should Immunonutrition Become Routine in Critically Ill Patients? A Systematic Review of the Evidence. *JAMA* 2001; 286:944-950.
30. Gadek JE, DeMichele SJ, Karlstad MD, Pacht ER et al. Effect of enteral feeding with eicosapentaenoic acid, (-linoleic acid, ant antoxidants in patients with acute respiratory distress syndrome. *Crit Care Med* 1999; 27:1409-1420.
31. van Haren FMP, van der Hoeven JG. Early enteral nutrition in the Intensive Care Unit. In: Yearbook of Intensive Care and Emergency Medicine 2002, J-L Vincent Ed., Springer-Verlag, Berlin 2002, pp 481-491.
32. Woodcock NP, Ziegler D, Palmer MD, Buckley P, Mitchell CJ, MacFie J. Enteral versus parenteral nutrition: a pragmatic study. *Nutrition* 2001; 17:1-12.
33. Griffiths RD, Jones C, Palmer TE. Six-month outcome of critically ill patients given glutamine-supplemented parenteral nutrition. *Nutrition* 1997; 13:295-302.
34. Moore FA, Moore EE, Jones TN, McCroskey BL, Peterson VM. TEN versus TPN following major abdominal trauma-reduced septic morbidity. *J Trauma* 1989; 29:916-922.
35. Kudsk KA, Croce MA, Fabian TC, Minard G, Tolley EA, Poret HA, Kuhl MR, Brown RO. Enteral versus parenteral feeding. Effects on septic morbidity after blunt and penetrating abdominal trauma. *Ann Surg* 1992; 215:503-511.
36. Borzotta AP, Pennings J, Papasadero B, Paxton J, Mardesic S, Borzotta R, Parrott A, Bledsoe F. Enteral versus parenteral nutrition after severe closed head injury. *J Trauma* 1994; 37:459-468.

37. McClave SA, Greene LM, Snider HL, Makk LJ, Cheadle WG, Owens NA, Dukes LG, Goldsmith LJ. Comparison of the safety of early enteral vs parenteral nutrition in mild acute pancreatitis. *JPEN J Parenter Enteral Nutr* 1997; 21:14-20.
38. Gianotti L, Braga M, Vignali A, Balzano G, Zerbi A, Bisagni P, Di Carlo V. Effect of route of delivery and fomulation of postoperative nutritional support in patients undergoing major operations for malignant neoplasms. *Arch Surg* 1997; 132:1222-1229.
39. Bozetti F, Braga M, Gianotti L, Gavazzi C, Mariani L. Potoperative enteral versus parenteral nutrition in malnourished patients with gastrointestinal cancer: a randomised multicentre trial. *Lancet* 2001 Nov 3; 358(9292):1487-1492.
40. Tremel H, Kienle B, Weilemann LS, Stehle P, Furst P. Glutamine dipeptide-supplemented parenteral nutrition maintains intestinal function in the critically ill. *Gastroenterology* 1994; 107:1595-1601.
41. Weingartmann G, Fridrich P, Mauritz W, Gotzinger P, Mittlbock M, Germann P, Karner J, Roth E. Safety and efficacy of increasing dosages of glycyl-glutamine for total parenteral nutrition in polytrauma patients. *Wien Klin Wochenschr* 1996; 108:683-688.
42. Palmer TE, Griffiths RD, Jones C. The effect of parenteral L-glutamine on muscle in the very severely ill. *Nutrition* 1996; 12:316-320.
43. de Beaux AC, O'Riordain MG, Ross JA, Jodozi L, Carter DC, Fearon KC. Glutamine-supplemented total parenteral nutrition reduces blood mononuclear cell interleukin-8 release in severe acute pancreatitis. *Nutrition* 1998; 14:261-265.
44. Powell-Tuck J. Total parenteral nutrition with glutamine dipeptide shortened hospital stays and improved immune status and nitrogen economy after major abdominal surgery. *Gut* 1999; 45:82-88.
45. Jensen GL, Miller RH, Talabiska DG, Fish J, Gianferante L. A double-blind, prospective, randomized study of glutamine-enriched compared withh standard peptide-based feeding in critically ill patients. *Am J Clin Nutr* 1996; 64:615-621.
46. Long CL, Borghesi L, Stahl R, Clark JA, Geiger JW, DiRienzo DB, Weis JK, Laws HL, Blakemore WS. Impact of enteral feeding of a glutamine-supplemented formula on the hypoaminoacidemic response in trauma patients. *J Trauma* 1996; 40:97-102.
47. Houdijk AP, Rijnsburger ER, Jansen J, Wesdorp RI, Weiss JK, McCamish MA, Teerlink T, Meuwissen SG, Haarman HJ, Thijs LG, van Leeuwen PA. Randomised trial of glutamine-enriched enteral nutrition on infectious morbidity in patients with multiple trauma. *Lancet* 1998; 352:772-776.
48. Jones C, Palmer TE, Griffiths RD. Randomized clinical outcome study of criticallly ill patients given glutamine-supplemented enteral nutrition. *Nutrition* 1999; 15:108-115.

# Percutaneous Tracheostomy — A Review

DR AMOD KARNIK
Intensive Care Unit, Toowoomba Base Hospital

*After graduating in Bombay, Dr Karnik completed his general medical and ICU training in the UK and Australia. While working in the UK, he was involved in the development of translaryngeal tracheostomy and has conducted workshops in the technique in the UK, Europe, India and South Africa, as well as Australia.*

## HISTORY

"A senseless, frivolous invention of Asceplaides" and "A suitable punishment for a sinner in the depths of the inferno" are two of the ways a surgical tracheostomy has been described.[1] This ancient procedure was practiced in India and Egypt around 3000 BC. By the Renaissance, the technique had been refined and its reputation had improved. In the 17th century, Fabricius commented, "This operation redounds to the honour of the physician and places him on a footing with the gods".[1] The modern surgical technique was described by Chevalier Jackson in 1909 and has undergone very few modifications since.[2]

Tracheostomy was the prerogative of surgeons and usually performed in the operating room, until recently, but the introduction of "percutaneous dilatational tracheostomy" has moved it into the realms of anaesthetists and intensivists. This concept was first introduced by Sheldon in 1955[3] and stimulated by Toye and Weinstein in 1969.[4] Their techniques created a passage using instruments combining dilatation and incision of tissues. Although effective, the procedure fell into disrepute following early complications, notably perforation of the posterior tracheal wall. The introduction by Ciaglia in 1985[5] of the technique using multiple dilators made it possible for anaesthetists and intensivists to perform the procedure in the intensive care unit.

In 1990, Schachner[6] described a technique using a device which slid into the trachea over a metal guide wire. When opened, the device resulted in dilatation of the intercartilaginous space allowing placement of a tracheostomy cannula. In the same year, Griggs modified a Howard Kelly forceps, allowing it to be passed over a guide wire and to be used as a blunt dilator to create a tracheostomy stoma.[7] Fantoni, building on his experience in dilatation of tracheal stenosis in children, fashioned a unique tracheostomy tube with an integral dilator, the basis of the translaryngeal tracheostomy technique[8] introduced in 1997. A modified version of the Ciaglia multiple dilator technique was introduced in 1999, the "Ciaglia Blue Rhino". This allows one-step dilation by means of a curved dilator with a hydrophilic coating.

## TECHNIQUES

At the time of writing, there are four commercially available tracheostomy kits: Ciaglia multiple dilators (Cook), Blue Rhino (Cook), Griggs forceps (Portex, Boots) and Fantoni's translaryngeal technique (Tyco Healthcare).

All percutaneous dilatational tracheostomies are performed on anaesthetised, paralysed and ventilated patients after adequate preoxygenation. The thyroid cartilage, cricoid cartilage and suprasternal notch are identified and the tracheostomy is introduced between the first and second, or second and third tracheal rings.

**Ciaglia and Griggs techniques**

The initial steps of these techniques are similar and will be described together. The intended site of the stoma is prepared and draped and lignocaine and adrenaline injected locally. The existing endotracheal tube is withdrawn so that the cuff is positioned above the vocal cords and a seal achieved by hyperinflation. An assistant should hold the endotracheal tube in this position.

*Cook multiple dilator*

The Cook kit consists of a short 11F dilator, an 8F guiding catheter, and six dilators from 12F to 36F.[5] After skin incision and initial dilatation with the 11F short dilator, the guiding catheter is introduced over the wire, maintaining a ridge on the catheter at skin level, thus making a double guide system. Progressively larger dilators (12F to 36F) are passed over this double guide, to dilate the soft tissues and trachea. A tracheostomy tube mounted over an appropriate sized dilator, is then inserted over the guiding catheter, through the stoma and secured.

*Cook "Blue Rhino"*

This kit consists of the 11F short dilator, the 8F guiding catheter, and a single dilator ("Blue Rhino") to dilate up to 38F. This is coated with a hydrophilic coating, which greatly facilitates the introduction of the instrument. Three introducers are also provided to insert an appropriately sized tracheostomy tube. The procedure is the same as with the other Cook kit, with the "Blue Rhino" inserted over the double guide system. It is advanced until the 38F mark is at the skin. Once removed, a tracheostomy tube mounted over an appropriate sized introducer is inserted over the guiding catheter, through the stoma and secured.

*Griggs*

The Portex kit consists of a short 14F dilator, guidewire dilating forceps (modified Howard Kelly clamps) and tracheostomy tube mounted on an obturator.[7] After initial dilatation with the 14F short dilator, the closed dilating forceps are introduced over the guidewire into the subcutaneous tissues, opened and removed. The procedure is repeated until the anterior tracheal wall is reached. The closed forceps are then advanced into the trachea, indicated by a loss of resistance. The handles are raised to align the jaws in the long axis of the trachea, the forceps opened to dilate the anterior tracheal wall and then removed. The tracheostomy tube assembly is then introduced over the guide wire into the trachea. With the guidewire and obturator removed, the tracheostomy tube is secured.

With these three procedures, the routine use of an end-tidal $CO_2$ monitor and intra-procedural bronchoscopy to confirm proper placement of the tracheostomy tube are recommended.[9]

**Fantoni technique**

For this technique, airway control and ventilation are maintained using a small bore ventilation tube with its cuff inflated between the proposed stoma site and the carina.

A curved needle is introduced into the trachea, under bronchoscopic control, and a guide wire is advanced cranially into the oropharynx. A special tracheostomy tube with an integral dilator is then attached to the oral end of the guide wire and traction applied at the neck end. As it is drawn down and out, this tube pierces and dilates the anterior tracheal wall and paratracheal tissues in a retrograde manner. A 2-3 mm incision is necessary to assist it to perforate the skin. The integral dilator is cut off from the rest of the tracheostomy tube, which can then be straightened, rotated and passed caudally to its final position using a special obturator. The tube is secured in place and its position verified with a bronchoscope.

## PERCUTANEOUS v SURGICAL TRACHEOSTOMY

**Table 1**
Reasons for developing percutaneous tracheostomy

| |
|---|
| Complications with surgical technique |
| Avoid transportation of critically ill patients within the hospital |
| Quicker to organise and perform |
| Cost |

There have been a number of reasons for the development of the percutaneous techniques (Table 1), but the benefits of percutaneous over surgical tracheostomy, have not been conclusively proven. To date, there have been nine studies[10-18] and three meta-analyses[19-21] comparing the two techniques. The studies are summarized in Table 2 and the meta-analyses in Table 3. It should be borne in mind that most of the studies

**Table 2**
Studies comparing percutaneous and surgical tracheostomy

| Author | Technique | Procedures (no.) | Operative complications (%) | Bleeding (%) | Post op complications (%) | Infection |
|---|---|---|---|---|---|---|
| Melloni[11] | Ciaglia | 25 | 8 | 8 | 4 | 0 |
| 2002 | Surgical | 25 | 0 | 0 | 36 | 28 |
| Westphal[18] | Ciaglia | 40 | 12.5 | 2.5 | 2.5 | 0 |
| 1999 | Fantoni | 40 | 0 | 0 | 0 | 0 |
| | Surgical | 40 | 12.5 | 2.5 | 35 | 35 |
| Porter[16] | Ciaglia | 12 | 42 | 0 | 0 | 0 |
| 1999 | Surgical | 12 | 8 | 0 | 0 | 0 |
| Holdgaard[15] | Ciaglia | 30 | 63 | 20 | 23 | 10 |
| 1998 | Surgical | 30 | 87 | 87 | 100 | 63 |
| Graham[17] | Ciaglia | 31 | na | na | 61 | na |
| 1996 | Surgical | 29 | na | na | 58 | na |
| Friedman[14] | Ciaglia | 26 | 35 | 13 | 12 | 0 |
| 1996 | Surgical | 27 | 41 | 11 | 41 | 15 |
| Crofts[13] | Ciaglia | 25 | na | na | 25 | 0 |
| 1995 | Surgical | 28 | na | na | 36 | 4 |
| Griggs[12] | Griggs | 153 | na | na | 3.9 | na |
| 1991 | Surgical | 74 | na | na | 18.9 | na |
| Hazard[10] | Ciaglia | 22 | na | na | 12 | 4 |
| 1991 | Surgical | 25 | na | na | 36 | 28 |

**Table 3**
Meta-analyses of studies comparing percutaneous and surgical tracheostomy

| Author | Studies (no.) | Patients (no.) | | Duration (mean mins) | | Peri op complications | | Post op complications | | Superiority percutaneous |
|---|---|---|---|---|---|---|---|---|---|---|
| | | PDT | SCT | PDT | SCT | PDT | SCT | PDT | SCT | |
| Dulguerov[19] | 48 | 1817 | 3512 | 11.7 | 26.9 | 56 | 8.8 | 38 | 27 | Inconclusive |
| Freeman[20] | 5 | 115 | 121 | 9.85 | 22 | 46.6 | 45.3 | 18 | 55.7 | Some advantages |
| Cheng[21] | 4 | 103 | 109 | 8 | 20.9 | 48.2 | 63.1 | 14 | 60 | Safer procedure |

looking at percutaneous tracheostomy excluded patients with distorted anatomy, neck trauma, morbid obesity, difficult airway, or marked coagulopathy. These are accepted as relative contraindications to percutaneous tracheostomy. Surgical tracheostomy is the procedure of choice for these patients.

There is considerable variation in definitions and reporting of complications during the procedure that may partly be responsible for ambiguous results. Quantifying the severity of bleeding also varies amongst the studies. However, percutaneous tracheostomy appears to have a higher incidence of perioperative complications, especially cardiorespiratory events,[9] but postoperative bleeding and stomal infections are less common.[19, 20]

There are few comparative studies of long term complications, including tracheal stenosis. Melloni et al[11] followed up a subgroup of 28 patients for six months with clinical evaluation and videobronchoscopy study at the end of that period. There were no late tracheal complications in the surgical tracheostomy group. Two patients in the percutaneous dilational (PDT) group (n=15) developed long term problems. Tracheomalacia with airway collapse on expiration was observed in a patient who had dyspnea only on exertion. No surgical correction was needed. The second patient failed a trial of decannulation and was shown to have a >50% tracheal stenosis. These authors proposed a classification of endoscopy findings to distinguish between granulomatous and fibrous stenosis, as the treatment of these are different. On the other hand Hazard et al[10] identified seven patients with tracheal stenosis, using a radiological follow up. Five of these had a surgical procedure (n=8), while two had undergone a percutaneous procedure (n=11). With such findings it is unclear whether either technique results in fewer long term complications.

Thus it appears that the speed and ease with which percutaneous tracheostomy can be organised and performed in the ICU, and the elimination of the need to transport critically ill patients out of the ICU environment appear the main reasons for the increasing use and popularity of the non-surgical approach amongst ICU physicians.

## WHICH PERCUTANEOUS TRACHEOSTOMY?

There have been very few studies comparing the various techniques of performing PT. The Ciaglia and Griggs techniques are the most commonly used and have been most frequently compared in the literature. Unfortunately, not all are randomized studies and it is very difficult to have such studies blinded. Bronchoscopy was not a standard in all studies. More importantly, none of the studies was of adequate power

to show a difference in any of the complications, given their relatively low incidence.

Van Heerden et al[2] compared the Ciaglia and Griggs techniques in a serial manner in 54 patients. They used bronchoscopic guidance in the first 15 patients and then only in patients with difficult landmarks. There was no statistically significant difference in the complication rates in the two groups (6/29 for Ciaglia method and 6/25 for Griggs). The first tube change was more difficult in the Ciaglia group. Another study comparing the Griggs and Ciaglia techniques[23] in 80 patients did not show any major complications with either. In a prospective, randomized trial with one hundred patients, Nates et al[24] reported a higher rate of perioperative haemorrhage with the Griggs rather than the Ciaglia technique (25% vs 2%). Anon et al[25] prospectively studied the Ciaglia technique in 25 and the Griggs technique in 38 patients. Neither intraoperative (12% Ciaglia group and 13% in the Griggs group) nor postoperative complications (12% vs 10%) were statistically different. Thus it seems from the limited data that there is little difference in the complication rates of these methods.

A study by Leinhardt et al[26] compared the Rapitrac kit (a kit based on the Schachner system,[6] and no longer available), the Ciaglia method and conventional surgical tracheostomy. This.showed an unacceptably high complication rate in the Rapitrac group (4/5 cases) as compared to Ciaglia (3/20 cases) and conventional surgical tracheostomy (2/16 cases). There were three cases of cuff leak in the Rapitrac group, necessitating an immediate tube change, two episodes of paratracheal insertion of the tube, and one of tube blockage and bilateral pneumothorax. Apart from three cases of minor skin edge bleeding, there were no major complications in the PT group. In the 16 patients who had a conventional tracheostomy, there was one episode of tube occlusion and misplacement and two cases of wound infection.

The Fantoni trans laryngeal technique (TLT) has theoretical advantages in patients with difficult or abnormal anatomy, and coagulopathies.[27] Walz et al[28] presented the first prospective randomized trial comparing it with the Ciaglia method. They found both techniques to be fast and safe in performance. Although there was a rise in the $P_aCO_2$ during the TLT method, it must be emphasized that they used the apnea method rather than ventilating through the microlaryngeal tube provided. Despite this, there were no episodes of hypoxia in either group. Another study compared the Ciaglia, Fantoni and surgical techniques in 120 post cardiac surgical patients.[18] The perioperative complication rate was 12.5% for the Ciaglia and surgical techniques, but there were no complications in the TLT group. A 35% infection rate was noted in the surgical group, with none in the other two groups. In a prospective comparison by Westphal et al, apart from reporting a few technical difficulties, the Fantoni technique was found to be as safe as the Ciaglia and maintained a higher $P_aO_2/FiO_2$ ratio during the procedure.[29]

The "Blue Rhino" is a recent addition to the kits available on the market. A prospective randomized trial of 25 patients in each arm compared the Blue Rhino with the Ciaglia multiple dilator set.[30] The Blue Rhino was quicker to perform (3 v 7 min). There were 11 (44%) complications with it as compared to 7 (28%) with the multiple dilator method. Three complications associated with the latter technique were life threatening (two posterior wall injuries, one pneumothorax).

Neither the Blue Rhino nor the Fantoni methods have been studied as extensively as the Ciaglia. More comparative studies need to be published before any recommendations can be made. A summary of the advantages and limitations of the commonly available techniques is shown in Table 4.

**Table 4**
Advantages and limitations of commonly used techniques for percutaneous tracheostomy

| Technique | Advantages | Limitations |
|---|---|---|
| Ciaglia (Multiple Dilators) | • Seldinger technique<br>• Repetitive steps<br>• Widely practised and taught<br>• Serially larger dilators help tamponade bleeding | • Risk of damage to posterior tracheal wall exists with passage of each dilator<br>• Technique necessitates overdilatation and insertion of a smaller trache tube<br>• Risk of aspiration of blood from stoma site |
| Blue Rhino | • Single step dilation<br>• Risk of posterior wall damage is reduced | • Technique necessitates overdilatation and insertion of a smaller trache tube<br>• A high incidence of tracheal ring fracture has been noted in the literature |
| Griggs | • Single dilator<br>• Exact dilatation to the size of the trache tube<br>• Mandatory bronchoscopy control<br>• Less chances of damage to the tracheal rings due to pulling out of the dilator | • Multiple steps<br>• Looks complicated<br>• Requires reintubation with a microlaryngeal tube<br>• Tube and dilator unable to be reused if trache tube pulled out inadvertently during the procedure |

As experience has accumulated, the complication rate with percutaneous techniques has decreased compared to surgical techniques. However, to show a statistically significant difference in the complication rates for any two different types of percutaneous techniques would require a larger study sample than has been reported so far in the literature. A meta-analysis is needed. A review[31] evaluating 40 series discussing complications in 1684 percutaneous tracheostomy patients looked at four different techniques. The overall complication rates were: Rapitrac 22.9%, Toye (an early technique combining dilatation and incision of tissues) 10.9%, Ciaglia 7.6% and Griggs 1.2%.

The Ciaglia method was the one reported most commonly. A total of 1074 procedures were analyzed. The overall complication rate was 5.5% with the most common complication being peri- and postoperative bleeding (2.1%). A 1% rate of long term tracheal stenosis was reported. There were 248 Griggs procedures reviewed, with the original investigators providing most of the data. There were very few complications reported. A 2% incidence of bleeding was seen, which is similar to the Ciaglia method.

## ADJUNCTS

### Bronchoscopy

During percutaneous dilatational tracheostomy (PDT), bronchoscopy has only been used as a teaching aid during the early part of the learning curve in many recent series,[32] or not at all.[14, 33-37] However, the accuracy of blind percutaneous puncture of the trachea has been doubted[38] and routine bronchoscopy during the procedure would allow precise placement of the stoma, demonstrate perforation of the posterior

tracheal wall and prevent dilatation or tearing of the posterior wall. It would also allow easy reintubation in case of a complication or loss of airway. There is growing evidence to support routine bronchoscopic control while performing dilatational tracheostomy.[19, 39-42]

**LMA**

The laryngeal mask airway has been used during these procedures. It does not protect against aspiration, the struts need to be cut off to allow introduction of a bronchoscope, and ventilating patients with poorly compliant lungs may pose a problem. There are varied reports of its usefulness in the literature[43, 44] and, at present, use of an LMA in preference to an endotracheal tube cannot be recommended.

**Ultrasound**

Ultrasound imaging of the neck has been used to delineate the anatomy of the thyroid isthmus and delineate aberrant vessels, identify a potential stoma site and estimate the distance from the surface of the skin to the trachea.[45] In a prospective study of trauma patients, the intended stoma site was changed in 24% following an ultrasound study.[46] It does appear to be a useful adjunct to these procedure, but needs further investigation to prove any potential benefits.

## LONG TERM COMPLICATIONS

Table 5 lists some of the persistent symptoms found during three years follow up of patients who have undergone a percutaneous tracheostomy in a teaching hospital.

**Table 5**
Complications post-decannulation over 3 years

| |
|---|
| Noisy breathing |
| Difficulty in breathing |
| Change in voice quality |
| Persistent cough |
| Swallowing problems |
| Unsightly scar |

Clinically significant tracheal stenosis remains the most dreaded long-term complication. The risk of this is higher with fracture of the cricoid cartilage or tracheal rings, or posterior displacement of the anterior tracheal wall.[11, 47] These injuries can result from sub-cricoid puncture, oblique puncture of the trachea and poor fit of the tracheostomy tube over the appropriate dilator, requiring more force to push the combination into the tracheal lumen. Some of these problems can be avoided if the procedure is done under bronchoscopy guidance.

Tracheal stenosis usually occurs 2-3 months after removal of the device.[48] It may be asymptomatic until the tracheal lumen is reduced by 75%.[49] Stridor does not occur until tracheal diameter is 5 mm or less.[50] A combination of symptoms (stridor, hoarse voice, difficulty breathing) and either evidence of inspiratory flow limitation or anatomical evidence of tracheal narrowing are necessary to diagnose clinically significant tracheal stenosis. Unfortunately spirometry has not been very useful in the diagnosis. Law and colleagues could find no correlation between symptoms and spirometry evidence of fixed airway obstruction.[49] A number of patients could not

perform spirometry. There is no documented correlation between symptoms and flow limitation on spirometry,[34, 49, 51-54] in any study of tracheal stenosis.

The incidence of tracheal stenosis following percutaneous tracheostomy is stated as 0-5%, depending on the diagnostic modality used and the degree of narrowing considered as stenosis. The possibility that tracheal injury may have been caused by trans-laryngeal intubation is a confounding factor.[55] The number of patients available for long term follow up is limited by mortality and morbidity following discharge from ICU. Hill and colleagues[34] reviewed eight patients who developed airway obstruction following removal of a tracheostomy tube. Four had stenosis at the level of the stoma site, one had stenosis at the level of the cricoid cartilage, one had excessive granulation at the stoma site and another had a subglottic web. Law et al[49] studied 41 patients who had a percutaneous tracheostomy and were decannulated. They considered >10% narrowing of the tracheal lumen as significant and four patients fell into this category, but there was no correlation between symptoms and tracheal narrowing.

There really are no established criteria for the diagnosis of clinically significant tracheal stenosis and all the present diagnostic evaluations aim at a structural definition. However, the diagnosis should be clinical and only those who are clinically symptomatic should be investigated further.

## EARLY TRACHEOSTOMY

Recently, there has been a trend to performing tracheostomy early (<6 days) in the ICU stay. A number of studies have shown benefits from this, by reduction in pulmonary septic complications[56] and rapid weaning from ventilatory support.[57] There is also a demonstrable reduction in use of resources with shorter length of stay and lower hospital costs, without an increase in morbidity or mortality in particular patient groups.[58, 59] Given the ease with which the procedure can be done in the ICU, such evidence will encourage an increase in the number of percutaneous tracheostomies being performed.

## ANAESTHETIC IMPLICATIONS

Managing the airway during a percutaneous tracheostomy is a very important and often difficult task.[60] The standard teaching is to position an endotracheal tube with the cuff just below the vocal cords. There is a high incidence of cuff rupture with this approach. There have also been reports of attempts at dilatation of a Murphy's eye. The alternative is to position the cuff above the cords, which is then hyperinflated to achieve a seal against the aryepiglottic folds and the interarytenoid fissure. This presents a risk of aspiration and an assistant is needed to stabilize the tube, especially if the anesthetist is also guiding the procedure using a bronchoscope. Obviously, the risk of cuff rupture or cannulation of the Murphy's eye is much reduced. The use of a microlaryngeal tube with its cuff inflated beyond the stoma site to protect the airway and maintain ventilation is an integral part of the Fantoni technique.[8] This method of airway control has been reported with other dilatational techniques and seems to work well.[61, 62]

## COST

Percutaneous tracheostomy has been proven repeatedly to be cheaper than surgical tracheostomy performed in the operating theatre.[18, 34, 36, 63, 64] However, if the surgical

tracheostomy is performed in the ICU, then it may offer greater savings than a percutaneous technique.[65]

## SUMMARY

There is a definite trend towards performing more percutaneous dilatational tracheostomies in the ICU. These have been shown to be quicker to organise and perform, as well as cheaper than surgical tracheostomy. However, it is still uncertain if they are safer and have fewer complications than surgical tracheostomy. Amongst the commercially available techniques, the Ciaglia is the one used most commonly, but more comparative studies are needed to prove the superiority of any one system. The use of a bronchoscope to guide the procedure should be a standard of care.

## REFERENCES

1. Sercer A. Tracheostomy through two thousand years of history. *Ciba Found Symp* 1962; 10:78-86.
2. Jackson C. Tracheotomy. *Laryngoscope* 1909; 19:285-290.
3. Sheldon ch PR, Freshwater DB. A new method for tracheostomy. *J Neurosurgery* 1955; 12:428.
4. Toye FJ, Weinstein J. A percutaneous tracheostomy device. *Surgery* 1969; 65:384.
5. Ciaglia P, Firsching R, Syniec C. Elective percutaneous dilatational tracheostomy. A new simple bedside procedure; preliminary report. *Chest* 1985; 87:715-719.
6. Schachner A, Ovil Y, Sidi J, Rogev M, Heilbronn Y, Levy MJ. Percutaneous tracheostomy — a new method. *Crit Care Med* 1989; 17:1052-1056.
7. Griggs WM WL, Gilligan JE, Thomas PD, Myburg J.A. A simple percutaneous tracheostomy technique. *Gynaecology and Obstetrics* 1990; 170:543-545.
8. Fantoni ADR. A non-derivative, non-surgical tracheostomy: the translaryngeal method. *Intensive Care Medicine* 1997; 23.
9. Marx WH, Ciaglia P, Graniero KD. Some important details in the technique of percutaneous dilatational tracheostomy via the modified seldinger technique. *Chest* 1996; 110:762-766.
10. Hazard P, Jones C, Benitone J. Comparative clinical trial of standard operative tracheostomy with percutaneous tracheostomy. *Crit Care Med* 1991; 19:1018-1024.
11. Melloni G, Muttini S, Gallioli G, Carretta A, Cozzi S, Gemma M et al. Surgical tracheostomy versus percutaneous dilatational tracheostomy. A prospective-randomized study with long-term follow-up. *J Cardiovasc Surg (Torino)* 2002; 43:113-121.
12. Griggs WM, Myburgh JA, Worthley LI. A prospective comparison of a percutaneous tracheostomy technique with standard surgical tracheostomy. *Intens Care Med* 1991; 17:261-263.
13. Crofts SL, Alzeer A, McGuire GP, Wong DT, Charles D. A comparison of percutaneous and operative tracheostomies in intensive care patients. *Canadian Journal of Anaesthesia* 1995; 42(9):775-779.
14. Friedman Y, Fildes J, Mizock B, Samuel J, Patel S, Appavu S et al. Comparison of percutaneous and surgical tracheostomies. *Chest* 1996; 110:480-485.
15. Holdgaard HO, Pedersen J, Jensen RH et al. Percutaneous dilatational tracheostomy versus conventional surgical tracheostomy. A clinical randomised study. *Acta Anaesthesiol Scand* 1998; 42:545-550.
16. Porter JM, Ivatury RR. Preferred route of tracheostomy — percutaneous versus open at the bedside: a randomized, prospective study in the surgical intensive care unit. *Am Surg* 1999; 65:142-146.
17. Graham JS, Mulloy RH, Sutherland FR, Rose S. Percutaneous versus open tracheostomy: a retrospective cohort outcome study. *J Trauma* 1996; 41:245-248; discussion 248-250.
18. Westphal K, Byhahn C, Rinne T, Wilke HJ, Wimmer-Greinecker G, Lischke V. Tracheostomy in cardiosurgical patients: surgical tracheostomy versus ciaglia and fantoni methods. *Ann Thorac Surg* 1999; 68:486-492.
19. Dulguerov P, Gysin C, Perneger TV, Chevrolet JC. Percutaneous or surgical tracheostomy: a meta-analysis. *Crit Care Med* 1999; 27:1617-1625.
20. Freeman BD, Isabella K, Lin N, Buchman TG. A meta-analysis of prospective trials comparing percutaneous and surgical tracheostomy in critically ill patients. *Chest* 2000; 118:1412-1418.
21. Cheng E, Fee WE Jr. Dilatational versus standard tracheostomy: a meta-analysis. *Ann Otol Rhinol Laryngol* 2000; 109:803-807.
22. Van Heerden PV, Webb SA, Power BM, Thompson WR. Percutaneous dilational tracheostomy — a clinical study evaluating two systems. *Anaes Intens Care* 1996; 24:56-59.

23. Ambesh SP, Kaushik S. Percutaneous dilational tracheostomy: the Ciaglia method versus the Portex [correction of Rapitrach] method. *Anesth Analg* 1998; 87:556-561.
24. Nates NL, Cooper DJ, Myles PS, Scheinkestel CD, Tuxen DV. Percutaneous tracheostomy in critically ill patients: a prospective, randomized comparison of two techniques. *Crit Care Med* 2000; 28:3734-3739.
25. Anon JM, Gomez V, Escuela MP et al. Percutaneous tracheostomy: comparison of Ciaglia and Griggs techniques. *Crit Care* 2000; 4:124-128.
26. Leinhardt DJ, Mughal M, Bowles B et al. Appraisal of percutaneous tracheostomy. *British Journal of Surgery* 1992; 79:255-258.
27. Byhahn C, Lischke V, Westphal K. Translaryngeal tracheostomy in highly unstable patients. *Anaesthesia* 2000; 55:678-682.
28. Walz MK, Peitgen K, Nimtz K, Eigler FW. Percutaneous dilatational tracheostomy VS. Translaryngeal Tracheostomy for critically ill patients — A prospective randomized trial. *Intensive Care Medicine* 1997; 23 Supp1:S42.
29. Westphal K, Byhahn C, Wilke HJ, Lischke V. Percutaneous tracheostomy: a clinical comparison of dilatational (Ciaglia) and translaryngeal (Fantoni) techniques. *Anesth Analg* 1999; 89:938-943.
30. Byhahn C, Wilke HJ, Halbig S, Lischke V, Westphal K. Percutaneous tracheostomy: ciaglia blue rhino versus the basic ciaglia technique of percutaneous dilational tracheostomy. *Anesth Analg* 2000; 91:882-886.
31. Powell DM, Price PD, Forrest LA. Review of percutaneous tracheostomy. *Laryngoscope* 1998; 108:170-177.
32. Petros S EL. Percutaneous dilatational tracheostomy in a medical ICU. *Intens Care Med* 1997; 23:630-634.
33. Cakar N TAS, Esen F, Telci L, Denkel T, Kesecioglu J. Percutaneous dilatational tracheostomy: safety and ease of performance at the bedside in the ICU. *Clin Intens Care* 1997; 4-9.
34. Hill BB, Zweng TN, Maley RH, Charash WE, Toursarkissian B, Kearney PA. Percutaneous dilational tracheostomy: report of 356 cases. *J Trauma* 1996; 41:238-243; discussion 243-244.
35. Chendrashekhar A., Ponnapalli SAD. Percutaneous Dilatational Tracheostomy: an alternative approach to surgical tracheostomy. *South Med J* 1995; 88:1062-1064.
36. Cobean R, Beals M, Moss C, Bredenberg CE. Percutaneous dilatational tracheostomy. A safe, cost-effective bedside procedure. *Archives of Surgery* 1996; 131:265-271.
37. Griggs WM, Myburgh JA, LIG W. A prospective comparison of a percutaneous tracheostomy technique with standard surgical tracheostomy. *Intensive Care Medicine* 1991; 17:261-263.
38. Dexter TJ. A cadaver study appraising accuracy of blind placement of percutaneous tracheostomy. *Anaesthesia* 1995; 50:863-864.
39. inkler WB KR, Seelman O, Havlicek J, Slany J. Bedside percutaneous dilational tracheostomy with endoscopic guidance: experience with 71 ICU patients. *Intens Care Med* 1994; 20:476-479.
40. Fernandez L, Norwood S, Roettger R, Gass D, Wilkins Hr. Bedside percutaneous tracheostomy with bronchoscopic guidance in critically ill patients. *Arch Surg* 1996; 131:129-132.
41. Barba CA, Angood PB, Kauder DR, Latenser B, Martin K, McGonigal MD et al. Bronchoscopic guidance makes percutaneous tracheostomy a safe, cost-effective, and easy-to-teach procedure. *Surgery* 1995; 118:879-883.
42. Hazard PB. Further refinement of percutaneous tracheostomy technique. *Intens Care Med* 1994; 20:466-467.
43. Dosemeci L, Yilmaz M, Gurpinar F, Ramazanoglu A. The use of the laryngeal mask airway as an alternative to the endotracheal tube during percutaneous dilatational tracheostomy. *Intens Care Med* 2002; 28:63-67.
44. Ambesh SP, Sinha PK, Tripathi M, Matreja P. Laryngeal mask airway vs endotracheal tube to facilitate bedside percutaneous tracheostomy in critically ill patients: a prospective comparative study. *J Postgrad Med* 2002; 48:11-15.
45. Muhammad JK, Patton DW, Evans RM, Major E. Percutaneous dilatational tracheostomy under ultrasound guidance. *Br J Oral Maxillofac Surg* 1999; 37:309-311.
46. Kollig E, Heydenreich U, Roetman B, Hopf F, Muhr G. Ultrasound and bronchoscopic controlled percutaneous tracheostomy on trauma ICU. *Injury* 2000; 31:663-668.
47. van Heurn LW, Theunissen PH, Ramsay G, Brink PR. Pathologic changes of the trachea after percutaneous dilatational tracheotomy. *Chest* 1996; 109:1466-1469.
48. Heffner JE, Miller KS, Sahn SA. Tracheostomy in the intensive care unit. Part 2: Complications. *Chest* 1986; 90:430-436.

49. Law RC, Carney AS, Manara AR. Long-term outcome after percutaneous dilational tracheostomy. Endoscopic and spirometry findings. *Anaesthesia* 1997; 52:51-56.
50. Weber AL, Grillo HC. Tracheal stenosis: an analysis of 151 cases. *Radiol Clin North Am* 1978; 16:291-308.
51. Hazard PB GHJ, Adams JW. Bedside percutaneous tracheostomy: Experience with 55 elective procedures. *Ann Thorac Surg* 1988; 46:63.
52. Ciaglia P GKD. Percutaneous Dilatational Tracheostomy:Results and Long-term Follow-up. *Chest* 1992; 101:464-467.
53. Friedman Y MA. Bedside percutaneous tracheostomy in critically ill patients. *Chest* 1993; 104:532.
54. Fischler MP, Kuhn M, Cantieni R, Frutiger A. Late outcome of percutaneous dilatational tracheostomy in intensive care patients. *Intens Care Med* 1995; 21:475-481.
55. Susanto I. Comparing percutaneous tracheostomy with open surgical tracheostomy. *Brit Med J* 2002; 324:3-4.
56. Kluger Y, Paul DB, Lucke J et al. Early tracheostomy in trauma patients. *Eur J Emerg Med* 1996;3:95-101.
57. Teoh WH, Goh KY, Chan C. The role of early tracheostomy in critically ill neurosurgical patients. *Ann Acad Med Singapore* 2001; 30:234-238.
58. Armstrong PA, McCarthy MC, Peoples JB. Reduced use of resources by early tracheostomy in ventilator-dependent patients with blunt trauma. *Surgery* 1998; 124:763-766; discussion 766-767.
59. Brook AD, Sherman G, Malen J, Kollef MH. Early versus late tracheostomy in patients who require prolonged mechanical ventilation. *Am J Crit Care* 2000; 9:352-359.
60. Schwann NM. Percutaneous dilational tracheostomy: anesthetic considerations for a growing trend. *Anesth Analg* 1997; 84:907-911.
61. Fisher L, Duane D, Lafreniere L, Read D. Percutaneous dilational tracheostomy: a safer technique of airway management using a microlaryngeal tube. *Anaesthesia* 2002; 57:253-255.
62. Sun KO. Use of microlaryngeal tube during percutaneous dilatational tracheostomy. *Anaesthesia* 1995; 50:660.
63. Marelli D PA, Manolidis S et al. Endoscopic guided percutaneous tracheostomy: Early results of a consecutive trial. *J Trauma* 1990; 30:433.
64. Freeman BD, Isabella K, Cobb JP et al. A prospective, randomized study comparing percutaneous with surgical tracheostomy in critically ill patients. *Crit Care Med* 2001; 29:926-930.
65. Wease GL, Frikker M, Villalba M, Glover J. Bedside tracheostomy in the intensive care unit. *Arch Surg* 1996; 131:55255-4; discussion 554-555.

# Stem Cells: Properties and Therapeutic Potential

REBECCA A. KEOUGH, BSc, PhD

JOY RATHJEN, BS, D.Phil

PETER D. RATHJEN, BSc, DPhil
School of Molecular and Biomedical Sciences, and National Stem Cell Centre, and ARC SRC in Molecular Genetics of Development, University of Adelaide, South Australia 5005, Australia

*Rebecca Keough is a Research Fellow and Joy Rathjen is a Senior Research Fellow in the School of Molecular and Biomedical Sciences at the University of Adelaide. Peter Rathjen is Executive Dean of the Faculty of Sciences at the University of Adelaide. This article is based on a paper presented at the 2002 Annual Scientific Meeting of the Australian Society of Anaesthetists in Adelaide, where it aroused a great deal of interest.*

The potential therapeutic applications of stem cells and the ethical issues surrounding the generation of human embryonic stem (ES) cell lines from human embryos have been a focus of recent media attention and political scrutiny. Proponents of this technology claim that human ES cells may revolutionise the fields of transplantation and regenerative medicine, promising cures for many of the major diseases facing Western society including diabetes, cardiovascular disease, neurodegenerative disease and cancer. However, the embryo destruction associated with derivation of ES cells from human embryos raises a number of ethical concerns and other stem cell sources, such as adult stem cells, are being championed as useful alternatives. This review examines these different types of stem cells and discusses their potential applications and issues that need to be addressed in order to develop products suitable for clinical application.

**Cell Therapy and Regenerative Medicine**

Many serious human diseases are caused by cell loss, damage or dysfunction. In its simplest form, cell therapy envisages treatment of such diseases by transplantation or replacement of the deficient population so as to restore normal function. Cell therapy has been in clinical use for several decades, most notably and successfully in the form of allogeneic bone marrow transplantation used to treat a variety of leukaemias and other haematological disorders. In broadening the application of cell therapy to treat other diseases, two main approaches are proposed: cellular therapy and tissue engineering. In simple cases, appropriate cell types will be transplanted, either directly at the site of damage or in a manner such that the cells will home to the appropriate sites (i.e. cellular therapy). Alternatively, the necessary cell types will be seeded and grown on a supportive bio-scaffold to produce a three-dimensional structure that is functionally and structurally able to replace damaged tissue or organs (i.e. tissue engineering). While the latter appears to be a significantly more complex undertaking,

preliminary data provide proof of concept for the treatment of monocellular deficiency diseases by cell replacement therapies. For example, transplantation of foetal tissue containing dopaminergic neurons has been shown to at least partially ameliorate symptoms in Parkinson's sufferers. Transplantation of pancreatic islet cells from cadavers can restore insulin levels in diabetes patients (reviewed in [1]). A wide range of additional potential targets has been identified including stroke, spinal cord injuries, multiple sclerosis, motor neuron disease, macular degeneration, liver and muscle regeneration and restoration of immune function.

The basic requirement for cell therapy is a supply of safe, highly pure, defined cell populations in adequate numbers for transplantation. A major limitation in the advance of such therapies has been problems associated with purification of sufficient numbers of the appropriate cell types from donors. This review describes the potential for production of therapeutically relevant cell types by differentiation of stem cells as a promising strategy for the realisation of cell therapy (see Figure 1).

**What is a stem cell?**

Stem cells are primitive, undifferentiated cell populations with the unique capacity to choose between two cell fates, self-renewal or differentiation, an essential requirement for maintaining stem cell pools throughout life. Stem cells can be isolated from numerous sources, notably the founder populations of the embryo (ES cells) or from later foetal and adult stem cell pools that reside within somatic tissues (adult stem cells). Adult stem cell populations have been identified in numerous organs, including skin and other epithelia, liver, pancreas, muscle, brain and blood.

Adult and embryonic stem cells appear to differ in their differentiation potentials. ES cells are isolated from the founder cell population of the embryo prior to its commitment to form any particular tissue lineage, and therefore retain the potential to differentiate into any of the tissues of the body. This property is described as pluripotence. In contrast, adult stem cells are generally thought to have a more restricted, multipotent differentiation potential, with the capacity to differentiate into the various cell lineages required to repair the tissues in which they reside in response to minor tissue damage. Alternatively, they can replenish tissues in which there is a constant turnover of cells, such as skin, intestinal epithelium and blood. Recently, however, the idea that adult stem cells have a limited developmental repertoire has been challenged, with several reports of adult stem cells derived from one tissue apparently able to differentiate into cells of another tissue type.[2] In combination, the capacity for unlimited expansion of the stem cells and widespread differentiation potential provide a potentially unlimited resource for the generation of cell types suitable for therapeutic transplantation.

**Embryonic stem (ES) cells**

ES cells (Figure 1A) are derived from the inner cell mass (ICM) founder cells of the pre-implantation mammalian embryo, or blastocyst, typically at 4.5 days post coitum (d.p.c.) in the mouse and 4-6 days in the human. During development, this small population of cells gives rise to all the tissues of the embryo and adult, including the germ lineage, and contributes to several of the extraembryonic tissues required for embryo support. ES cells were first isolated from mouse blastocysts more than twenty years ago[3,4], but it is only more recently that similar cell lines have been generated from primates, including humans[5-7].

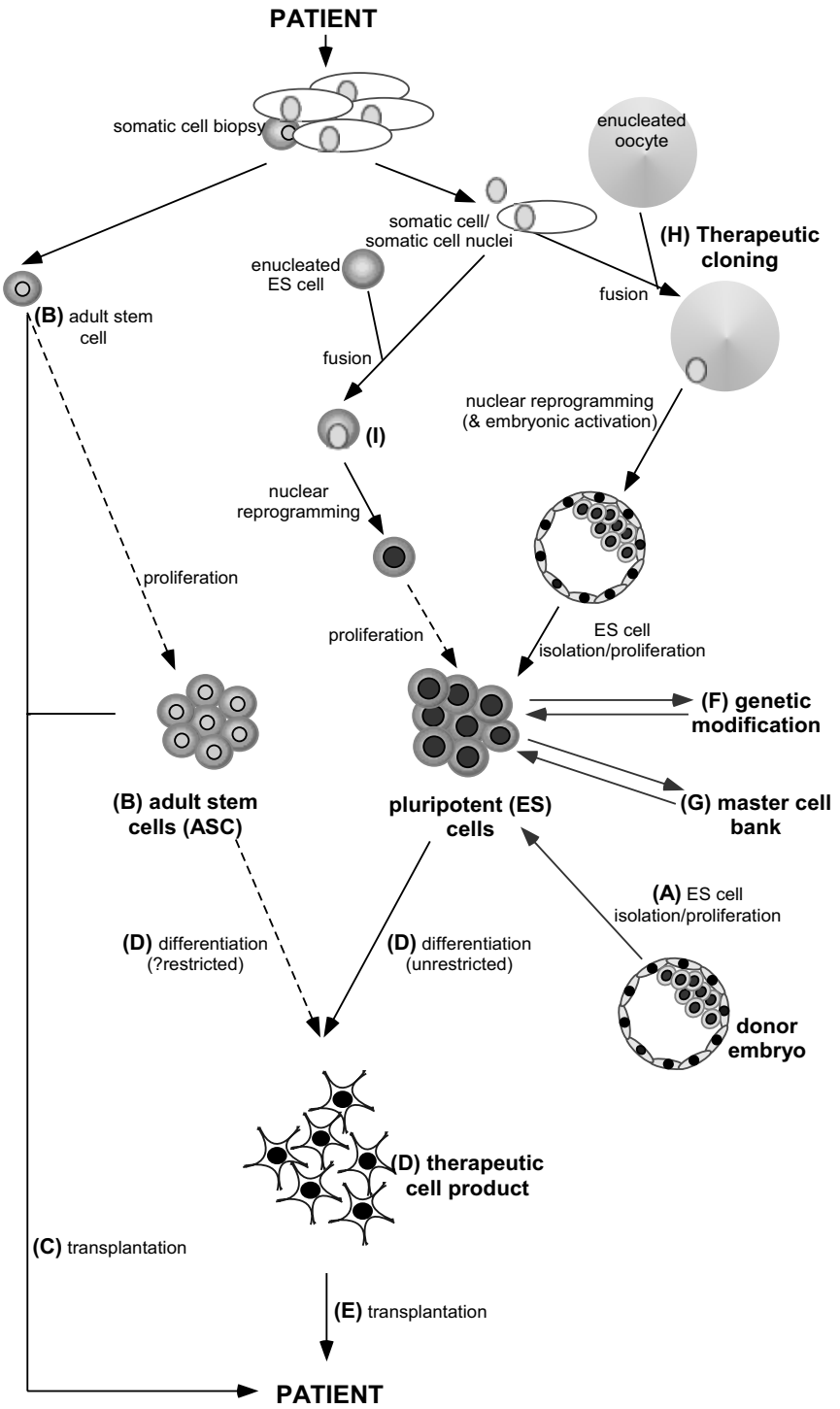


**Figure 1.** A schema summarising the multiple approaches to the application of embryonic and adult stem cell populations to stem cell therapies. Detailed explanations of individual processes (A-I) can be found within the text.

A remarkable property of ES cells is that, given the appropriate culture conditions, they can proliferate indefinitely while maintaining a normal karyotype and retaining pluripotency.[8, 9] It is therefore possible to generate an essentially unlimited supply of precursor cells for later differentiation into therapeutically relevant lineages. Maintenance of undifferentiated mouse ES cells in culture requires an exogenous source of IL6-family cytokines[10, 11] while human ES cells require fibroblast growth factor-2 (FGF-2) or growth on a feeder layer of mitotically inactivated fibroblasts[6, 7, 9] Removal of these growth factors, or addition of differentiation-inducing factors, results in a loss of pluripotence and differentiation into numerous cell types, including cardiac muscle, neurons, epithelial cells, haematopoietic cells and adipocytes.[7 12]

Definitive evidence that mouse ES cells are pluripotent, with a developmental potential equivalent to the ICM, has been demonstrated by the formation of chimaeric mice after reintroduction into host blastocysts. ES cells have been shown to contribute differentiated progeny to all tissues of the chimaeric mice, including the germ cells.[13] Direct injection of human or mouse ES cells into ectopic sites in rodents results in the formation of teratomas, benign tumours comprised of multiple differentiated cell types, indicative of the pluripotentiality of the injected stem cells.

A closely related, embryo-derived, pluripotent stem cell population, termed embryonic germ (EG) cells can be sourced from primordial germ cells in the genital ridge of the developing, post-implantation foetus, typically at 5-9 weeks gestation in the human[14] and 8.5-12.5 d.p.c. in mice.[15] EG cell lines share many properties of ES cells in culture, including pluripotence, but have received less attention as an experimental or therapeutic substrate.

*Human ES cells*

A number of human ES cell lines have now been established (see http://stemcells.nih.gov/registry), generally from excess embryos generated during in vitro fertilisation (IVF) procedures to treat infertility. Embryos are grown in vitro to the blastocyst stage and the ICMs are isolated and cultured on mitotically inactivated feeder cell layers. These cells are propagated as flat, compact colonies that are selected and expanded to produce ES cell lines.

Despite differences between human and mouse ES cell lines with respect to their morphology, growth rate, propagation and culture requirements, they are broadly similar in their gene expression profiles and differentiation potential. In vitro, both human and mouse ES cells can be differentiated into populations equivalent to the three germ layers of the embryo. Further differentiation results in the production of virtually all of the somatic cell types of the body[16-18] and human ES cells have been shown to respond to numerous growth factors resulting in tissue-specific gene activation.[19] However, unlike mouse ES cells, clonal propagation of human ES cells as homogeneous, undifferentiated, pluripotent populations is currently not possible. Despite this, their unique attributes indicate that human ES cells could potentially provide an unlimited supply of differentiated, specialised cells for transplantation.

**Adult stem cells**

Multipotent adult stem cells (Figure 1B) are found in many tissues and, unlike ES cells, these are generally slow cycling cells, residing in specific niches which provide the correct microenvironment to support self-renewal and appropriate differentiation upon receipt of particular environmental cues, such as wound healing or tissue

regeneration. The ability of adult stem cells to repair and regenerate the specific tissues in which they reside is well established and has been exploited therapeutically, most notably in the form of bone marrow and cord blood transplants using haematopoietic stem cells (HSC; Figure 1C).

Current limitations associated with the therapeutic use of adult stem cells include difficulties associated with identifying and purifying these cells from source material and, more particularly, the ability to expand these populations in vitro to produce sufficient quantities of cells for transplantation. Unlike ES cells, most adult stem cell types appear to have a limited capacity for self-renewal when cultured in vitro.[20] The present research effort dedicated to discovering methods for ex vivo expansion of the well characterised and studied HSC population is testament to the difficulties facing adult stem cell biologists.

Recent reports detailing the contribution of transplanted bone marrow cells to non-haematopoietic tissues, such as liver, epithelia and other tissues[21, 22] suggested that adult stem cells maintain (or regain) pluripotential rather than multipotential developmental capacity, a plasticity that was previously thought unlikely. Numerous, similar reports of transdifferentiation by other stem cells, including neural, muscle and mesenchymal stem cells, have generated both excitement and controversy in the stem cell field.[22]

The controversy surrounding adult stem cell plasticity results from difficulties associated with demonstrating conclusively that plasticity or transdifferentiation is the only explanation for the experimental observations. In particular, transplantation of heterogeneous populations that may include other stem cell types, or fusion of transplanted cells with recipient cells of a different lineage cannot yet be eliminated.[2] However, if adult stem cells are capable of transdifferentiation, the potential exists to produce virtually any cell type from a readily accessible source of adult stem cells, thereby markedly increasing their therapeutic value.

Of particular interest are so-called multipotential adult progenitor cells (MAPCs), a subset of mesenchymal stem cells from bone marrow, which can be differentiated in vitro (therefore, without cell fusion) into cells representing progeny of the three germ layers, including hepatocytes,[23] endothelial cells[24] and neurons.[25] Surprisingly, murine MAPCs can also contribute to a wide range of tissues in chimaeric mice produced from injection of MAPCs into mouse blastocysts,[25] suggesting that these cells may represent a pluripotential adult stem cell population. MAPCs were identified after prolonged culturing and it is unclear, therefore, whether they represent a cell population that is normally present in the bone marrow, or whether they are an artefact, generated by "de-differentiation" or "reprogramming" as a result of culturing. However, in terms of clinical exploitation, issues of whether MAPCs exist in vivo or whether adult stem cells are truly plastic in their potential may be of little importance if such cells can be used to generate the desired differentiated progeny which can engraft and function as required.

**Differentiation of stem cells**

Transplantation of adult stem cells to achieve therapeutic outcomes may not require prior differentiation in vitro (Figure 1C). The site of transplantation or homing to specific niches may provide the necessary cues for subsequent differentiation into the required cell types, and studies to date suggest that there is little risk of tumour formation. It is premature to say whether transplantation of potentially pluripotent

adult stem cells, such as MAPCs, require differentiation in vitro before therapeutic transplantation (Figure 1D, E). Initial studies involving direct transfer of undifferentiated MAPCs into mice resulted in engraftment and contribution to numerous tissues without tumour formation.

In contrast, ES-based cellular therapies will require pure, differentiated cell populations for transplantation in order to avoid inappropriate cell types, including undifferentiated, tumorigenic cells, being delivered to the graft site (Figure 1D). Recent work suggests that transplantation of more differentiated cell types decreases the risk of tumour formation.[26-28]

Protocols for the differentiation of ES cells generally involve the formation of embryoid bodies (EBs).[16] These cell aggregates, in which differentiation is heterogeneous and uncontrolled, recapitulate many of the processes of early embryonic development, including the formation of extraembryonic lineages and differentiation into cell types representative of all three germ layers of the embryo. However, EBs are largely disorganised, lacking many of the positional cues that specify body axes in the embryo. Accordingly, the differentiated cell populations resulting from EB differentiation are extremely heterogeneous, with individual, potentially useful cell types often present at low levels. Further, concerns have been raised[16] that inappropriate cell-cell interactions and exposure to inappropriate cell signalling within the disorganised environment may compromise the function of differentiated stem cell progeny.

Purification of differentiated cells from EBs has been used to produce populations enriched for specific cell types. Standard methodologies for purification include fluorescence-activated cell sorting (FACS) or magnetic-activated cell separation (MACS) techniques to isolate cells expressing specific cell surface antigens. Genetic modification has also been used to facilitate selection or purification of the cells of interest from a mixed population.[29, 30] Typically, this has been achieved by engineering ES cells to express selectable marker genes under the control of lineage-specific promoters coupled with selective culture conditions that favour the growth of particular lineages or cell purification. However, the application of these methodologies is limited to occasional circumstances in which the target differentiated cell can be identified by gene expression.

An alternative approach for achieving homogeneous differentiated populations from ES cells is the use of modified culture conditions to enrich for specific lineages. While selective conditions can provide some advantage,[31] directed differentiation of ES cells to homogeneous populations is particularly attractive. Homogeneous formation of progenitor cells, termed early primitive ectoderm-like (EPL) cells,[32, 33] provides an excellent starting population for further differentiation into numerous cell lineages. In particular, synchronous and homogeneous formation of neural precursor cells in a manner that recapitulates establishment of this lineage in the embryo has been achieved, and both glia and neural crest can be derived as near homogeneous populations from these cells in response to appropriate signalling environments.[34] An alternative approach is genetic modification of ES cells to ensure directed differentiation into particular cell types. This has generally involved the enforced expression of critical genes that direct cell fate.[27, 35, 36]

**What cell types can be generated?**

To date, a wide range of therapeutically relevant cell types has been produced from

mouse ES cells, utilising diverse approaches to enrich for the desired cells. Such cell types include haematopoietic cells,[36, 37] motor neurons,[38] insulin secreting cells,[35, 39] cardiomyocytes,[40] hepatocytes,[26, 28] chondrocytes,[41] osteoblasts,[18] oligodendrocytes,[29, 30, 42] and dopaminergic neurons.[27, 43] Directed differentiation of human ES cells is not yet as advanced, although the isolation of endothelial cells,[44] neurons[45, 46] and haematopoietic cells[47] has been reported.

Many of these ES-derived cell types have been shown to be equivalent to embryonic tissues using gene expression markers, differentiation potential and function in vitro. For example, ES-derived cardiomyocytes display structural and functional properties of early-stage or foetal cardiomyocytes as determined by gene expression, myofibrillar organisation and electrophysiological responses.[48, 49] Similarly, neurons produced from mouse ES cells express a variety of tissue-specific genes, ion channels and receptors, and can generate both excitatory and inhibitory synaptic connections in vitro, characteristic of functional post-mitotic neurons.[50]

*Are ES-derived cells functional in vivo?*

Several specialised cell types generated by ES cell differentiation have now been transplanted successfully into mice (Figure 1E), demonstrating that the cells can survive, are not teratogenic if purified away from pluripotent cells, and can integrate into host tissues. Examples include reconstitution of the haematopoietic system in mice with cells produced by ES cell differentiation in vitro[36, 51] and the integration of ES-derived cardiomyocytes into heart muscle.[52]

Even more promising, ES-derived dopaminergic neurons were shown to ameliorate symptoms in a rat model of Parkinson's Disease,[27] while ES-derived insulin secreting cells resulted in the normalisation of hyperglycaemia in a mouse model of diabetes.[35, 39] These experiments provide proof of principle of the functionality of ES-derived cells in vivo and their ability to treat disease in relevant animal models.

Cell therapy using adult stem cells other than HSC is less advanced, although the clinical use of human corneal stem cells is being investigated[53] and numerous studies to evaluate the use of neural stem cells in pre-clinical models of spinal cord injury and neurological disorders have been undertaken. A notable example is the recent work of Pluchino et al[54] showing that ex vivo expansion and transplantation of mouse neural stem cells can induce recovery in a chronic mouse model of multiple sclerosis.

**Combining gene therapy with cell therapy**

Gene therapy is the provision of specific genes to replace or supplement the activity of defective genes, or the introduction of a gene that provides a new functionality to treat a disease or disorder. Delivery of the corrective gene into target cells and control of its expression in these cells have been major difficulties associated with this technology. The ability to couple gene therapy with stem cell therapy is attractive as cells can be genetically modified in vitro and those cells demonstrating appropriate expression of the introduced gene can then be selected for transplantation. Transplantation of stem cells, with their capacity for self-renewal, is likely to eliminate the need for repeated gene therapy treatments.

Methods for genetically modifying adult stem cells for gene therapy applications are being pursued, particularly with HSC, but also with myoblasts, osteoblasts and neural stem cells. At present, these methods commonly rely on random integration of the introduced genes into the genome and, as such, there is little control over integration

site effects such as inappropriate expression of the introduced gene, and activation or inactivation of neighbouring genes.

Use of ES cells to treat genetic disease potentially provides a gene therapy of greater sophistication based on their demonstrated capacity to support modification of endogenous genes by homologous recombination[55, 56] (Figure 1F). Direct repair of defective gene loci would circumvent problems associated with gene introduction and random integration. The proliferative capacity of ES cells enables efficient selection, expansion and differentiation of the appropriately modified cells prior to transplantation. Proof of concept for this approach has been demonstrated by genetic manipulation of mouse ES cells to repair a single gene defect in immunodeficient mice. The modified ES cells were differentiated in vitro into haematopoietic cells and engrafted back into syngeneic immunodeficient mice to restore immune function.[51]

**Hurdles to overcome**

The potential of stem cells for the production of specific, specialised cell types for transplantation is clear and the benefits are likely to be enormous. However, a number of ethical and scientific challenges will need to be met before this technology can be translated into the clinic. Ongoing scientific investigation of both adult stem cells and ES cells is directed towards the development of methodologies for: (1) generation of sufficient numbers of pure cell populations by expansion of stem cell populations; (2) determination of the optimum cell type for transplantation, whether it be a committed progenitor cell or a terminally differentiated cell; (3) rigorous testing of the functionality and efficacy of the cells in predictive animal models; (4) the method of cell delivery, either transplantation directly to the site of engraftment or incorporation into a scaffold and (5) optimal delivery formulation, such as the inclusion of specific growth factors to enhance survival, engraftment or correct differentiation of transplanted cells.

*Immunorejection of transplanted cells*

Immunological incompatibility between stem cell-derived donor cells and the transplant recipient is an important issue that must also be addressed. In the event that transplantation of adult stem cell populations can be used for therapeutic benefit this is unlikely to be a serious limitation because purification and use of the patient's own cells would eliminate any requirement for immunosuppression.

However, differentiated cells generated from ES cells will express the histocompatibility antigens of the donor embryo from which they were derived and are therefore likely to elicit an immune response when grafted into patients. One proposed solution to this problem is the establishment of stem cell banks (Figure 1G), such as that launched recently in the United Kingdom (see http://www.nibsc.ac.uk/divisions/cbi/stemcell.html). These banks would contain large numbers of human ES cell lines, providing a wide range of histocompatibilities. Other strategies include: genetic modification of the primary histocompatibility antigen genes in the ES cell lines such that transplanted cells escape immunosurveillance, and the induction of immune tolerance in recipients by manipulation of T cell activity, an approach that holds promise for widespread use with allogeneic transplants.

Use of autologous ES cells is perhaps the most direct method to overcome this problem. The cloning of Dolly the sheep demonstrated that a differentiated, somatic mammalian cell could be "reprogrammed" to change both its differentiation state and

differentiation potential.[57] "Therapeutic cloning" envisages the creation of embryos by transfer of somatic cell nuclei into unfertilised oocytes (Figure 1H). ES cell lines derived from blastocysts produced in this manner would be autologous with the donor of the somatic nucleus, envisaged to be the cell therapy recipient. Therapeutic cloning has been used successfully in mice to repopulate the haematopoietic system.[51] However, the extent of nuclear reprogramming upon cloning is highly variable and incomplete,[58] as evidenced by the very low frequency of successful animal cloning. Therefore, it will be crucial to ensure that ES cells generated in this manner can differentiate normally into functional cells.

Alternative approaches for creating autologous sources of pluripotential stem cells which potentially circumvent the requirement for human embryonic material are envisaged but at an early stage of development. These include production of oocytes from mouse ES cells[59] and multipotent stem cells from unfertilised, parthenogenetic oocytes.[60] Of particular interest are experiments suggesting that it might be possible to reprogram somatic cells to a pluripotent state by fusion with pluripotent cell cytoplasm[61] (Figure 1I).

Despite the obvious benefits of using stem cells derived from the patients themselves for transplantation in order to avoid problems of graft rejection, the use of allogeneic cells may also have advantages.[62] The establishment of banks of stem cell lines, whether embryonic or adult, would allow standardisation and quality control of cell production, differentiation and grafting procedures. Large batches of differentiated cells could be produced and tested prior to administration to patients. Such an approach would be expected to facilitate the widespread application of stem cell therapies.

*Ethical issues*

Present technologies for the isolation of human ES cells result in the destruction of human embryos, a critical ethical issue that remains to be resolved. Similarly, therapeutic cloning necessitates the creation and subsequent destruction of large numbers of human embryos. Political and social debate about the merits of ES cell research has led to diverse legislative outcomes in different countries. However, the solution to these ethical issues probably lies in research programs directed towards the development of non-destructive methodologies for human pluripotent cell isolation and use.

## Conclusions

The expectation that stem cell therapy will revolutionise regenerative medicine has a sound scientific basis. The technology to differentiate stem cells into virtually any desired cell type is advancing rapidly, although significant scientific and ethical challenges remain. As we begin to understand more about the factors that control and influence cellular plasticity and differentiation, protocols for the production of specific, specialised cell types will become increasingly refined. This will result in precise, safe, and efficacious applications of this technology and provide new opportunities for the treatment of many debilitating diseases and traumas.

## Acknowledgements

We gratefully acknowledge support from the Australian Research Council (ARC), the ARC Special Research Centre for Molecular Genetics of Development, the National Health and Medical Research Council of Australia, Raymond Ryce, and from

past and present members of the laboratory who have contributed to our understanding of stem cells.

## References


1. Daley GQ. *Curr Opin Genet Dev* 2002; 12:607-613.
2. Verfaillie CM. *Trends Cell Biol* 2002; 12:502-508.
3. Evans MJ, Kaufman MH. *Nature* 1981; 292:154-156.
4. Martin GR. *Proc Natl Acad Sci USA* 1981; 78:7634-7638.
5. Thomson JA, Kalishman J, Golos TG et al. *Proc Natl Acad Sci USA* 1995; 92:7844-7848.
6. Thomson JA, Itskovitz-Eldor J, Shapiro SS et al. *Science* 1998; 282:1145-1147.
7. Reubinoff BE, Pera MF, Fong CY, Trounson A, Bongso A. *Nat Biotechnol* 2000; 18:399-404.
8. Smith AG. *Seminars in Cell Biology* 1992; 3:385-399
9. Amit M., Carpenter MK, Inokuma MS et al. *Dev Biol* 2000; 227:271-278.
10. Williams RL, Hilton DJ, Pease S et al. *Nature* 1988; 336:684-687.
11. Smith AG, Heath JK, Donaldson DD et al. *Nature* 1988; 336:688-690.
12. Bradley A. *Curr Opin Cell Biol* 1990; 2:1013-1017.
13. Beddington RS, Robertson EJ. *Development (Cambridge, England)* 1989; 105:733-737.
14. Shamblott MJ, Axelman J, Wang S et al. *Proc Natl Acad Sci USA* 1998; 95:13726-13731.
15. Matsui Y, Zsebo K, Hogan BL. *Cell* 1992; 70:841-847.
16. Rathjen J, Rathjen PD. *Curr Opin Genet Dev* 2001; 11:587-594.
17. Odorico JS, Kaufman DS, Thomson JA. *Stem Cells* 2001; 19:193-204.
18. Itskovitz-Eldor J, Schuldiner M, Karsenti D et al. *Mol Med* 2000; 6:88-95.
19. Schuldiner M., Yanuka O., Itskovitz-Eldor J, Melton D, Benvenisty N. *Proc Natl Acad Sci USA* 2000; 97:11307-11312.
20. Verfaillie CM, Pera MF, Lansdorp PM. *Hematology (Am Soc Hematol Educ Program)* 2002; 369-391.
21. Grant MB, May WS, Caballero S et al. *Nat Med* 2002; 8:607-612.
22. Krause DS, Theise ND, Collector MI et al. *Cell* 2001; 105:369-377.
23. Schwartz RE, Reyes M, Koodie L et al. *J Clin Invest* 2002; 109:1291-1302.
24. Reyes M, Dudek A, Jahagirdar B et al. *J Clin Invest* 2002; 109:337-346.
25. Jiang Y, Vaessen, Lenvik T et al. *Exp Hematol* 2002; 30:896-904.
26. Yin Y, Lim YK, Salto-Tellez M et al. *Stem Cells* 2002; 20:338-346.
27. Kim JH, Auerbach JM, Rodriguez-Gomez JA et al. *Nature* 2002; 418:50-56.
28. Chinzei R, Tanaka Y, Shimizu-Saito K et al. *Hepatology* 2002; 36:22-29.
29. Billon N, Jolicoeur C, Ying QL et al. *J Cell Sci* 2002; 115:3657-3665.
30. Xian HQ, McNichols E, St Clair A, Gottlieb DI. *Stem Cells* 2003; 21:41-49.
31. Ying QL, Stavridis M, Griffiths D et al. *Nat Biotechnol* 2003; 21:183-186.
32. Lake J, Rathjen J, Remiszewski J, Rathjen PD. *J Cell Sci* 2000; 113( Pt 3):555-566.
33. Rathjen J, Lake JA, Bettess MD et al. *J Cell Sci* 1999; 112( Pt 5):601-612.
34. Rathjen J, Haines BP, Hudson KM et al. *Development* 2002; 129:2649-2661.
35. Blyszczuk P, Czyz J, Kania G et al. *Proc Natl Acad Sci USA* 2003; 100:998-1003.
36. Kyba M, Perlingeiro RC, Daley GQ. *Cell* 2002; 109:29-37.
37. Wiles MV, Keller G. *Development* 1991; 111:259-267.
38. Wichterle H, Lieberam I, Porter JA, Jessell TM. *Cell* 2002; 110:385-397.
39. Soria B, Roche E, Berna G et al. *Diabetes* 2000; 49:157-162.
40. Hughes S. *J Pathol* 2002; 197:468-478.
41. Kramer J, Hegert C, Guan K et al. *Mech Dev* 2000; 92:193-205.
42. Liu S, Qu Y, Stewart TJ et al. *Proc Natl Acad Sci USA* 2000; 97:6126-6131.
43. Bjorklund L, Sanchez-Pernaute R, Chung S et al. *Proc Natl Acad Sci USA* 2002; 99:2344-2349.
44. Levenberg S, Golub JS, Amit M et al. *Proc Natl Acad Sci USA* 2002; 99:4391-4396.
45. Schuldiner M, Eiges R, Eden A et al. *Brain Res* 2001; 913:201-205.
46. Carpenter MK, Inokuma MS, Denham J et al. *Exp Neurol* 2001; 172:383-397.
47. Kaufman DS, Hanson ET, Lewis RL et al. *Proc Natl Acad Sci USA* 2001; 98:10716-10721.
48. Kehat I, Kenyagin-Karsenti D, Snir M et al. *J Clin Invest* 2001; 108:407-414.
49. Doevendans PA, Kubalak SW, An RH et al. *J Mol Cell Cardiol* 2000; 32:839-851.
50. Guan K, Chang H, Rolletschek A, Wobus AM. *Cell Tissue Res* 2001; 305:171-176.
51. Rideout WM 3rd, Hochedlinger K, Kyba M et al. *Cell* 2002; 109:17-27.
52. Klug MG, Soonpaa MH, Koh GY, Field LJ. *J Clin Invest* 1996; 98:216-224.

53. Daniels JT, Dart JK, Tuft SJ, Khaw PT. *Wound Repair Regen* 2001; 9:483-494.
54. Pluchino S, Quattrini A, Brambilla E et al. *Nature* 2003; 422:688-694.
55. Zwaka TP, Thomson JA. *Nat Biotechnol* 2003; 21:319-321.
56. Thomas KR, Capecchi MR. *Cell* 1987; 51:503-512.
57. Campbell KH, McWhir J, Ritchie WA, Wilmut I. *Nature* 1996; 380:64-66.
58. Hochedlinger K, Jaenisch R. *Curr Opin Cell Biol* 2002; 14:741-748.
59. Hubner K, Fuhrmann G, Christenson LK et al. *Science* 2003; 300:1251-1256.
60. Lin H, Lei J, Winnger D et al. *Stem Cells* 2003; 21:152-161.
61. Tada M, Takahama Y, Abe K, et al. *Curr Biol* 2001; 11:1553-1558.
62. Strom TB, Field LJ, Ruediger M. *Curr Opin Immunol* 2002; 14:601-605.

# Tetanus and the Anaesthetist

MARY PINDER, MB, BS, MRCP(UK), FRCA, FJFICM
Staff Specialist, Director of Research, Intensive Care Unit,
Sir Charles Gairdner Hospital, Perth

A/PROFESSOR JEFFREY LIPMAN, MB, BCh, DA, FFA(Crit Care), FJFICM
Director, Department of Intensive Care Medicine, Royal Brisbane Hospital, and
Head, Department of Anaesthesiology and Critical Care, University of Queensland

*Mary Pinder trained in intensive care in the UK, Fiji, South Africa and Australia. She is currently working as Staff Specialist in Intensive Care at the Sir Charles Gairdner Hospital. Her research interests include crisis management training, alveolar recruitment strategies, hormonal response to sepsis and critical care in developing countries.*

*Since qualifying as an anaesthetist in Johannesburg, Jeffrey Lipman has been working full-time in Intensive Care Units since1979. Prior to moving to Brisbane in 1997, he was Head of Intensive Care at Baragwanath Hospital, where his interest in tetanus originated. He has published over 100 articles in peer-reviewed journals on a wide variety of topics. His research interests most recently have been concentrated on sepsis and related topics, particularly antibiotic pharmacokinetics.*

Tetanus is an acute, often fatal disease caused by the exotoxins produced by *Clostridium tetani*. It was first described more than two thousand years ago. The disease is characterised by generalised muscle rigidity, autonomic instability and sometimes convulsions. It is preventable by adequate immunisation programs, but remains common world wide. Although predominantly a disease of developing countries, the management of tetanus remains a challenge to critical care specialists in developed countries because,[1, 2]

- Tetanus presenting in its early stages, when it is most amenable to treatment, is often unrecognized. Severe and even fatal cases are under reported.
- The incidence in developed countries is increasing and at risk populations include the elderly, the immunocompromised and intravenous drug abusers.
- Severe tetanus requires prolonged intensive care and organ support, with the associated complications including nosocomial sepsis.
- Autonomic dysfunction in severe tetanus is difficult to control and is a significant cause of mortality.
- Reports in the literature relating to the management of severe tetanus are sometimes conflicting. They constitute only level 3 or 4 evidence.

**Pathophysiology**

Tetanus is caused by *Clostridium tetani*, an anaerobic Gram-positive bacillus. The mature organism forms a terminal spore which is resistant to oxygen, moisture and extremes of temperature and is capable of surviving decades. Spores exist ubiquitously

in soil and in animal and human faeces. Following entry into a wound, spores proliferate as the vegetative form, if the local tissue oxygen tension is sufficiently low. The usual mode of entry is through a puncture wound or laceration, although tetanus may follow surgery, burns, gangrene, chronic ulcers, dog bites, injections such as with drug users, dental infection, abortion and childbirth. Tetanus neonatorum usually follows infection of the umbilical stump. The injury itself may be trivial, and in as many as 29% of cases there is no history or evidence of a wound.[1] The oxidation-reduction potential of the tissues is lowered by factors such as the presence of necrotic tissue, pus and foreign bodies.

In the vegetative form, spores produce two toxins: tetanospasmin and tetanolysin. Tetanolysin is capable of further damaging otherwise viable tissue surrounding an infected wound, optimizing conditions for bacterial multiplication. Tetanospasmin causes the clinical manifestations of tetanus. This toxin is distributed widely by the blood stream, entering the nervous system at the neuromuscular junction (NMJ) of alpha-motor neurons. Once tetanospasmin is bound to the NMJ, inhibition of transmitter release is permanent and recovery depends on the formation of a new synapse. Toxin is then carried by retrograde axonal transport to the cell body and migrates trans-synaptically to other neurones. The cells most commonly affected by tetatanospasmin are the alpha-motor neurone, parts of the sympathetic nervous system and the spinal inhibitory cells.

The action of spinal inhibitory cells is mediated at a local level by glycinergic interneurones. Release of glycine from these cells is directly inhibited by tetanospasmin, preventing the damping down of reflex muscle action. The consequent loss of inhibition leads to the clinical manifestations of tetanus, the characteristic rigidity and spasms.

A further inhibitory system connecting the brain stem to the spinal cord, mediated by the transmitter gamma-aminobutyric acid (GABA) is also affected by tetanospasmin. There is thought to be blood borne spread of toxin from the entry site to the brain at the area of the floor of the fourth ventricle, where the blood-brain barrier is non-existent. This may explain why trismus and neck rigidity are early manifestations of the disease.

**Active immunoprophylaxis**

Natural immunity to tetanus does not occur. The disease may both relapse and recur. Victims of tetanus must be actively immunised. Tetanus toxoid is a cheap and effective vaccine, which is thermally stable.[3] It is a non-toxic derivative of the toxin which, nevertheless, elicits and reacts with antitoxic antibody. A serum antibody titre of 0.01 units/ml is universally quoted as being protective. However, this figure was extrapolated directly from animal experiments and there are reports of tetanus in individuals with much higher serum antibody titres.[1, 3] The diagnosis of tetanus should not be excluded on the basis of toxin-neutralising antibody $\geq$ 0.01units/ml, if the history and clinical features are supportive.

In adults, a full immunization course consists of three toxoid doses, given at an optimal interval of 6-12 weeks between the first and second doses, and 6-12 months between the second and third doses. Neonates have immunity from maternal antibodies. Children over 3 months should be actively immunized, and need four doses in total. Two or more doses to child-bearing females over 14 years will protect any child produced within the next five years. Pregnant women who are not immunised should

thus be given two spaced-out doses two weeks to two months before delivery. Booster doses should be given routinely every ten years, but the full immunisation course should not be repeated.

Side-effects of tetanus toxoid are uncommon and not life-threatening. They are associated with excessive levels of antibody due to indiscriminate use. Common reactions include urticaria, angio-oedema and diffuse, indurated swelling at the site of injection.

**Diagnosis**

The diagnosis of tetanus relies on clinical assessment. There is no diagnostic laboratory test and blood cultures are positive in less than half of cases. The absence of an obvious wound does not exclude the diagnosis.[1] The most common differential diagnosis is dystonic reaction to tricyclics. Others include strychnine poisoning, local temporomandibular disease, local oral disease, convulsions, tetany, intracraninal infections or haemorrhage and psychiatric disorders.

Suggested criteria for the clinical diagnosis of tetanus are:

1. An illness characterized by the acute onset of hypertonia and/or painful muscular contractions (usually of the jaw and neck) and generalized muscle spasms, without any other apparent causes such as drug reactions, other central nervous system disorders, or hysteria.
2. No history of contact with strychnine.
3. Subsequent disease course consistent with tetanus.

**Clinical presentation**

The incubation period (i.e. time from injury to onset of symptoms) varies from 2 to 60 days. The period of onset (i.e. from first symptom to first spasm) similarly varies. However, nearly all cases (90%), present within 15 days of infection.[4] The incubation period and the period of onset are of prognostic importance, with shorter times signifying more severe disease.

The presenting symptoms are pain and stiffness. Stiffness progresses to rigidity, with difficulty in mouth opening — trismus or lockjaw. The rigidity becomes generalised, with the facial muscles producing a characteristic clenched-teeth expression called risus sardonicus. The disease then descends, to affect the trunk and limbs. Typical spasms, with flexion and adduction of the arms, extension of the legs and opisthotonos, are very painful and may be so intense that fractures and tendon separations occur.[1] Spasms are caused by external stimuli, e.g. noise and pressure. As the disease worsens, even minimal stimuli produce more intense and longer-lasting spasms. Spasms are life threatening when they involve the larynx and/or diaphragm.

Neonatal tetanus presents most often on day 7 of life,[5] with a short (one day) history of failure of the infant to feed. The neonate displays typical spasms that can be easily misdiagnosed as convulsions of another aetiology. In addition, because these infants vomit (as a result of the increased intra-abdominal pressure) and are dehydrated (because of their inability to swallow), meningitis and sepsis are often considered first.

Autonomic dysfunction occurs in severe cases,[5, 6] and begins a few days after the muscle spasms (see below). If tetanus of uterine origin is associated with autonomic dysfunction, the prognosis is very poor.

Local tetanus is an uncommon mild form of tetanus with a mortality of 1%. The signs and symptoms are confined to a limb or muscle, and may be the result of

immunization. Cephalic tetanus is also rare. It results from head and neck injuries, eye infections and otitis media. The cranial nerves, especially the seventh are frequently involved, and the prognosis is poor. This form may progress to more generalized disease. In heroin addicts, tetanus seems to be severe with a high mortality, although reported numbers are small.

**Autonomic dysfunction**

Sympathetic overactivity in tetanus was first described by Kerr and others in 1968.[5] The hallmarks were reported to be fluctuating tachycardia and hypertension, sometimes followed by hypotension, cardiac dysrhythmias, peripheral pallor, sweating and pyrexia. These are associated with increased plasma and urinary concentrations of catecholamines, increased oxygen consumption and carbon dioxide production.

Proposed mechanisms for autonomic dysfunction include:[1, 7]

- The effect of toxin on the brainstem and autonomic interneurones causing impairment of inhibitory pathways. This is the most likely mechanism.
- A direct effect of toxin on the myocardium.
- Loss of inhibition of the adrenal medulla with increased adrenaline secretion.
- Direct inhibition of the release of endogenous opiates by tetanospasmin.
- Increased release of thyroid hormone.[1, 2]

Fluctuations in blood pressure and heart rate appear to be related to changes in systemic vascular resistance, rather than cardiac output or left ventricular filling pressure.[6] Noradrenaline is increased disproportionately relative to adrenaline, suggesting a direct sympathetic mechanism rather than humoral release. Another possible cause is a seizure-like discharge causing a sudden tachycardia and vasoconstriction, increasing central venous and arterial pressure. Cessation of this seizure-like activity results in hypotension, bradycardia, decreased central venous pressure and vasodilation.

Sympathetic overactivity seems to predominate in this process, but it is possible that effects on parasympathetic nervous system have been largely overlooked and under-reported. For example, sudden cardiac arrest is a common complication of severe tetanus but its cause has not always been ascribed to autonomic dysfunction.[7] Preterminal bradycardia, salivation and increased bronchial secretions have also been reported. It should be noted that such patients have often also been on beta-adrenergic blocking agents, which confuses the interpretation of their signs.

**Management**

The management of patients with tetanus has five goals: neutralisation of circulating toxin (i.e. passive immunisation), eradication of the organism and removal of the source of toxin, control of rigidity and spasms, management of autonomic dysfunction and general supportive care. Many anaesthetic drugs and techniques may be of value in the control both of spasms and autonomic dysfunction.

**Passive immunization**

Human antitetanus serum has now largely replaced equine serum (ATS), as it is less antigenic. Although never prospectively tested, the present dosage recommendation is 3000-6000 units IM (or IV if available).[1, 2] It has been suggested that patients who are unimmunised or whose immunisation status is unknown should be given human rich antiserum on presentation with contaminated wounds. No controlled study has shown

this to be more effective than wound toilet and penicillin administration. Intrathecal administration of antitetanus serum is still controversial. A large meta-analysis reported it to have no advantage over the intramuscular route.[8]

The side effects of human antitetanus serum include fever, shivering and chest or back pains. Cardiovascular parameters need to be monitored, and the infusion may need to be stopped temporarily if significant tachycardia and hypotension present. If human antiserum is not available, equine ATS can be used after testing and desensitization.[1]

**Eradication of the organism**

*Wound care*

Once human antitetanus toxin has been given, the infected site should be thoroughly cleaned and all necrotic tissue extensively debrided. Hysterectomy is advocated for tetanus associated with septic abortion or childbirth. Wound debridement may not always be practicable, for example in drug addicts or those with no obvious wound.

*Antibiotics*

Tetanus spores are destroyed by antibiotics. The vegetative form (bacillus) is sensitive to antibiotics in vitro. However, in vivo efficacy depends on the antibiotic concentration at the wound site, and large doses may be required. Recommended antibiotic regimens include:

1. Metronidazole 500 mg IV 8-hourly for 10 days: The drug has a spectrum of activity against anaerobes, is able to penetrate necrotic tissue, and has been shown to be more effective than penicillin in this situation.[9]
2. Penicillin G 1-3 Mu IV 6-hourly for 10 days. Penicillin is a GABA antagonist in the CNS, and may aggravate the spasms.

Overall penicillin is the antibiotic of choice. Erythromycin, tetracycline, chloramphenicol and clindamycin are alternatives.[2]

**Control of rigidity and spasms**

In the early stages of tetanus, the patient is most at risk from laryngeal and other respiratory muscle spasm. Therefore, if muscle spasms are present, the airway should be urgently secured by endotracheal intubation or tracheostomy. If respiratory muscles are affected, mechanical ventilation is instituted. In severe tetanus, spasms usually preclude effective ventilation and muscle relaxants may be required. No one drug or class of drugs has been shown to be consistently effective in controlling rigidity and spasms in severe disease. A wide range of drugs suppressing central and/or peripheral nervous activity has been used, both singularly and in combination. Neuromuscular blocking agents (NMBAs) are often combined with heavy sedation to control severe spasms. There is no consensus on the most appropriate NMBA to use in tetanus. Pancuronium bromide may cause tachycardia and hypertension, by stimulating noradrenaline release from sympathetic nerve endings, but has been used safely in tetanus. Neuromuscular blockade has been avoided by using agents such as dantrolene, intrathecal baclofen and propofol.[10-12] Heavy sedation alone may prevent muscle spasms and improve autonomic dysfunction (see below).

**Management of autonomic dysfunction**

More than thirty years since the syndrome was first described, management of

autonomic dysfunction is still sub-optimal. As with control of spasms, no one drug or class of drug has been shown to be consistently effective in the control of autonomic dysfunction. Treatments include the use of heavy sedation, peripheral adrenergic blocking agents, morphine, magnesium sulphate, baclofen, clonidine, chlorpromazine, ganglion-blockers and atropine, often in combination. Other strategies include epidural analgesia and atrial pacing.

Evaluation of treatment is complicated by the lack of defined criteria for the diagnosis of autonomic dysfunction in tetanus. Studies differ in the haemodynamic variables monitored and their method of measurement. Catecholamine concentrations are not always reported and the method of measurement may have been inaccurate in some of the older studies. Lastly, there is no standardisation of other aspects of management. In retrospective studies, it is impossible to distinguish cases of true autonomic dysfunction from, for example, drug withdrawal syndromes, stress response in the setting of inadequate sedation, labile blood pressure in a patient with poorly-controlled pre-existing cardiovascular disease, or the systemic inflammatory response syndrome.

*Beta-adrenergic blocking agents*

These were the earliest drugs advocated for the management of autonomic dysfunction. Propanolol was used at first, but it may result in profound hypotension, severe pulmonary oedema and cardiac arrest.[7] Labetalol has been used for its combined alpha- and beta effects, either alone or in combination with propanolol.[5] However, it has less effect on alpha blockade than on beta blockade. From the literature, it appears that combined alpha- and beta-blockade denervates the circulation and makes circulatory support during periods of hypotension difficult. Esmolol, an ultra-short acting beta-blocker, reduces catecholamine release in shocked animals and has been used in tetanus.[13] In one report, it provided haemodynamic stability with good control of tachycardia and hypertension. However, arterial catecholamine concentrations remained markedly increased. This may be of concern as control of the haemodynamic disturbance without reduction in catecholamine levels has been associated with fatal myocardial necrosis in phaeochromocytoma. This situation might be aggravated in tetanus by the direct action of the toxin on the myocardium.

*Post-ganglionic and alpha-adrenergic blocking agents*

Bethanidine, guanethidine and phentolamine were used as adjuncts to propranolol by Prys-Roberts and Kerr, with apparent beneficial effect.[14] Other similar agents that have been used include trimetaphan, phenoxybenzamine and reserpine. A disadvantage of this group of drugs is that induced hypotension may be difficult to reverse and withdrawal can result in rebound hypertension.

*Morphine*

Morphine acts centrally to reduce sympathetic tone to the heart and vascular system. Morphine helps control cardiovascular instability, but it is unclear whether catecholamine concentrations are reduced. Gastrointestinal effects such as constipation and ileus may compound the effects of autonomic dysfunction.

*Magnesium sulphate*

Magnesium inhibits release of humoral and neuronal catecholamines and reduces

the sensitivity of alpha-adrenergic receptors. It has been used to treat autonomic dysfunction with varying success, appearing to reduce pulse rate, vasoconstriction and catecholamine concentrations when used with sedatives.[15, 16] Negative inotropic effects are minimised, but not eliminated, if hypocalcaemia is avoided. Marked neuromuscular blockade is associated with the use of magnesium; with the simultaneous use of NMBAs, this may predispose to prolonged weakness during the recovery phase. Intravenous magnesium sulphate is given in doses similar to those used for the treatment of pre-eclampsia, keeping serum levels in the range 2.5-4.0 mmol/l.

*Epidural blockade*

Lumbar epidural blockade with local anaesthetic agents has controlled cardiovascular instability and reduced serum catecholamine concentrations. Epidural blockade provides sympathetic blockade, neuromuscular blockade and analgesia.[17] There are only individual case reports of this technique and they provide insufficient evidence to draw firm conclusions on its value.

*Clonidine*

Clonidine is a selective partial agonist for alpha$_2$-adrenergic receptors with a 200-fold selectivity for the alpha$_2$-receptor over the alpha$_1$. Alpha$_2$ agonists have both peripheral and central effects on the cardiovascular system. Acting centrally, clonidine causes hypotension and bradycardia. The precise mechanisms of action is not fully understood, but is thought to involve inhibition of sympathetic outflow and potentiation of parasympathetic nervous activity, increasing vagal tone. Acting peripherally, clonidine inhibits the release of nor-adrenaline from prejunctional nerve endings. It also produces marked sedation and anxiolysis, decreases spontaneous motor activity and potentiates the sedative and anaesthetic actions of other drugs. Two case studies report the use of clonidine in the management of autonomic dysfunction in tetanus with conflicting results.[16, 18] Clonidine is available as an oral, intravenous, intrathecal and transdermal preparation. Dexmedetomidine, a superselective alpha$_2$-adrenergic agonist related to clonidine but more potent and more selective, may prove to be more effective in the management of autonomic dysfunction.

*Inhalational anaesthetic agents*

Nitrous oxide, halothane and isoflurane have all been reported as being used in the management of tetanus.[14, 19] Nitrous oxide and halothane were shown to adequately control haemodynamic instability. Isoflurane, in relatively high concentrations, was shown to control spasms but the case reported did not demonstrate significant autonomic dysfunction. Many intensive care units lack the facilities for the administration, monitoring and scavenging of such agents, limiting the potential usefulness of this treatment strategy.

*Other drugs and treatment strategies*

A continuous infusion of atropine has been advocated for the management of parasympathetic dysfunction.[20] Atrial pacing may be of use in the treatment of hypotension and bradycardia associated with autonomic dysfunction.[21] Other treatments that have been reported as isolated case studies include sodium nitroprusside, steroids, hypothermia and hyperbaric oxygen.[1, 2]

Drugs of potential benefit include sodium valproate, angiotensin converting enzyme

(ACE) inhibitors and adenosine. Sodium valproate blocks GABA-aminotransferase and is therefore an inhibitor of GABA metabolism. Angiotensin II increases noradrenaline synthesis and facilitates its release from nerve endings. Inhibition of angiotensin II synthesis may therefore reduce sympathetic overactivity in tetanus. Adenosine has actions that include reduction of presynaptic noradrenaline release and antagonism of the inotropic effects of catecholamines. Adenosine, therefore, has potential for use in sympathetic overactivity. However, to date its use has been largely experimental with clinical use limited to isolated bolus doses.[22]

**Supportive treatment**

Steps should be taken to prevent contractures, nosocomial pneumonias and deep vein thrombosis. The patient (including the mother if a neonate is afflicted) must be actively immunised. Where possible, supportive psychotherapy should be offered to both patient and family.

**Complications**

Muscle spasms disappear after 1-3 weeks, but residual stiffness may persist. Although most survivors recover completely by 6 weeks, cardiovascular complications, including cardiac failure, arrhythmias, pulmonary oedema and hypertensive crises can be fatal. No obvious cause of death can be found at autopsy in up to 20% of deaths.

**Conclusion**

Severe tetanus remains a clinical challenge. Although autonomic dysfunction was first described more than thirty years ago and, despite the extensive range of pharmacological agents available, its control remains difficult in severe cases. Current evidence is often anecdotal as controlled trials are difficult to conduct. There are advocates for the use of epidural blockade, magnesium sulphate and esmolol to control spikes of hypertension and tachycardia, in addition to sedation and neuromuscular blockade. Further assessment of the use of dexmedetomidine, to control autonomic dysfunction, and dantrolene and intrathecal baclofen, to control spasms, is needed. Finally, it should not be forgotten that tetanus is completely preventable by adequate immunisation programs.

**References**

1. Bleck TP. Tetanus: Pathophysiology, management and prophylaxis. *Dis Mon* 1991; 37:556-603.
2. Trujillo MH, Castillo A, Espana JV, Manzo A, Zerpa R. Impact of intensive care management on the prognosis of tetanus: analysis of 641 cases. *Chest* 1987; 92:63-65.
3. Passen EL, Anderson BR. Clinical tetanus despite "protective" level of toxin-neutralizing antibody. *JAMA* 1986; 255:1171-1173.
4. Sun KO, Chan YW, Cheung RT, So PC, Yu YL, Li PC. Management of tetanus: a review of 18 cases. *JR Soc Med* 1994; 87:135-137.
5. Kerr JH, Corbett JL, Prys-Roberts C, Crampton Smith A, Spalding JMK. Involvement of the sympathetic nervous system in tetanus. *Lancet* 1968; 2:236-241.
6. Domenighetti GM, Savary G, Stricker H. Hyperadrenergic syndrome in severe tetanus: extreme rise in catecholamines responsive to labetalol. *Br Med J* 1984; 288:1483-1484.
7. Buchanan N, Smit L, Cane RD, De Andrade M. Sympathetic overactivity in tetanus: Fatality associated with propanolol. *Br Med J* 1978; 2:254-255.
8. Abrutyn E, Berlin JA. Intrathecal therapy in tetanus: A meta-analysis. *JAMA* 1991; 266:2262-2267.
9. Ahmadsyah I, Salim A. Treatment of tetanus: An open study to compare the efficacy of procaine penicillin and metronidazole. *Br Med J* 1985; 291:648-650.

10. Checketts MR, White RJ. Avoidance of intermittent positive pressure ventilation in tetanus with dantrolene therapy. *Anaesthesia* 1993; 48:969-971.
11. Saissy JM, Demaziere J, Vitris M et al. Treatment of severe tetanus by intrathecal injections of baclofen without artificial ventilation. *Intensive Care Med* 1992; 18:241-244.
12. Borgeat A, Popovic V, Schwander D. Efficacy of a continuous infusion of propofol in a patient with tetanus. *Crit Care Med* 1991; 19:295-297.
13. King WW, Cave DR. Use of esmolol to control autonomic instability of tetanus. *Am J Med* 1991; 91: 425-428.
14. Prys-Roberts C, Corbett JL, Kerr JH, Crampton Smith A, Spalding JMK. Treatment of sympathetic overactivity in tetanus. *Lancet* 1969; 1:542-546.
15. Lipman J, James MFM, Erskine J, Plit ML, Eidelman J. Autonomic dysfunction in severe tetanus: Magnesium sulphate as an adjunct to deep sedation. *Crit Care Med* 1987; 15:987-988.
16. Sutton DN, Tremlett MR, Woodcock TE, Nielsen MS. Management of autonomic dysfunction in severe tetanus: the use of magnesium sulphate and clonidine. *Intensive Care Med* 1990; 16:75-80.
17. Southorn PA, Blaise GA. Treatment of tetanus-induced autonomic nervous system dysfunction with continuous epidural blockade. *Crit Care Med* 1986; 14:251-252.
18. Brown JL, Sinding H, Mathias CJ. Autonomic disturbance in severe tetanus: failure of parenteral clonidine to control blood pressure. *J Infect* 1994; 67-71.
19. Stevens JJWM, Griffen RM, Stow PJ. Prolonged use of isoflurane in a patient with tetanus. *Brit J Anaesth* 1993; 70:107-109.
20. Dolar D. The use of continuous atropine infusion in the management of severe tetanus. *Intensive Care Med* 1992; 18:26-31.
21. Brand GR, Breheney FX. Atrial pacing in tetanus. *Anaesth Intens Care* 1984; 12:63-65.
22. Kamibayashi T, Maze M. Clinical uses of alpha$_2$-adrenergic agonists. *Anesthesiology* 2000; 93: 1345-1349.

# The Management of Traumatic Brain Injury: Is There a Place for Therapeutic Hypothermia?

YUGAN M. MUDALIAR, FCP(SA), FRACP, FJFICM
Intensive Care Unit, Westmead Hospital, Westmead, NSW

*Yugan Mudaliar is Medical Director of the Intensive Care Unit at Westmead Hospital and a Clinical Senior Lecturer in the Faculty of Medicine at the University of Sydney. He is an Associate Fellow of the Australasian College of Physical Scientists of Engineering in Medicine and of the Australian Institute of Physics. In his clinical role, Dr Mudaliar has a special interest in neurological resuscitation.*

**Current Standards for the management of Traumatic Brain Injury**

The guidelines for the management of severe traumatic brain injury (TBI), were published as a joint project of the Brain Trauma Foundation and the American Association of Neurological Surgeons in 1995 and updated in 2000.[1] Standards were established for intracranial pressure (ICP) monitoring, ventricular CSF drainage and maintenance of adequate arterial oxygenation, with guidelines for sedation, paralysis and seizure prophylaxis. Treatment of ICP >20-25 mmHg and maintaining cerebral perfusion pressure (CPP) at more than 70 mmHg were established as standards.

First tier therapeutic interventions for raised ICP include hyperventilation to a $P_aCO_2$ between 30-35 mmHg and CSF drainage (Figure 1). Mannitol (0.25-1.0 g/kg) is also recommended for increased ICP. However, to achieve a serum osmolarity of <320 mOsm/l with mannitol therapy while simultaneously trying to maintain euvolaemia is clinically difficult. Second tier therapies for persistently elevated ICP include the use of barbiturates and hyperventilation to a $P_aCO_2$ <30 mmHg (Fig 1). Consideration was given to the use of jugular bulb oxygen saturation ($SjvO_2$) and cerebral blood flow (CBF) monitoring for patients with persistently elevated ICP.

The European Brain Injury Consortium (EBIC) has also established practical guidelines for the management of patients with TBI. The EBIC guidelines emphasise the need for monitoring of ICP, $SjvO_2$ and EEG (to detect and treat seizure activity and during high dose barbiturate therapy). Recommendations were made for treatment of ICP at or above 20 mmHg, CPP >60 (and preferably >70) mmHg, establishing the absolute threshold for $P_aCO_2$ not <30 mmHg, maintaining $SpO_2$ >95% and maintaining arterial haemoglobin level >60 g/l.[2]

These guidelines[1, 2] have attracted considerable criticism in the medical literature. Existing trials have not recruited a sufficient number of patients to confirm or refute the existence of a real benefit from the use of hyperventilation, mannitol, CSF drainage, and/or barbiturate therapy.[3, 4] A large-scale multi-centre study is currently being undertaken to evaluate the role of corticosteroid therapy after significant TBI.[5]

**Therapeutic Hypothermia for Traumatic Brain Injury?**

The use of therapeutic hypothermia for TBI is the subject of considerable

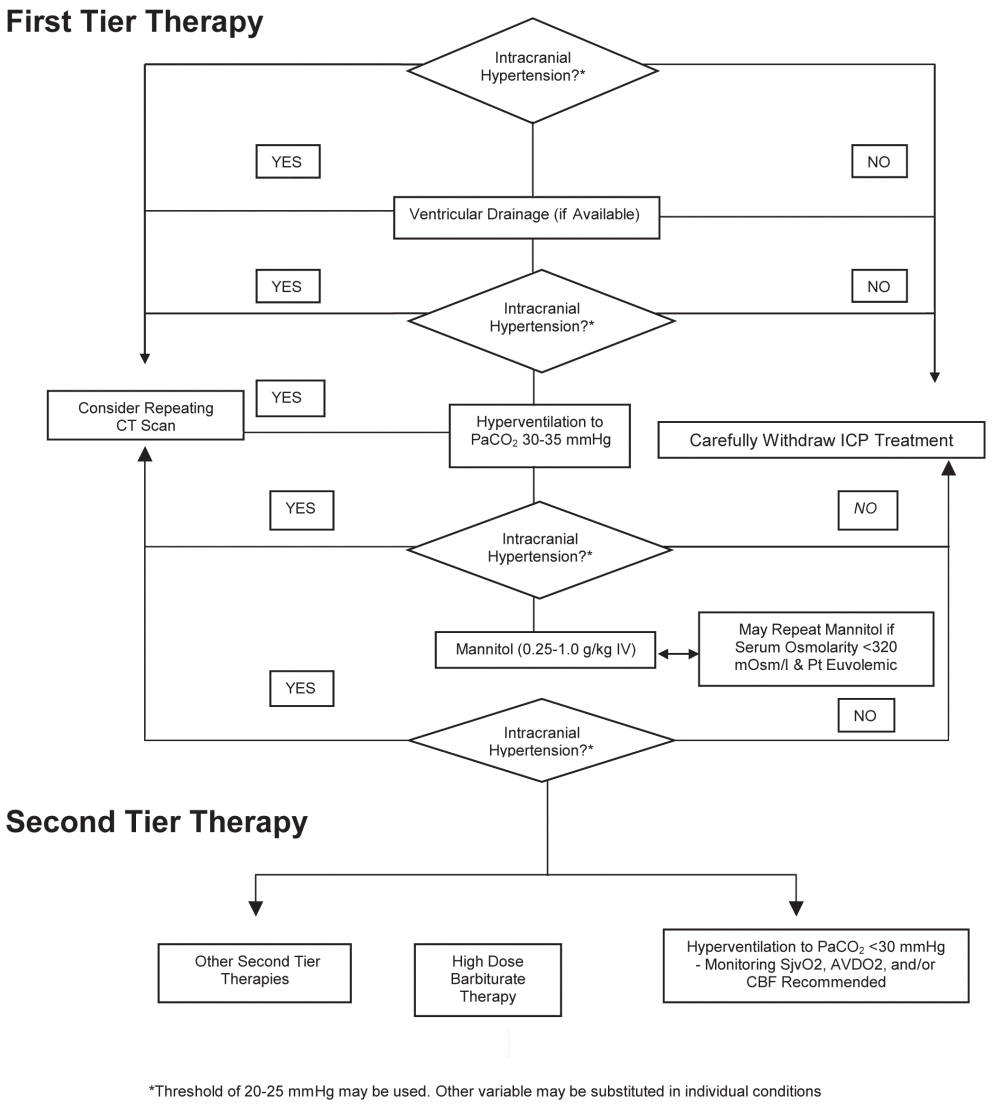


**Figure 1.** First tier therapy.

controversy in the medical literature. Despite several enthusiastic reports,[6, 7, 8, 9] a recent multi-centre trial by Clifton et al,[10] comparing the effects of hypothermia with normothermia among patients with severe TBI reported no clinical benefit. The Clifton study, enrolled 392 patients randomly assigned to be treated, and concluded that treatment with hypothermia, with the body temperature reaching 33°C within 8 hours after injury, is not effective in improving outcomes in patients with severe TBI. The patients in the hypothermia group had more hospital days with complications than the patients in the normothermia group.

**Physiological considerations**

Normal young adults, have a morning oral temperature of approximately 36.7°C (SD+0.2°C). Oral temperature in 95% of normal young adults varies between 36.3-37.1°C.[11] The normal core temperature has circadian fluctuations of 0.5-0.7°C and is lowest at 0600 hours. In women, there is in addition a monthly temperature variation characterised by a rise in basal body temperature at the time of ovulation. During anaesthesia and critical care, the body temperature is typically measured at the tympanic membrane, pulmonary artery, nasopharynx, oesophagus, rectum and/or skin. Intracranial temperature can be measured by using an epidural or intra-parenchymal probe. When intracranial temperature is measured using either a ventricular or epidural thermistor, the temperature gradient fluctuates between 0.4-1°C.[12] A temperature gradient of up to 2°C between core temperature and brain temperature has been described.[13]

Hypothermia is defined as a body temperature, below normal in a homeothermic mammal. Hypothermia is classified as:

- Mild: 35-36°C
- Moderate: 32-34°C
- Severe: <32°C

**Historical perspectives**

Aristotle said that the brain was an organ for cooling the blood.[14] Of course, the brain does cool or warm the blood and the body quite efficiently in the course of its operations. Hypothalamic reflexes initiated by cold cause shivering. Reflex responses activated by warmth are primarily controlled from the anterior hypothalamus, which triggers cutaneous vasodilatation and sweating. The signals which activate the hypothalamic temperature regulating centres come from two sources: temperature sensitive cells in the anterior hypothalamus and cutaneous cold receptors.

Hippocrates advocated snow and ice to check haemorrhage and was aware of the analgesic effects of cold. Baron Larrey, a surgeon in Napoleon's army, also became aware of the therapeutic effects of hypothermia in soldiers who had undergone painless amputation of limbs. Refrigeration anaesthesia was first described in 1866.[15]

Modern interest in the use of therapeutic hypothermia began in 1938-1940, with the reports by Smith and Fay[16] from the Temple University School of Medicine in Philadelphia. These researchers reported that treatment of human cancers with locally applied cooling to reduce temperatures to 75-90°F arrested tumour growth. These authors observed that a reduction in temperature induced physiologic changes in tumour cells that were comparable to those induced by irradiation therapy, i.e. nuclear destruction and cytoplasmic disintegration. Smith and Fay also noted the adverse effects of hypothermia on renal function, blood chemistry, and basal metabolic rate. They advised that medical personnel should only attempt the use of therapeutic hypothermia for cancer therapy in medical institutions with the capability of monitoring an individual patient's clinical course. They emphasised the benefit of low temperatures in producing pain control when applied to tissue compartments. Smith and Fay concluded that: "its (therapeutic hypothermia) usefulness in the therapeutic field otherwise remains a problem to be solved in the future". In a subsequent report, Fay[17] described his earlier experiences of localised and generalised refrigeration of the human brain to control a variety of clinical conditions including cancer, leukaemia, glioblastoma, Hodgkins Disease, filariasis, and syphilis.

The horrors of human experimentation, including the use of hypothermia, in the Nazi concentration camps between 1939 and 1945 were highlighted in the Nuremberg Trial and in Lifton's publication titled "Medical Killing and the psychology of genocide".[18]

**Patho-physiological rationale for the use of therapeutic hypothermia**

Hypothermia reduces intracranial pressure, increases cerebral perfusion pressure, lowers the cerebral metabolic rate, and reduces the cerebral spinal fluid concentration of interleukin-1β and glutamate. Mechanisms by which therapeutic hypothermia may be beneficial in, at least, sub-groups of patients with TBI include a significant reduction in excitatory amino acids during the period of cooling and sustained suppression of cytokines, particularly interleukin-1β. Stabilisation of the blood brain barrier and a general reduction in the post-traumatic hypermetabolic state also occur.

**Statistical Criticisms**

Despite these patho-physiological considerations, Hartung and Cottrell[19] made critical editorial comments regarding the lack of useful data to support the use of hypothermia in patients with TBI. Both these authors provide powerful insights into the methodologies and statistical analytical methods which were employed in the clinical trials for the use of therapeutic hypothermia for TBI. This critique, highlights the importance of sample size calculations in clinical studies; the role of inadvertent bias and perhaps inadvertent misinterpretations of *P* value probability differences. The authors conclude that: "Only regression and co-variant analyses address such differences statistically, referring to p values which are intended to serve the purpose of highlighting important differences from differences that should be dismissed, and then only when the study sample size is sufficiently large to make them meaningful."

**Experimental studies in Brain Injury**

In 1950, Bigelow et al first described hypothermic protection following cerebral ischaemia in dogs after total cardiac arrest for 15 minutes.[20] A comprehensive review of Phase I and Phase II animal studies supporting the use of therapeutic hypothermia in the treatment of severe TBI can be found in a review by Clifton and Hayes.[21] In the study on dogs by Rosomoff et al[22, 23] hypothermia was induced by immersion up to the shoulders in ice water. Rewarming was achieved by immersion in a water bath in which the temperature was maintained 10°C higher than body temperature until normothermic levels were reached. The CSF pressure measurements in the four groups of animals plotted against time are shown in Figures 2 and 3. In this study, it was demonstrated that application and maintenance of reduced body temperature clearly changed the pathologic character of experimental brain injury. Hypothermia prevented the development of progressive fulminating brain oedema associated with injuries at normal body temperature. Hypothermia also altered the post-traumatic inflammatory cellular response. Despite these beneficial effects, all the animals died when they were rewarmed. Their survival times were, however, five times longer than the normothermic controls.

In the study, the CSF pressure following injury was noted to rise slowly after the first hour to a peak by about the fourth hour, whereupon it remained relatively stable for up to twelve hours. Quite paradoxically, CSF pressure measurements did not reflect the continuing sizeable increase in brain volume due to cerebral oedema that occurred

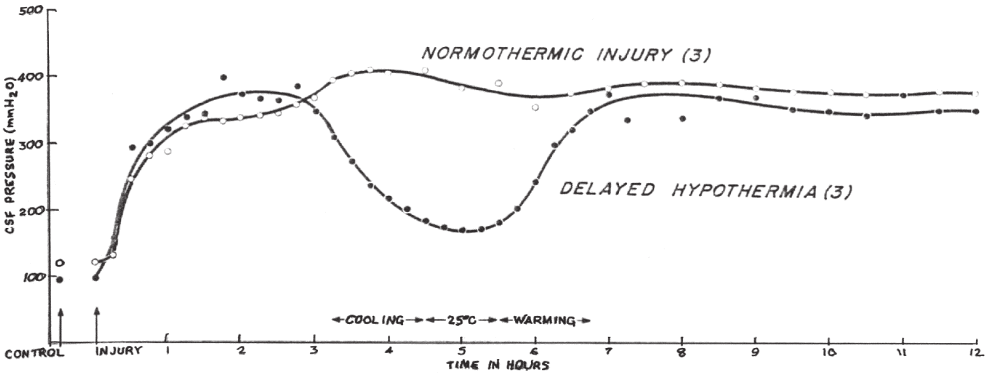


**Figure 2.** Cerebrospinal fluid pressure after brain injury at normal body temperature and with CSF Pressure after TBI at normothermia and during delayed hypothermia. Redrawn from data in the original publication.

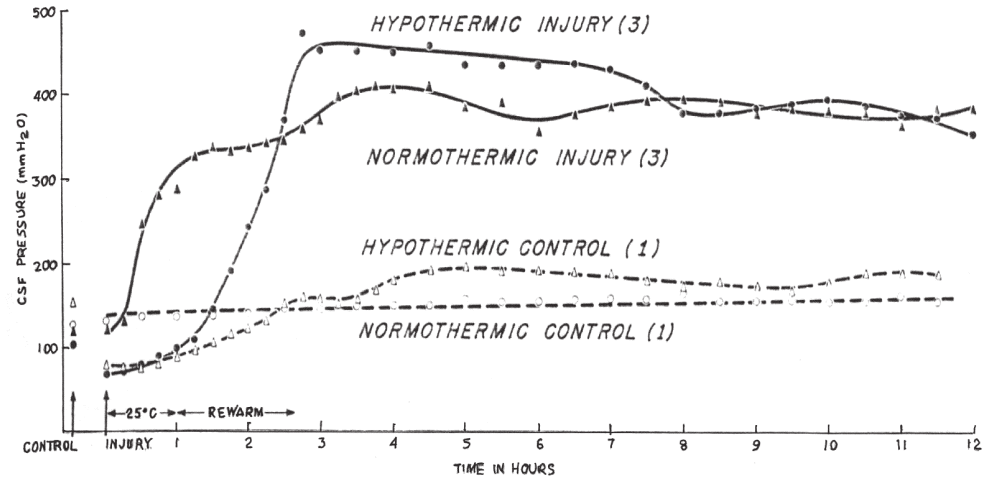


**Figure 3.** CSF Pressure at normothermia and hypothermia, with standard brain injuries and controls. Redrawn from the data in the original publication.

between the fourth and twelfth hour. Furthermore, there was no correlation between the level of CSF pressure and mortality. It was noted that the CSF pressure was reduced during the induction and maintenance of hypothermia in these experiments. With rewarming, the pressure reverted to levels established for normothermic controls, seemingly affected only temporarily by the reduction in temperature. This suggested that the CSF pressure alone is not an indication of prognosis, since there was no correlation between CSF pressure measurements and mortality.

In a similar study by Civalero et al[24] on dogs, the temperature in various organs was recorded during internal cooling. The true brain temperature was higher during cooling than the mean recorded temperature in other organs of the body. Brain

temperature was found to be 8-10°C (up to max of 14°C) above the recorded oesophageal temperature. In addition, the oxygen consumption during rewarming was greater than prior to cooling when the temperature was 37°C. These experimental observations are not only important but are also physiologically consistent with the hypothalamic thermoregulatory response increasing its neuro-humoural output in an attempt to re-establish normothermia.[25]

Earlier animal studies showed a reduction in the rate of cerebral oedema formation and mortality after injury to the cerebral cortex.[26]

**Adverse effects of hypothermia**

Therapeutic hypothermia can be associated with significant adverse effects.[27]

*CNS*

Hypothermia results in progressive reduction of neuronal function, which may manifest as confusion, disorientation, stupor and coma. This reduction results in changes in electrophysiological monitoring. The EEG shifts to slower frequencies while the somatosensory evoked potentials (SSEP) shows prolongation of latencies and a reduction in amplitude.

*Respiratory, Shivering and Oxygen Consumption*

Hypothermia causes a transient increase in ventilation before the more characteristic depression. Both respiratory rate and tidal volume are reduced; this appears top be due to a central CNS effect. Experimentally selective rewarming of the brain stem reverses these respiratory effects. Shivering is the most widely recognised side effect of hypothermia. More recent carefully controlled studies suggest that total body oxygen consumption increases by 40-100% during shivering. Such increases in oxygen consumption could have adverse effects on organs such as the heart or the brain, that may have fixed vascular obstructions to the arterial blood flow.

*Cardiovascular System*

Hypothermia results in a substantial activation of the sympathetic nervous system as is evident by an increase in circulating noradrenaline and associated vasoconstriction. The sympathetic effects may cause myocardial ischaemia both through peripheral and coronary vasoconstriction. The cardiac conduction system is cold sensitive and hypothermia causes bradycardia, prolonged PR intervals, widening of the QRS complex and prolongation of the QT interval resulting in the typical “J-wave”. Atrial fibrillation is common below 32°C.

*Renal*

There is progressive depression of renal tubular function with a marked reduction in the tubular reabsorptive capacity that results in “cold diuresis”. Sympathetic neuronal stimulation with cutaneous and splanchnic vasoconstriction also contributes to diuresis.

*Immune System and Infection*

Experimentally induced hypothermia reduces the function of neutrophils, lymphocytes, and macrophages. This has been shown to be associated with a high incidence of wound infections and pneumonia.

*Haemotological effects*

There is an increase in blood viscosity and a marked defect in coagulation parameters and platelet function that results in increased bleeding diathesis.

*Altered Pharmocokinetics and Pharmocodynamics*

Hypothermia potentiates the effects of central nervous system depressants and prolongs the duration of action of drugs dependent on enzymatic systems for their clearance, eg the duration of action of non-depolarising neuromuscular blockers is prolonged.

*Electrolyte abnormalities*

Hypophosphataemia, hypomagnesaemia, hyper- and hypo-calcaemia, and other electrolyte abnormalities have been reported during therapeutic hypothermia.[28]

**Hyperthermia in patients with TBI**

Fever is common in critically ill patients with neurotrauma, especially those patients with a prolonged length of stay in the ICU. Hyperthermia is associated with adverse outcomes in patients with TBI. It is therefore common practice for body temperature to be reduced using antipyretic medications (paracetamol) and external body cooling when the body temperature is elevated beyond 38.5°C in these patients, at least so as to maintain normothermia.[29]

**Ethical Considerations**

In his editorial discussing the multi-centre randomised study by Clifton et al,[10] Narayan[30] considers why so much laboratory and earlier clinical data struck an "optimistic note in favour of therapeutic hypothermia", while the Clifton study revealed no benefits. He noted that 38% of the patients involved in this study were enrolled without consent. Arguments in favour of consent waiver included the inability of severely head injured patients to consent themselves and relatives being unavailable within the short timeframe (less than 6 hours) required to institute treatment.

It can be ethically argued that further studies are required, with the consent of relatives or next-of-kin, to identify categories of patients who would benefit from the use of therapeutic hypothermia in traumatic brain injury. Does a single negative study mean that therapeutic hypothermia should not be used clinically? Perhaps not. The Clifton study found that hypothermia may be beneficial in a certain proportion of patients below the age of 45 years, with intractable high ICP. Two separate non-randomised studies by Jiang[8] and Shiozaki[9] support this observation.

Therapeutic hypothermia for TBI cannot be regarded as standard treatment and protocols for its use must incorporate the uncertainties associated with its routine application. Enthusiastic clinicians, who wish to institute therapeutic hypothermia for severe TBI as a life-saving measure, should obtain informed consent from a responsible person on behalf of the patient. The Informed Consent information sheet and the treating clinician should clearly explain that therapeutic hypothermia is a life-threatening intervention used in an attempt to save the patient's life, despite the fact that a medical publication has failed to confirm the benefit in a clinical trial. Such Informed Consent procedures would provide safeguards for clinicians and institutions against clinical and legal criticisms if, and when, the patient develops disabilities or death, which may be attributable to the therapeutic hypothermia. Conservative

clinicians must await further studies to determine the sub-population of patients most likely to benefit from the treatment.

**Future considerations**

Any future studies must specify the population of patients to be studied and to compare them with standard guideline directed treatments as control. Further, it would be necessary to define the duration and the extent of the therapeutic hypothermia, the target starting and definitive temperature endpoints. New protocols should be produced with firm indications and exclusions for its use. The treatment time frame must be specified and a rigorous schedule for cooling and rewarming (duration of hypothermia treatment and target rate of rise of temperature). Since this treatment cannot be administered in a double blind fashion, enthusiasm for its use should be tempered by conservative analysis. Protocols must have a clear understanding of the outcomes and serious adverse advents, including death, must be recorded and published. The study must enrol a sufficiently large number of patients to detect the possibility of a statistically significant difference in treatment.

The reader is reminded that this critique pertains to the use of therapeutic hypothermia for TBI. This review specifically excludes analysis of the publications on the use of therapeutic hypothermia in other clinical circumstances during anaesthesia, resuscitation and critical care.

**References**

1. The Management and Prognosis of Severe Traumatic Brain Injury. A Joint Project of the Brain Trauma Foundation, (Improving the Outcome of Brain Trauma Nationwide) and the American Association of Neurological Surgeons (ISBN 0-9703144-0-X) 1995, 2000.
2. Maas AIR, Dearden M et al. Current recommendations for Neurotrauma. *Current Opinions in Critical Care* 2000, 6:281-292.
3. Roberts I, Schierhout G, Alderson P. Absence of evidence for effectiveness of five interventions routinely used in Intensive Care management of severe head injury: A systematic review. *J Neurol Neurosurg Psychiatry* 1998; 65:729-733.
4. Cochrane Library (Database on disk and CD-Rom) Oxford: The Chochrane Collaboration, 1998.
5. Corticosteroid randomisation after significant head injury (CRASH Protocol) — Coordinating Centre. London School of Hygiene and Tropical Medicine, 49-51 Bedford Square, London UK WC1B 3 DP. Email CRASH@ishtm.ac.uk www.CRASH.ishtm.ac.uk
6. Clifton GL, Allen S et al A Phase II study of moderate hypothermia in severe brain injury. *J Neurotrauma* 1993 10(3): 263-271. Ed. comment: Muizelaar JP; p. 273.
7. Marion DW, Obrist WD et al. The use of moderate therapeutic hypothermia for patients with severe head injuries: A preliminary report. *J Neuro Surg* 1993; 79:354-362.
8. Jiang J-Y, Yu M-K et al. Effect of long-term mild hypothermia therapy in patients with severe traumatic brain injury: One-year follow-up Review of 87 cases. *J Neurosurg* 2000; 93:546-549.
9. Shiozaki T et al. A multi-centre prospective randomised controlled trial of the efficacy of mild hypothermia for severely head injured patients with low intracranial pressure. *J Neurosurg* 2001; 94:50-54
10. Clifton GL, Emmy RM et al. *N Engl J Med* 2001; 344:556-563.
11. Ganong WF. Review of Medical Physiology. 10th Edition. p. 193.
12. Mellegard P, Nordstrom CH. Epidural temperature and possible intracerebral temperature gradients in man. *Br J Neurosurg* 1990; 4:31-38.
13. Schwab S, Schwarz S et al. Moderate Hypothermia and brain temperature in patients with severe middle cerebral artery infarction. *Acta Neurochirur Suppl (Wien)* 1998; 71:131-134.
14. Hofstadter DR, Dennett DC. The Mind's I, BasicBooks P6, 1981. ISBN 014 00.6253X.
15. Little DM. Hypothermia — review article. *Anaesthesiology* 1959; 20:842-877.
16. Smith LW, Fay T. Observations on human being with cancer, maintained at a reduced temperature of 75-90°F. *Am J Clin Pathol* 1940; 10:1-11.
17. Fay T. *J Neurosurg* 1959; 16:239-260.

18. Lifton Robert Jay. The Nazi Doctors, Medical Killing and the Psychology of Genocide. BasicBooks 1986. ISBN 0-465-04905-2.
19. Hartung J, Cottrell JE. Editorial. Statistics and hypothermia. *J Neurosurg Anaesthesiol* 1998; 10(1):1-4.
20. Bigelow WG, Callaghan JC et al. Generalised hypothermia for experimental intracardiac surgery. Use of electrophrenic respirations, artificial pacemaker for cardiac standstill, and radio frequency rewarming in generalised hypothermia. *Ann Surg* 1950; 132:531-543.
21. Clifton Gl, Hayes RL. Hypothermia for the treatment of head injury. In Bk. Neurotrauma. Eds. Narayan RK, Wilberger JE Jr) Povlishock JT.1995 Pub. McGraw Hill. Chapter 29; 393-412.
22. Rosomoff HL, Gilbert R. Brain volume and cerebral fluid pressure during hypothermia. *Am J Physiol* 1955; 183:19-22
23. Rosomoff HL, Shulman K et al. Experimental brain injury and delayed hypothermia. *Surg Gyn Obst* 1960; 110:27-32.
24. Civalero L A, Moreno JR et al. Temperature conditions and oxygen consumption during deep hypothermia. *Acta Chir Scand* 1962; 123:179-188.
25. Laskowski EJ, Klatzo I et al. Experimental study of the effects of Hypothermia on local brain injury. *Neurology* 1960; 10:499-505.
26. Clasen RA, Pandolfi S et al. Hypothermia and hypotension in experimental cerebral oedema. *Arch Neurol* 1968; 19:472-486.
27. Gelb AW. Side effects of hypothermia. Proceedings of the Second International Symposium on therapeutic hypothermia: April 30-May 2, 1999. Editors Spiss CK, Bacher A et al. University of Vienna, Austria.
28. Polderman KH et al., Hypophosphaetemia and hypomagnasesaemia induced by cooling in patients with severe head injury. *J Neurosurg* 2001; 94:697-705.
29. Kilpatrick MN, Lowry DW et al. Hyperthermia in the Neurosurgical Intensive Care Unit. *Neurosurg* 2000; 47:850-856.
30. Narayan RK. Editorial comment. *N Eng J Med* 2001; 34:602-603.

# Bacterial Cognition and Consensual Pathogenesis: The Microbiology of Intensive Care

JON IREDELL, MB, BS, PhD, FRACP(Infect Dis), FRCPA(Microbiology)
Centre for Infectious Diseases and Microbiology, and
Department of Intensive Care, Westmead Hospital, University of Sydney

*Jon Iredell trained in intensive care, infectious diseases and microbiology in Brisbane and Adelaide, completing his PhD at the University of Adelaide and the Max Planck Institute for Microbiology, Tuebingen, Germany. He was appointed as staff specialist in the Centre for Infectious Diseases and Microbiology and as a general intensivist at Westmead in 1998. His research interests relate to bacterial pathogenesis and horizontal gene transfer in resistant Gram-negative bacteria.*

Most Intensive Care Units have a close working relationship with the Clinical Microbiologist or Infectious Diseases specialist. This is largely driven by the provision and interpretation of culture-based data from relevant clinical samples, as well as the difficulty and complexity of infection management in the critically ill. It is therefore somewhat disconcerting that much of the routine diagnostic yield from the microbiology laboratory is perceived as unhelpful.[1] While information content and delivery clearly can be improved, new understandings in bacteriology may help us to better manage the host-pathogen relationship. In this review, we revisit some traditional paradigms and look at new approaches to Intensive Care microbiology.

The human host environment varies enormously, and is subject to constant monitoring in the critically ill patient. Extracellular fluid distribution, protein binding, tissue perfusion and changing drug metabolism and elimination during shock and resuscitation are familiar variables. Local factors such as antibiotic penetration and activity in abscess cavities or ischaemic tissues are also factors we must consider when planning antibiotic therapy for sepsis. We will not deal with these topics further, but they are important to contextualise the discussion which follows, since changes in the host environment directly affect the behaviour of the invading pathogen and its response to therapy.

An anthropocentric view of microbiology has served us well for many years. We have traditionally defined medically important bacteria on the basis of criteria articulated by Robert Koch in the late nineteenth century for tuberculosis,[2] and the introduction of specific therapies against agents such as *Mycobacterium tuberculosis* and *Streptococcus pneumoniae* are milestones in modern medicine. In general terms, Koch's postulates for defining the causative link of organism to disease state that the organism must be present in every case of the disease, and be isolated in pure culture from infected patients. The pure isolate must cause the disease when introduced into a new host, and be isolated again from that secondary host. The concept of the bacterial pathogen as a single species invader, armed with toxins and specific resistances to antibiotics and/ or immune defences such as phagocytes, has been

further modified to acknowledge that we cannot isolate every relevant pathogen in cell-free culture and that many organisms are diagnosable now only by culture-independent methods.[3]

**Environmental sensing from the periplasmic space (Figure 1)**

The Gram-negative/Gram-positive divide reflects basic physical differences in bacteria. In general terms, the envelope of the Gram-negative organism (e.g. "enterics" such as *E.coli* and *Klebsiella spp*., and "non-fermenters" such as *P. aeruginosa* and *Acinetobacter spp*.) is a much more sophisticated structure than that of the Gram-positive organism (e.g. *S. aureus*, *S. pneumoniae*). The Gram-positive envelope, designed to withstand physical stress, dessication and environmental insult, simply comprises a cytoplasmic membrane surrounded by a thick layer of peptidoglycan, giving physical strength to a typically spherical organism. The Gram-negative envelope is dual, with a unique hydrophobic outer membrane. Transduction of energy to this is difficult and access through it is controlled by membrane channels (porins). These porins may allow facilitated diffusion for desirable nutrients or act as simple aqueous channels of relatively fixed diameter and charge. The expression and assembly of these channels are actively regulated, to alter overall "porosity" of the outer membrane, which in certain Gram-negative organisms such as *Acinetobacter spp*. (relatively tolerant of drying) may be several logs less than that of, say, *E. coli*. This allows rapid downregulation of porins to confer extreme antibiotic resistance at minimal biological cost,[4] and is also reflected in greater tolerance of drying and, therefore, enhanced survival on environmental surfaces and fomites.

The other practical consequence is the tiny but critically important area between the inner and outer membranes: the periplasmic space. This is protected by the hydrophobic Gram-negative outer membrane, and houses a number of important structures which give the Gram-negative bacteria unique advantages. A primary example is the array of delicate two-component sensor-regulator systems which "sense" the environment. Conformational change in cognate sensor proteins, upon binding relevant substrate, is transduced to the paired regulatory protein responsible for adjusting DNA transcription of messenger molecules in the cell. The sensed molecule thus triggers a relevant reaction or cascade of responses. The periplasm also provides a sheltered space for assembly of specialised structures such as protein adhesins, while enzymes such as beta-lactamases need not be highly active nor produced in large amounts to be effective, but can be contained in high concentration where they are most needed.

The innner (cytoplasmic) membrane is an active structure, easily energised and supplied with amino acids, and is the site for macromolecular assemblies such as transporter pumps and flagellar motors. Environmental sensing from the periplasm permits chemotaxis down specific concentration gradients, and regulatory systems facilitate adjustment to a new environment. *V. cholerae*, for example, senses certain specific sugars and amino acids, pH and temperature and, through a cascade of regulatory networks, switches off motility and switches on adhesins and toxin formation when the time is right.[5] This allows the organism to be vigorously motile in the mucus layer of the colon and the rice water stool shed into the environment, but to grow in adherent microcolonies utilising specific adhesins when in the gut. Tangled pilin bundles around the flagellar shaft illustrate the incompatibility of these systems when the organism is forced *in vitro* to simultaneously express both adhesion and

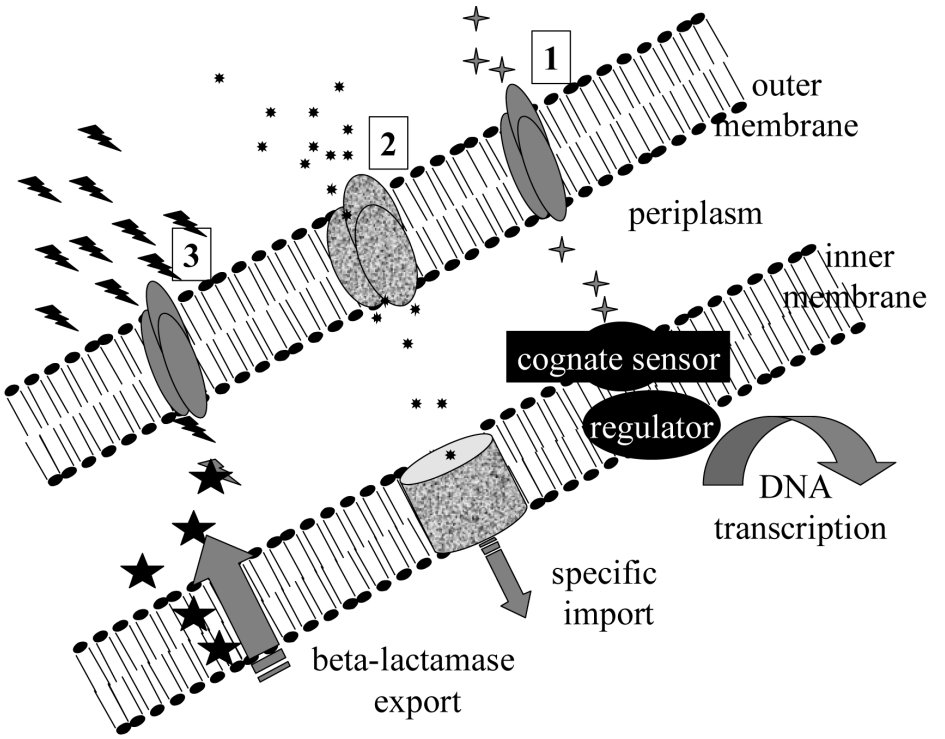


**Figure 1.** The periplasmic space. Sensing of the environment within the periplasm is by interaction with cognate sensor/s (1), which mediate DNA transcription via their paired regulatory protein. Specific porins may facilitate diffusion of desirable substances (2) into the periplasmic space for importation. Beta-lactamases are exported into the periplasm (3), to attack antibiotics which have got through the porin channels.

motility phenotypes (Figure 2). Bacteria are thus continually niche-adapting in real time, with a highly flexible repertoire of responses often poorly reflected by the *in vitro* phenotype reported from the diagnostic Microbiology laboratory.

**Cognition and consensual pathogenesis: quorum sensing in bacteria**

The idea of bacteria as cognisant beings is fundamental to current concepts of bacterial pathogenesis. A truly cognisant being is not only able to sense its environment and adapt immediately, but is able to communicate with peers to coordinate action to mutual advantage. In bacteria this is referred to as quorum sensing.[6] Small secreted molecules send messages between bacteria, allowing them to co-ordinate responses in newly arriving cells. In its original form, this local signalling was recognised as population-density sensing, and it is now known to be widespread in eubacteria. These signals are complex and variable, and may also be recognised by other organisms. Signals may even be used as disinformation networks, to alter behaviour of arriving competing species so as to suit the incumbents. Many systems coexist in a single organism and coordinate sophisticated responses to like and competing organisms in real time.[5,6] In our cholera example, the bacillus is vigorously motile until it reaches the

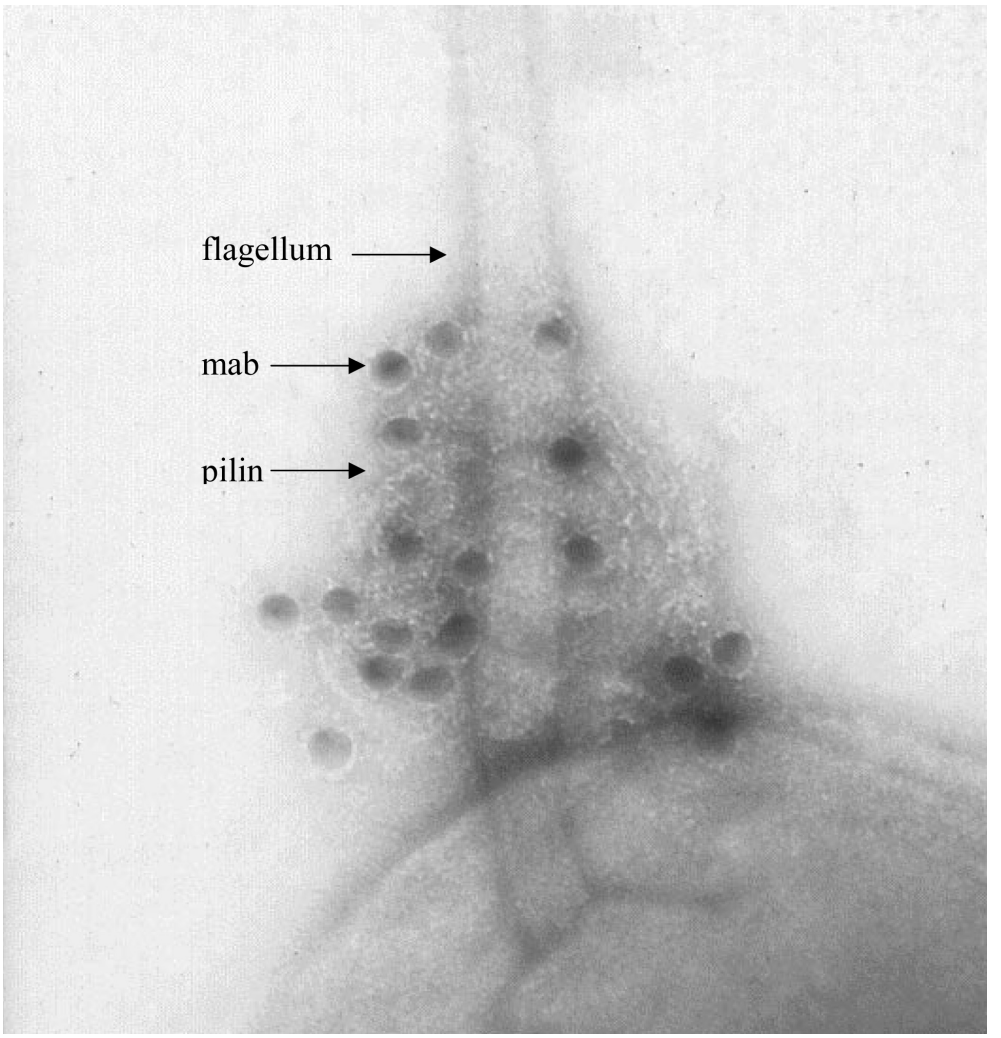


**Figure 2.** 10 nm Au particle-labelled monoclonal antibodies (mab) can be seen decorating adhesin fibres (pilin) tangled around the polar flagellum of a cholera bacillus forced to simultaneously express both adhesion and motility phenotypes in vitro.

mucosa, where it stops and develops a microcolony. Once high population density is sensed, expression of cholera toxin causes the gut to flush the amplified microcolonies out to begin the cycle again.[7]

**Antibiotic sensitivity is context-specific**

Bacterial populations thus wax and wane in relatively short growth cycles. Clinically apparent bacterial infection can be visualised as a continuum from rapid replication to extinction and it is helpful to think in terms of the two extremes of the growth curve. An organism which is not actively growing and dividing is less susceptible to antibiotics, which characteristically target the machinery on which these processes are dependent.

The minimal inhibitory concentration (MIC) is the antibiotic concentration required to prevent exponential growth of the organism in question. It may be two or three logs higher for some drugs (e.g. aminoglycosides) in the stringent or stationary phase, than that which is derived in the diagnostic laboratory for bacteria in optimal growth conditions. It becomes easier to understand why early bacteraemic Staphylococcal endocarditis responds effectively to a combination of beta-lactam (cell wall-active) and aminoglycoside (protein synthesis inhibition) therapy for a few days, with rapid sterilisation of the bloodstream,[8] but why weeks of therapy with a cell-wall active agent are subsequently needed to prevent relapse (Figure 3).

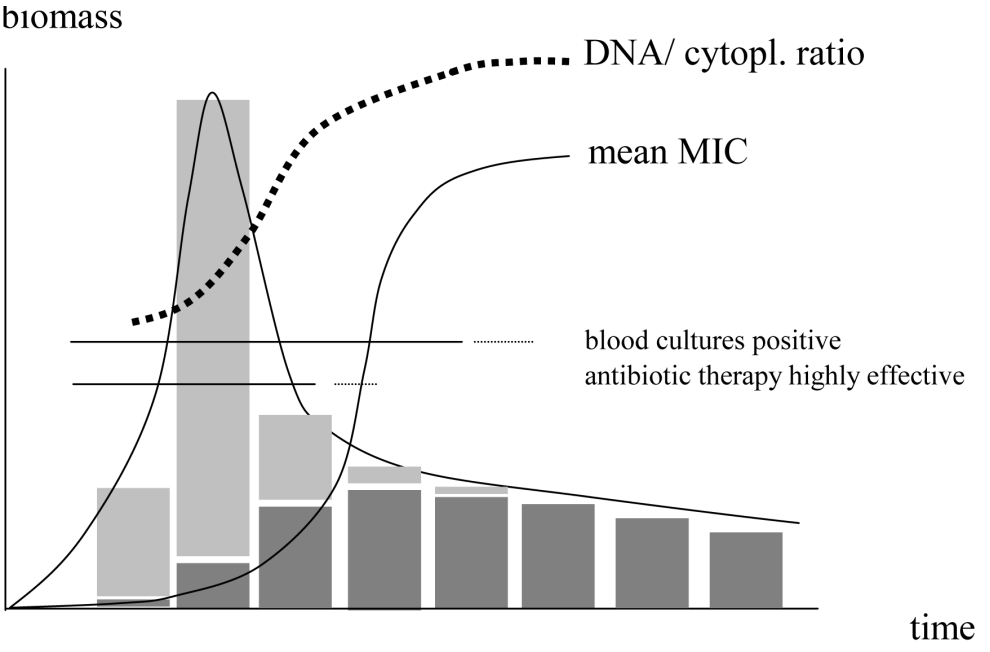


Figure 3. Hypothetical schema of the transition from vegetative growth to stringent state, with impact upon blood culture positivity and antibiotic susceptibility. During initial rapid replication, vegetative organisms (light grey) spread through the bloodstream and are readily cultured. The sensitive (vegetative) biomass quickly dies, and stringent state organisms (dark grey) become the dominant surviving biomass. Stringent state organisms have higher DNA/ cytoplasmic ratios, reflecting reductive division to preserve DNA mass at the expense of cytoplasmic activity (and thus, antibiotic sensitivity).

Likewise, infections of prosthetic material or the poorly vascularised long bone of an adult with chronic osteomyelitis is largely a stringent-state infection. This term means, as it implies, an adaptation to stressful and/or nutrient-limited environments. Reductive division to conserve DNA at the expense of cytoplasmic material and activity is driven by similar processes in Gram-positive and Gram-negative bacteria.[9] At its extreme, this can be seen as the sporulation response in organisms with that capacity (e.g. Clostridia) and the viable non-cultivable state in Gram-negative organisms, which is a normal response to physiologic stress.[10] In the stressed or stringent, most of the organisms in the population are very slowly growing and slowly dividing and so

relatively insensitive to growth and replication inhibitors. Bacteria have evolved to live on the brink of survival, and to launch into explosive growth (and a frenzy of genetic exchange, in many cases) when the opportunity presents itself. We see this at the bedside in an overwhelming infection, as clinical sepsis, which is exquisitely sensitive to appropriately targeted growth and replication inhibitors (antibiotics). Reversion to "ticking-over" (and less antibiotic-sensitive) growth is actually the more usual state.

Importantly, the common site for this stressed- or stringent-state growth is on surfaces which provide moisture and warmth but little nutrients. In a clinical environment, this can be the endotracheal tube, intravesical and intravascular catheters and so on. This helps to explain why growth of a microorganism from a prosthesis or catheter often cannot be eradicated by the antibiotic to which it seems exquisitely sensitive in vitro.[11] A dependent patient in the operating theatre or the Intensive Care Unit has a number of such ecological niches which can be opportunistically colonised, with the formation of a biofilm. Surgical and supportive care devices not only seed infection, as they are usually simultaneously breaching anatomical defences, but provide a reservoir of organisms which may be subjected to repeated or ongoing antibiotic selection pressure and are often competent for genetic exchange. Close similarities between the type of adhesins needed for biofilm formation[12] and those associated with specific DNA uptake[13] are probably no coincidence.

The biofilm growth characteristic of inert surface infection (e.g. catheter infection) is a normal adaptation for many Gram-negative bacteria and, when conditions are optimal, may be logarithmic and therefore quite antibiotic-sensitive. However, the biofilm is a loose three-dimensional structure with aqueous channels designed to optimise nutrient access in a nutrient-poor environment; it is an efficient system for stressed organisms and often supports stringent growth.[11] Secondly, many biofilms are associated with exopolysaccharide secretion, producing antiphagocytic barriers which protect the organisms directly (e.g. *Strep. viridans* in endocarditis). Finally, biofilms may form along with adherent fibrin or blood clot (e.g. endovascular catheter lumens), making antibiotic delivery to the organism more difficult again. All of these factors contribute to the difficulty of eradicating biofilm infections in the clinical environment.

**Reconsidering the species concept in bacteria: the floating genome**

It is also important to remember that bacterial genomes are uniquely fluid. Small subunit (16S) ribosomal (rDNA) sequences have long been considered the most significant and reliable genotypic basis for bacterial speciation. However, this belief is undermined by work suggesting a role for horizontal gene exchange in generating intra-species diversity in bacteria,[14, 15] and by recognition of the importance of mobile genetic elements such as gene cassettes in (especially Gram-negative) bacteria. This complexity has led to calls for complete review of the notion of species, which has translated poorly from Mendelian concepts of eukaryotic evolution.[16] The bacterial genome is both uniquely tolerant of extrachromosomal material, and uniquely flexible in genomic structure and content. A bacterial genome of between two and six million base pairs is not only several logs larger than a simple viral genome (e.g. around ten thousand base pairs for HIV or about twice that size for human coronavirus), but may house literally as many as hundreds of copies of extrachromosomal plasmids with more genetic information than the average virus in each plasmid. In addition to this huge increase in complexity, bacteria have a unique ability to exchange these extra-

chromosomal vehicles, to pick up DNA from outside the cell, and to integrate and excise DNA for biological advantage. This DNA pool is the "floating genome",[17] the smallest common component of which is the integron, a gene capture and management system which is apparently unique to and ubiquitous in bacteria.[18, 19]

The floating genome can be considered as a genetic pool shared between bacteria, flux within which is determined by the transferability of the DNA, and the competence of various bacteria to receive it. Some of this genetic material is specifically packaged in plasmids or bacteriophages, which use a variety of mechanisms for entry, replication, and transmission. Their ability to move from one bacterium to another is termed the "host range". Some vehicles (e.g. bacteriophages) may be extremely specific and confer unique virulence characteristics,[20] while others (e.g. "broad host-range" plasmids) may be highly promiscuous and can be picked up and transmitted by distantly related bacteria.

Bacteria which are able to survive antibiotic attack are usually swimming in a soup of DNA released by lysis of susceptible organisms, and thus exposed to potentially useful genetic material if competent to take it up. The highly dependent surgical or ICU patient usually has a greatly increased burden of microbes in the oropharynx, upper respiratory tract, wound sites and drains, monitoring devices etc, and the intensity and complexity of microbial interactions are greatly increased over that ordinarily seen in a healthy individual with all anatomical defences (such as skin, airway, urogenital tract) intact.

**Getting it right: the essential tension in prescribing**

The necessity to get it right first time in sepsis is well established, with worse outcomes in bacteraemia and ventilator-associated pneumonia as a result of delays in appropriate antibiotic therapy.[21-23] Empiric broad spectrum antibiotics is the mainstay of therapy in the critically ill patient with suspected infection,[1] since only a minority of prescriptions are informed by a timely and relevant microbiological diagnosis.[24-26] Antibiotic usage is well documented to be associated with emergence of specific resistance,[27-30] and reduced usage of an antibiotic in the ICU setting has been shown to be associated with reduced resistance to that antibiotic in subsequently cultured isolates over a period of months.[30] Attempts to "cycle" antibiotics in an effort to reduce selection for resistance have met with some success,[30, 31] but only cautious acceptance as yet.[32, 33] Part of this caution reflects our imperfect understanding of the processes underlying resistance development and transmission, and the failure of traditional teaching in microbiology to provide us with more relevant paradigms.

Most Australian intensivists state that they need to cover 95% or more of the possible pathogens in critically ill patients with serious infection.[1] Protocol-driven prescribing is thus widely adopted as the primary principle, based on local and regional data about appropriate targeting (organisms, sensitivities). The competent prescriber for critically ill patients is therefore always trying to reconcile an inherent conflict between the need to "get it right first time" and the need to minimise the selection for resistance. There is a need for better quality information at the bedside if we are to manage this paradox more successfully in the future.

The majority (typically, around 70%) of Intensive Care Unit patients are on broad spectrum beta-lactam antibiotics such as cephalosporins and antipseudomonal penicillins,[24] often in combination with aminoglycosides and other classes. Antibiotic susceptibility profiles are often quite specific to a Unit, even within the same hospital

department.[35] When surveyed directly, most intensivists in Australia agree that antibiotic resistance is a problem and that local profiles are essential, but less than 40% state that they actually use local laboratory data to guide their decision-making. A significant percentage find "expert" advice from Infectious Diseases and Microbiology colleagues to be completely unhelpful.[1] Decision support systems are of proven value, with objective criteria applied in the scoring of ventilator-associated pneumonia having been shown to be effective in reducing antibiotic use without compromising efficacy.[35]

**Where do we go from here?**

Thus while the use of bedside information systems is increasingly popular, it is heavily dependent on the information quality itself, and the models we use to interpret the data. It may be that we should be looking at different data altogether. If much of what we do in Intensive Care is empiric, and protocols need to be based on resistance potential and selection pressures, then perhaps we need to be using existing molecular techniques as Infection Control tools. Paired with a knowledge of the vehicles of genetic flux (the plasmids and transposons, and even the bacteria themselves), twice-weekly surveillance of representative sites such as the bronchial tree (e.g. via non-bronchoscopic BAL),[36] and perineum, wounds, drains and "usual" sites (sterile sites such as blood and CSF) when indicated, augmented by molecular tools to examine mobile resistance elements, should provide a relevant local census of the microflora and its resistance potential. Ultimately, advances such as real-time PCR and micro-array (gene chip) technology, might even allow relevant patient-specific studies. Improved understandings of bacteria and how they respond to selection pressure and niche modification in the clinical context should promote better husbandry, not only of the microflora in an individual patient but the entire ecosystem in which we all participate.

**References**

1. Sintchenko V, Iredell JR, Gilbert GL, E. Coiera. What do physicians think about evidence-based antibiotic use in critical care? A survey of Australian intensivists and infectious disease practitioners. *Intern Med J* 2001; 31:462-469.
2. Koch R. Die Aetiologie der Tuberkulose. *Mittbeihgen aus dem Kaiserlichen Gesundbeitsamte* Vol. 2, 1884;1-88.
3. Fredericks D, Relman D. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin Microbiol Rev* 1996; 9:18-33.
4. Limansky AS, Mussi MA, Viale AM. Loss of a 29-kilodalton outer membrane protein in Acinetobacter baumannii is associated with imipenem resistance. *J Clin Microbiol* 2002; 40:4776-4778.
5. Gardel CI, Mekalanos JJ. Alterations in Vibrio cholerae motility phenotypes correlate with changes in virulence factor expression. *Infect Immun* 1996; 64:2246-2255.
6. Bassler BL. Small talk. Cell-to-cell communication in bacteria. *Cell* 2002; 109:421-424.
7. Zhu J, Miller MB, Vance RE et al. Quorum-sensing regulators control virulence gene expression in Vibrio cholerae. *Proc Nat Acad Sci USA* 2002; 99:3129-3134.
8. Mylonakis E, Calderwood S. Infective endocarditis in adults. *N Engl J Med* 2001; 345:1318-1330.
9. Lewis K. Programmed death in bacteria. *Microbiol Mol Biol Rev* 2000; 64:503-514.
10. Defives C, Guyard S, Oulare M et al. Total counts, culturable and viable, and non-culturable microflora of a French mineral water: a case study. *J Appl Microbiol* 1999; 86:1033-1038.
11. Prince AS. Biofilms, antimicrobial resistance, and airway infection. *N Engl J Med* 2002; 347:1110-1111.
12. Watnick P, Kolter R. Biofilm, city of microbes. *J Bacteriol* 2000; 182:2675-2679.
13. Iredell JR, Manning PA. The toxin-co-regulated pilus of Vibrio cholerae O1: a model for type 4 pilus biogenesis? *Trends Microbiol* 1994; 2:187-192.
14. Dillon B, Valenzuela V, Don R et al. Limited diversity among human isolates of Bartonella henselae. *J Clin Microbiol* 2002; 40:4691-4699.

15. Kersulyte D, Chalkauskas H, Berg DE. Emergence of recombinant strains of Helicobacter pylori during human infection. *Mol Microbiol* 1999; 31:31-44.
16. De la Cruz F, Davies J. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* 2000; 8:128-133.
17. Liebert CA., Hall RM, Summers AO. Transposon Tn21, flagship of the floating genome. *Microbiol Mol Biol Rev* 1999; 63:507-522.
18. Collis CM, Grammaticopoulos G, Briton J, Stokes HW, Hall RM. Site-specific insertion of gene cassettes into integrons. *Mol Microbiol* 1993; 9:41-52.
19. Recchia GD, Hall RM. Origins of the mobile gene cassettes found in integrons. *Trends Microbiol* 1997; 5:389-394.
20. Karaolis DKRS, Somara S, Maneval DR Jr et al. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* 1999; 399:375-379.
21. Ibrahim EH, Sherman G, Ward S et al. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. *Chest* 2000; 118:146-155.
22. Iregui M, Ward S, Sherman G et al. Clinical importance of delays in the initiation of appropriate antibiotic treatment for ventilator-associated pneumonia. *Chest* 2002; 122:262-268.
23. Heyland DK, Cook DJ, Griffith L et al. The attributable morbidity and mortality of ventilator-associated pneumonia in the critically ill patient. *Am J Respir Crit Care Med* 1999; 159:1249-1256.
24. Bellomo R, Bersten AD, Boots RJ et al. The use of antimicrobials in ten Australian and New Zealand intensive care units. The Australian and New Zealand Intensive Care Multicentre Studies Group Investigators. *Anaesth Intens Care* 1998; 26:648-653.
25. Luna C, Vujacich P, Niederman M et al. Impact of BAL data on the therapy and outcome of ventilator-associated pneumonia. *Chest* 1997; 111:676-685.
26. Luna CM, Videla A, Mattera J et al. Blood cultures have limited value in predicting severity of illness and as a diagnostic tool in ventilator-associated pneumonia. *Chest* 1999; 116:1075-1084.
27. Villers D, Espaze E, Coste-Burel M et al. Nosocomial Acinetobacter baumannii infections: microbiological and clinical epidemiology. *Ann Intern Med* 1998; 129:182-189.
28. Oh JY, Yu HS, Kim SK et al. Changes in patterns of antimicrobial susceptibility and integron carriage among Shigella sonnei isolates from southwestern Korea during epidemic periods. *J Clin Microbiol* 2003; 41:421-423.
29. Rahal JJ,Urban C, Horn D et al. Class restriction of cephalosporin use to control total cephalosporin resistance in nosocomial Klebsiella. *JAMA* 1998; 280:1233-1237.
30. Moss WJ, Beers MC, Johnson E et al. Pilot study of antibiotic cycling in a pediatric intensive care unit. *Crit Care Med* 2002; 30:1877-1882.
31. Kollef MH, Vlasnik J, Sharpless L, Pasque C, Murphy D, Fraser V. Scheduled change of antibiotic classes: a strategy to decrease the incidence of ventilator-associated pneumonia. *Am J Respir Crit Care Med* 1997; 156:1040-1048.
32. John JFJ, Rice LB. The microbial genetics of antibiotic cycling. *Infect Control Hosp Epidemiol* 2000; 21: S22-31.
33. Farr BM, Salgado CD, Karchmer TB et al. Can antibiotic-resistant nosocomial infections be controlled? *Lancet Infect Dis* 2001; 1:38-45.
34. O'Connell B, Marcus R, Broadbent V et al. Can a single antibiotic policy for the empirical treatment of febrile neutropenic patients be used for all categories of haematology/oncology patient in the same institution? *J Antimicrob Chemother* 1998; 42:677-678.
35. Singh N., Rogers P, Atwood CW et al. Short-course empiric antibiotic therapy for patients with pulmonary infiltrates in the Intensive Care Unit. A proposed solution for indiscriminate antibiotic prescription. *Am J Respir Crit Care Med* 2000; 162:505-511.
36. Arora SC, Mudaliar YM, Lee C, Mitchell D, Iredell J., Lazarus R. Non-bronchoscopic broncho-alveolarlavage in the microbiological diagnosis of pneumonia in mechanically ventilated patients. *Anaesth Intens Care* 2002; 30:11-20.

# Author's Index

# Key Words