from the neuromuscular blocking agent when the concentration of anaesthesia2. Consequent from this small reduction in vecuronium dose of vecuronium required for a given effect during sevoflurane an important clinical question. We established recently that the magnitude of potentiation of the dose requirements there was a much greater degree of recovery effects of neuromuscular blocking agents by volatile anaesthesia is shown that direct comparison of pH values obtained by different analyzers. Riddington and colleagues' and Takala and colleagues'1 have demonstrated any significant difference between our study and that of Camprubi and colleagues is a result of the use of different blood-gas analysers. It also cannot be excluded that the apparent difference in baseline pH, values between our data and that of Vanlithout and colleagues' article, we were unable to strictly test the hypothesis that VVB would maintain mucosal blood flow as VVB was used routinely at the institution. Therefore, we cannot exclude the possibility that pH, values might have decreased to even lower values without VVB. The main purpose of our study was to assess the ability of tonometry to detect intraoperative mucosal hypoxia, to measure gastric and sigmoid pH, and to relate the observed changes in pH, to the occurrence of endotoxemia and primary graft function.

We interpret their study with caution. The results of their study are entirely in keeping with there being a clinically important difference in the degree of neuromuscular blocker potentiation manifest by sevoflurane and isoflurane.

M. WELTE
Institute of Anaesthésiology
University of Munich
Klinikum Großhadern
Munich, Germany

1. Welte M, Pichler B, Groh J, Anthuber M, Jauch K-W, Riddington M, Pratschke E, Lenhart FP, Haller M, Frey L, Peter K. Portal perfusion was maintained by the use of VVB. 'The main purpose of our study was to assess the ability of pH, values measured early during hepatectomy in our study (7.28 vs 7.32). In fact, low pH, values (>7.32) before the second postoperative day, an observation that further supports the presence of chronic impairment of intestinal perfusion in our patients with end-stage cirrhosis. This is supported by the lower pH, values measured early during hepatectomy in our study (7.28 vs 7.32). In fact, low pH, values (>7.32) before the second postoperative day, an observation that further supports the presence of chronic impairment of mucosal microcirculatory perfusion. Hence, pre-existing chronic mucosal hypoxia might explain why pH, decreased during the anhepatic stage, although overall portal flow was maintained by the use of VVB.

Sir,—Clinical interaction studies evaluating the influence of different anaesthetic techniques on the action of neuromuscular blocking agents often compare two or more groups of patients treated separately but concurrently as part of the same study. These studies are termed "parallel" to emphasize their difference from other clinical studies in which patients are their own controls. Parallel comparisons include most of the usual forms of interaction studies investigating the abilities of different inhalation anaesthetics to potentiate the neuromuscular effects of neuromuscular blocking agents. However, the large between-patient variability in dose-response and dose-duration studies with neuromuscular blocking agents may preclude demonstration of subtle differences in susceptibility to these agents. Paired crossover studies can minimize the effect of inter-individual variability1. However, these paired crossover studies are difficult to perform with surgical patients. Healthy ASA I or II patients that are anaesthetized two or more times, receiving on each occasion a different anaesthetic technique, are rare. Therefore, these paired crossover studies have to be performed with volunteers.

Our clinical study was designed as a parallel comparison of different anaesthetic techniques, that is, opioid-nitrous oxide—oxygen, opioid-nitrous oxide—oxygen—isoflurane and opioid-nitrous oxide—oxygen—sevoflurane, in their ability to potentiate the neuromuscular effects of the most commonly used non-depolarizing neuromuscular blocking agents in surgical patients. Using oneway analysis of variance (ANOVA) and subsequently the Ryan-Einot-Gabriel-Welch (REGW) multiple range test to identify eventual sources of difference, there was a significant difference in magnitude and duration of neuromuscular action between the control group, receiving no volatile agent, and the groups receiving inhalation anaesthetics. We were unable to demonstrate any significant difference between opioid—nitrous oxide—oxygen—isoflurane and opioid—nitrous oxide—oxygen—sevoflurane anaesthesia on the effects of neuromuscular blocking agents. In the population investigated, the differences between opioid—nitrous oxide—oxygen—isoflurane and opioid—nitrous oxide—oxygen—sevoflurane anaesthesia were less important than the differences between opioid—nitrous oxide—oxygen—isoflurane and opioid—nitrous oxide—oxygen—sevoflurane anaesthesia on the effects of neuromuscular blocking agents in the population investigated, the differences between opioid—nitrous oxide—oxygen—isoflurane and opioid—nitrous oxide—oxygen—sevoflurane anaesthesia on the effects of neuromuscular blocking agents were small compared with the between-patient differences in these responses to neuromuscular blocking agents (<10% for the mean values of both magnitude and duration of action). In order to demonstrate significant differences
in the potentiating abilities of isoflurane and sevoflurane on the effects of neuromuscular blocking agents, that is a significant difference in both depth and duration, sample sizes of several hundreds of patients per neuromuscular blocking group would have been required. Such sample sizes are beyond practicable limits. We selected our population carefully, using rigid inclusion criteria, as described in the text, to reduce patient variability. This inevitably introduced constraints such as the availability of large numbers of eligible patients within a reasonable period of time. Any investigator is limited by both resources and time. Additionally, our study was intended as a phase II trial to evaluate the interaction between sevoflurane and neuromuscular blocking agents within a limited number of surgical patients.

Both the investigator and clinician should be concerned about the ability to detect an important clinical difference. Different investigators disagree on what is a clinically significant difference. They also disagree on the risk they are willing to take of missing a meaningful effect caused by drug interaction. The choice of a 20% difference in ED₅₀ as a clinically significant value is arbitrary. Such a difference is small compared with the large inter-individual differences in the response to neuromuscular blocking agents. With the degree of variability and sample sizes used in our study, a 30% reduction in the ED₅₀ of pancuronium and atracurium and a 45% reduction in the ED₅₀ of vecuronium had an 80% chance of being significant at the 0.05 level.

Finally, we still feel that further interaction studies on the influence of higher concentrations of volatile anaesthetics on larger doses of neuromuscular blocking agents are warranted. To elucidate small but significant differences in the ability of isoflurane and sevoflurane to potentiate the action of neuromuscular blocking agents, these interaction studies need to be performed in volunteers who are studied twice, receiving isoflurane on one occasion and sevoflurane on the other.

L.E.H. VANLINTHOOT
T. DE BOO
L.H.D.J. BOU
Institute for Anaesthesiology
University Hospital of Nijmegen
Nijmegen, The Netherlands


Cusum: a statistical method to evaluate competence in practical procedures

Sir,—It was interesting to read the commentary by Kestin describing the use of cusum analysis to measure the competence of anaesthetic trainees at practical procedures. The cusum is a useful graphical tool for discerning trends from a series of observations. The derivation of boundary lines for sequential tests allows comparison of the observed proportions of success or failure against predetermined standard criteria. This could be developed into a continuous performance monitor in anaesthesia.

In his article, Kestin discussed the problems of keeping paper records and plotting fractions on the graphs. In response to this we set out to develop a computerized personal log book system for practical procedures to run in parallel with the electronic anaesthetic log book on the Psion 3a. However, in setting up the algorithms we have encountered a number of problems.

The values for s, h₀ and h₁ calculated using the formulae in the appendix do not agree with the values in table 1 of the article. On reviewing the original articles on the application of the cusum, the value of Q should read Q = ln[(1 – p₀)/(1 – p₁)]. The values in the table are indeed correct if this calculation is used for Q. The null hypothesis in the article is stated as “true failure rate is NOT different from the acceptable failure rate” and if the cusum exceeds h₁ then it is rejected. This does not imply that the true failure rate exceeds the unacceptable failure rate which is the performance indicator which interests us. Similarly the alternative hypothesis is stated as “the true failure rate is equal to or exceeds the unacceptable failure rate” and if the cusum decreases to less than h₀ then it is accepted. Surely this would imply that the failing cusum of registrar B in figure 1 has a true failure rate that is equal to or exceeds the unacceptable failure rate.

h₀ corresponds to the failure rate under the null hypothesis which should surely read “the true failure rate is not different from the unacceptable failure rate”. If the cusum exceeds h₁ then this hypothesis is accepted and the trainee’s performance is unacceptable with reference to the agreed unacceptable failure rate. Similarly, h₁ corresponds to the failure rate under the alternative hypothesis which should surely read “the true failure rate is equal to or exceeds the acceptable failure rate”. If the cusum decreases to less than h₀ then the alternative hypothesis is rejected and the trainee’s performance is no worse than the accepted failure rate.

Furthermore, there appear to be clearcut criteria in the definition of success or failure for a particular procedure. This is confusing and it is difficult to see why type 1 and 2 errors of 10% (0.1) where chosen.

For reasons of convenience for the plotting of the graph, the values of s, h₀ and h₁ where multiplied by 10. It is not clear if, in the event of a success (1–r) or failure (r) or if the cusum exceeds i·, then it is rejected. This does not imply that the true failure rate is NOT different from the acceptable failure rate which is the performance indicator which interests us. Similarly the alternative hypothesis is stated as “the true failure rate is equal to or exceeds the unacceptable failure rate” and if the cusum decreases to less than h₀ then it is accepted. Surely this would imply that the failing cusum of registrar B in figure 1 has a true failure rate that is equal to or exceeds the unacceptable failure rate.

The statistical method described by Kestin appears to be a very powerful analytical tool and may have a wide range of applications in the assessment of trainees. However, the errors and inconsistencies in the appendix are a source of confusion. It would be useful if the derivation of the definitions, variables, hypotheses and theory behind the calculations were explained more clearly.

E. J. HAMMOND
Department of Anaesthesia
Poole General Hospital
Poole
A. K. McINDOE
Royal Devon and Exeter NHS Trust


Sir,—I would like to thank Drs Hammond and McIndoe for their interest, and in particular for correcting the error in the formula for Q in the appendix—p₀ and p₁ have been transposed.

In answer to their other comments, the advantages of using α and β of equal magnitude is explained in the methods section; the choice of 0.1 was a compromise between the common values for α of 0.05 and for β of 0.1 or 0.2 used in clinical studies. The software on the Psion 3a can be used to record and display the cusum; in this case, it is not as helpful as the increments. The software referred to by Kestin appears to be a very powerful analytical tool and may have a wide range of applications in the assessment of trainees. However, the errors and inconsistencies in the appendix are a source of confusion. It would be useful if the derivation of the definitions, variables, hypotheses and theory behind the calculations were explained more clearly.

I. G. KESTIN
Department of Anaesthesia
Derriford Hospital
Plymouth

Sir,—We appreciate the letter of Camprubi and Sabate which relates to our recent article on mucosal pH (pH₆) during orthotopic liver transplantation (OLT). Their preliminary data suggest that the use of venovenous bypass (VVB) during the anhepatic stage of OLT may indeed be beneficial for mucosal oxygenation: in contrast with our results, gastric pH, was preserved with VVB, but decreased transiently without VVB. As pointed out in our article, we were unable to strictly test the hypothesis that VVB would maintain mucosal blood flow as VVB was used routinely at our institution. Therefore, we cannot exclude the possibility that pH₆ values might have decreased to even lower values without VVB. The main purpose of our study was to assess the ability of tonometry to detect intraoperative mucosal hypoxia, to measure gastric and sigmoid pH₆, and to relate the observed changes in pH₆ to the occurrence of endotoxemia and primary graft function.

Although in the preliminary report of Camprubi and colleagues, important information on patient physical status and determinants of tissue oxygenation is lacking, the fact that cardiac index, oxygen extraction and lactate concentration are comparable with the values measured in our study suggests that the different results for gastric pH₆ during the anhepatic stage may be attributed to the more severe chronic impairment of intestinal perfusion in our patients with end-stage cirrhosis. This is supported by the lower pH₆ values measured early during hepatectomy in our study (7.28 ± 0.06) in fer, low pH₆ values (<7.32) before the second postoperative day, an observation that further supports the presence of chronic impairment of mucosal microcirculatory perfusion. Hence, pre-existing chronic mucosal hypoxia might explain why pH₆ decreased during the anhepatic stage, although overall portal flow was maintained by the use of VVB. It also cannot be excluded that the apparent difference in baseline pH₆ values between our study and that of Camprubi and colleagues is a result of the use of different blood-gas analysers. Riddington and colleagues and Takala and colleagues have shown that direct comparison of pH₆ values obtained by different analysers is not valid. Therefore, it was recommended that each institution should determine its own reference values for pH₆.

M. WELTE
Institute of Anaesthesiology
University of Munich
Klinikum Großhadern
Munich, Germany


Sir,—Clinical interaction studies evaluating the influence of different anaesthetic techniques on the action of neuromuscular blocking agents often compare two or more groups of patients treated separately but concurrently as part of the same study. These studies are termed “parallel” to emphasize their difference from other clinical studies in which patients are their own controls. Parallel comparisons include most of the usual forms of interaction studies investigating the abilities of different inhalation anaesthetics to potentiate the neuromuscular effects of neuromuscular blocking agents. However, the large between-patient variability in dose-response and dose-duration studies with neuromuscular blocking agents may preclude demonstration of subtle differences in susceptibility to these agents. Paired crossover studies can minimize the effect of inter-individual variability. However, these paired crossover studies are difficult to perform with surgical patients. Healthy ASA I or II patients that are anaesthetized two or more times, receiving on each occasion a different anaesthetic technique, are rare. Therefore, these paired crossover studies have to be performed with volunteers.

Our clinical study was designed as a parallel comparison of different anaesthetic techniques, that is opioid-nitrous oxide-oxygen, opioid-nitrous oxide-oxygen-isoflurane and opioid-nitrous oxide-oxygen-sevoflurane, in their ability to potentiate the neuromuscular effects of the most commonly used non-depolarizing neuromuscular blocking agents in surgical patients. Using a crossover design of variance (ANOVA) and subsequently the Ryan-Einot-Gabriel-Welch (REGW) multiple range test to identify eventual sources of difference, there was a significant difference in magnitude and duration of neuromuscular action between the control group, receiving no volatile agent, and the groups receiving inhalation anaesthetics. We were able to demonstrate any significant difference between opioid-nitrous oxide–oxygen–isoflurane and opioid-nitrous oxide–oxygen–sevoflurane anaesthesia on the effects of neuromuscular blocking agents. However, these interventions had no influence on the neuromuscular blocking agents, the magnitude and duration of neuromuscular blocking agents were small compared with the between-patient differences in these responses to neuromuscular blocking agents (<10% for the mean values of both magnitude and duration of action). In order to demonstrate significant differences


Dr Vanlinthout's group set out to clarify inconsistencies in the literature on the interaction between sevoflurane and neuromuscular blocking agents. Unfortunately their study merely adds to the confusion. They concluded that sevoflurane and isoflurane potentiated neuromuscular blocking drugs to a similar degree. It is not reasonable to draw such a conclusion from the data presented in their article.

For vecuronium alone, they found an ED₅₀ value of 16.9 μg kg⁻¹. With the degree of variability reported and assuming that a 20% reduction in dose requirements is clinically relevant, their study found, at best, only an 11% advantage of demonstrating such a reduction. That is, if the reduction in vecuronium dose reported in the presence of sevoflurane (ED₅₀ of 14.4 μg kg⁻¹, 17% less than with isoflurane) was a real (rather than chance) occurrence, then 204 patients would have to have received each anaesthetic to have an 80% chance of demonstrating such a difference at the 0.05 significance level.

We should interpret their study with caution. The results of their study are entirely in keeping with there being a clinically important difference in the degree of neuromuscular blocker potentiation manifest by sevoflurane and isoflurane.

P. M. C. WRIGHT
Department of Anaesthesia
University of Newcastle upon Tyne


Correspondence
in the potentiating abilities of isoflurane and sevoflurane on the effects of neuromuscular blocking agents, that is a significant difference in both depth and duration, sample sizes of several hundreds of patients per neuromuscular blocking group would have been required. Such sample sizes are beyond practicable limits. We selected our population carefully, using rigid inclusion criteria, as described in the text, to reduce patient variability. This inevitably introduced constraints such as the availability of large numbers of eligible patients within a reasonable period of time. Any investigator is limited by both resources and time. Additionally, our study was intended as a phase II trial to evaluate the interaction between sevoflurane and neuromuscular blocking agents within a limited number of surgical patients.

Both the investigator and clinician should be concerned about the ability to detect an important clinical difference. Different investigators disagree on what is a clinically significant difference. They also disagree on the risk they are willing to take of missing a meaningful effect caused by drug interaction. The choice of a 20% difference in ED$_{50}$ as a clinically significant value is arbitrary. Such a difference is small compared with the large inter-individual differences in the response to neuromuscular blocking agents. With the degree of variability and sample sizes used in our study, a 30% reduction in the ED$_{50}$ of pancuronium and atracurium and a 45% reduction in the ED$_{50}$ of vecuronium had an 80% chance of being significant at the 0.05 level.

Finally, we still feel that further interaction studies on the influence of higher concentrations of volatile anaesthetics on larger doses of neuromuscular blocking agents are warranted. To elucidate small but significant differences in the ability of isoflurane and sevoflurane to potentiate the action of neuromuscular blocking agents, these interaction studies need to be performed in volunteers who are studied twice, receiving isoflurane on one occasion and sevoflurane on the other.

L.E.H. VANLINTHOUT  
T. DE BOO  
L.H.D.J. BOUW  
Institute for Anaesthesiology  
University Hospital of Nijmegen  
Nijmegen, The Netherlands

4. Freedman JA, Chalmers TC, Smith H, Kuebler RR. The importance of beta, the type II error and sample size in the assessment of trainees. However, the errors and inconsistencies in the appendix are a source of confusion. It would be useful if the derivation of the definitions, variables, hypotheses and theory behind the calculations were explained more clearly.

Sir,—It was interesting to read the commentary by Kestin describing the use of cusum analysis to measure the competence of anaesthetic trainees at practical procedures. The original article in the British Journal of Anaesthesia 1995; 75: 805-809. 2. Williams SM, Parry BR, Schlup MMT. Quality control: an application of the cusum. British Medical Journal 1992; 304: 1359-1360.

Cusum: a statistical method to evaluate competence in practical procedures

Sir,—It was interesting to read the commentary by Kestin describing the use of cusum analysis to measure the competence of anaesthetic trainees at practical procedures. The cusum is a useful graphical tool for discerning trends from a series of observations. The derivation of boundary lines for sequential data can aid comparison of the observed proportions of success or failure against predetermined standard criteria. This could be developed into a continuous performance monitor in anaesthesia.

In his article, Kestin discussed the problems of keeping paper records and plotting fractions on the graphs. In response to this, we noted that a computerized personal log book system for practical procedures to run in parallel with the electronic anaesthetic log book on the Psion 3a. However, in setting up the algorithms we have encountered a number of problems.

The values for $s$, $h_{0}$ and $h_{1}$ calculated using the formulae in the appendix do not agree with the values in table 1 of the article. On reviewing the original articles on the derivation of the cusum, the value of $Q$ should read $Q=\ln [(1-p_{1})(1-p_{2})]$. The values in the table are incorrect if this calculation is used for $Q$. The null hypothesis in the article is stated as "the true failure rate is NOT different from the acceptable failure rate" and if the cusum exceeds $h_{0}$ then it is rejected. This does not imply that the true failure rate exceeds the unacceptable failure rate which is the performance indicator which interests us. Similarly the alternative hypothesis is stated as "the true failure rate is equal to or exceeds the unacceptable failure rate" and if the cusum decreases to less than $h_{0}$ then it is accepted. Surely this would imply that the failing cusum of registrar B in figure 1 has a true failure rate that is equal to or exceeds the unacceptable failure rate.

$p_{0}$ corresponds to the failure rate under the null hypothesis which should surely read "the true failure rate is not different from the unacceptable failure rate". If the cusum exceeds $h_{0}$ then this hypothesis is accepted and the trainee's performance is unacceptable with reference to the agreed unacceptable failure rate. Similarly $p_{1}$ corresponds to the failure rate under the alternative hypothesis which should surely read "the true failure rate is equal to or exceeds the acceptable failure rate". If the cusum decreases to less than $h_{0}$ then this hypothesis is rejected and the trainee's performance is no worse than the accepted failure rate.

Furthermore, there appear to be clearest criteria in the definition of success or failure for a particular procedure. This is confusing and it is difficult to see why type 1 and 2 errors of 10% (0.1) where chosen.

For reasons of convenience for the plotting of the graph, the values of $s$, $h_{0}$ and $h_{1}$ where multiplied by 10. It is not clear if, in the event of a success $F_{1}$ or failure $F_{2}$ or $F_{3}$ then the hypothesis is rejected, then the trainee's performance is lower than the acceptable failure rate.

The statistical method described by Kestin appears to be a very powerful analytical tool and may have a wide range of applications in the assessment of trainees. However, the errors and inconsistencies in the appendix are a source of confusion. It would be useful if the derivation of the definitions, variables, hypotheses and theory behind the calculations were explained more clearly.

E. J. HAMMOND  
Department of Anaesthesia  
Poole General Hospital  
Poole


Sir,—I would like to thank Drs Hammond and McIndoe for their interest, and in particular for correcting the error in the formula for $Q$ in the appendix—$p_{0}$ and $p_{1}$ have been transposed.

In answer to their other comments, the advantages of using $\alpha$ and $\beta$ of equal magnitude is explained in the methods section; the choice of 0.1 was a compromise between the common values for $\alpha$ of 0.05 and $\beta$ of 0.1 or 0.2 used in clinical studies. The software on the Psion 3a can be used to record and display the cusum; in this case, it is not as helpful to have $\alpha$ and $\beta$ equal than if graphical methods are used. The graphs used by our trainees were plotted using the nearest integers to 10 and $10^{-1}$ as the increments.

The original article in the British Medical Journal referred to by Drs Hammond and McIndoe was the article that stimulated my interest in this topic. However, I found it difficult to understand the basic statistical concepts of cusum analysis from this article. I found it even more difficult to write an explanation myself, and this article is the only one of mine which the reviewers have ever suggested that more text be included to clarify the argument! I would refer them and others to Davies' text for a better explanation than I could provide.

I. G. KEFFIN  
Department of Anaesthesia  
Derriford Hospital  
Plymouth