Pick up the Pieces

Depth of Anesthesia and Long-term Mortality

IN this issue of ANESTHESIOLOGY, Kertai *et al.*[1] report on the association between anesthetic depth, as measured electroencephalographically with a bispectral index (BIS) monitor, and long-term mortality. It is the fifth observational study on a continuing controversy, which began in 2005 with a paper by Monk *et al.*[2] We will try to pick up the pieces here.

Nonrandomized studies have indicated potential advantages of administering anesthesia using processed electroencephalograms. These potential advantages include rapid emergence from anesthesia, decreased anesthetic requirements, less postoperative nausea and vomiting, and earlier discharge from the postanesthesia care unit.[3] When the first randomized multicenter trial, the B-Aware Trial,[4] allocated 2,463 patients at high risk of intraoperative awareness to BIS-guided anesthesia or routine care, they found an 80% reduction in awareness incidence, albeit with few endpoints (2 *vs.* 11 cases of definite awareness). In contrast, an American monocenter trial, the B-Unaware Trial,[5] compared BIS monitoring with end-tidal gas monitoring in 2,000 patients at high risk for awareness and found two cases within each study arm.

Without attempting to address why the anesthesia community has remained skeptical of brain-function monitoring, the most controversial question remains whether there is a possible causal, rather than purely statistical, association between deep anesthesia and long-term outcomes.

In 2005, Monk *et al.*[2] used multivariable time-to-event analysis to create a predictive model for 1-yr mortality in 1,046 patients undergoing noncardiac surgery.[2] The overall mortality in that study was 5.5%, where cancer was the cause of 52% of these deaths. Independent predictors were comorbidity (included as dichotomized Charlson comorbidity index), cumulative "deep hypnotic time" (number of minutes BIS values were less than 45), and intraoperative hypotension. The paper by Monk *et al.*[2] generated a storm of reactions, largely in response to the authors' suggestion of causality: "These associations suggest that intraoperative anesthetic management may affect outcomes over longer time periods than previously appreciated." Main criticisms included: confusing causation and statistical association[6]; combining comorbidities to a single score and dichotomizing Charlson comorbidity score and BIS values, both resulting in information loss and risk of residual confounding[7]; and coauthorship by an employee of Aspect Medical Systems, Newton, MA (manufacturer of BIS® monitor), suggesting a possible conflict of interest.[8] Other readers commented that the BIS value range of 40–45 that prevented awareness in trials was actually considered "deep anesthesia" in the study by Monk *et al.*[2]

The pathophysiologic link between deep anesthesia and mortality is still unknown. Monk *et al.*[2] postulated that deep anesthesia might interfere with the immune response, resulting in organ damage by (as yet) undefined mechanisms, such as inflammation and hypercoagulation-induced disturbances of the microcirculation.

Lindholm *et al.*[9] subsequently reported an association between low intraoperative BIS values and 1-yr mortality in 4,087 noncardiac surgery patients. Without malignancy status as a covariable in the multivariable model, the relationship established was similar to that observed by Monk *et al.*[2] However, with the addition of malignancy status, the relationship between low BIS and mortality disappeared. The authors concluded:

> … the statistical relation between 1-yr mortality and [time] BIS <45. . . is sensitive to the selection of co-variates in the statistical model, and a randomized study is required to demonstrate that there really is a causal impact … and, if it does, the effect is probably very weak in comparison with co-morbidity as assessed by ASA [American Society of Anesthesiologists] physical score, the pre-existing malignancy status at surgery and age.[9]

Recently, Leslie *et al.*[10] studied B-Aware Trial[4] data to determine whether incidence of death, myocardial infarction, and stroke were lower among patients allocated to BIS-guided management compared with routine care. After median 4.1-yr follow-up, mortality was similar in both groups. In a *post hoc* cohort analysis, the authors[10] investigated the "causal" association of low BIS values and outcomes in the BIS-monitored group only (as BIS values were not recorded in the control group). They used a propensity score to adjust for other confounders. For mortality, the hazard ratio of BIS values lower than 40 for more than 5 min (as compared with other BIS values) was 1.41 ($P = 0.039$). They also reported statistically significant higher odds ratios for myocardial infarction (1.94) and stroke (3.23) among patients with low BIS values. Unfortunately, data on preoperative malignancy status, a large confounder in previous studies, was not available.

In 2010, Kertai *et al.*[11] reported on the BIS-mortality relationship using a cohort of 460 patients from the B-Unaware

◆ This Editorial View accompanies the following article: Kertai MD, Palanca BJA, Pal N, Burnside BA, Zhang L, Sadiq F, Finkel KJ, Avidan MS, in behalf of the B-Unaware Study Group: Bispectral index monitoring, duration of bispectral index below 45, patient risk factors, and intermediate-term mortality after noncardiac surgery in the B-Unaware Trial. ANESTHESIOLOGY 2011; 114:545–56.

Accepted for publication December 16, 2010. The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

 Anesthesiology 2011; 114: 485–7

Trial[5] that had undergone cardiac surgery. Eighty-two patients (17.8%) died within 3 yr. Cumulative duration of low BIS values was independently associated with an adjusted hazard ratio of 1.29. It is noteworthy that there was no difference in the relationship between end-tidal anesthetic gas concentrations during the anesthetic maintenance phase and BIS values. More β blocker use and left ventricular dysfunction were present among patients with low BIS values for more than 4 h, which prompted the authors[1] to speculate that cumulative duration of BIS values lower than 45 is likely to be a marker of systemic illness, poor cardiac function, and complicated intraoperative course.

In this issue of ANESTHESIOLOGY, Kertai *et al.*[1] report on the cohort of the remaining 1,473 noncardiac surgery patients of the B-Unaware Trial.[1] Within 3 yr, 358 patients (24.3%) had died. Two hundred eighteen of these deaths were attributed to cancer. Male sex, stroke history, type II diabetes, surgery for malignancy, ASA Physical Status Classification System category, and intensive care time were statistically associated with increased mortality, whereas increased body mass index and hemoglobin concentration as well as intermediate risk surgery were associated with decreased mortality. BIS-monitored and unmonitored patients had similar mortality (24.9 *vs.* 23.7%). In contrast to previous results from cardiac surgery patients,[11] within this cohort, extended low BIS values of less than 45 were not associated with mortality, nor were increasing mean and cumulative end-tidal anesthetic concentrations.

In a domain full of scientific and political controversy, it is imperative that the design and analysis of the study are suited to test the underlying hypotheses. First, although some authors indeed used terminology as "predicting" and "predictors," the actual aim of all the aforementioned studies was to investigate causal associations between intraoperative anesthetic intervention factors and long-term outcomes. However, quantifying a predictive or causal association requires different approaches in terms of study design and statistical analysis.

When one truly aims to claim a primary causal relationship, the ultimate challenge is to ensure that all other potential causes of the outcome, which may be related to the factor of interest (*e.g.*, deep anesthesia), are ruled out. Ideally, this goal is achieved using a randomized study design. Randomization ensures that—provided there is a large enough sample population—all other known and, more importantly, unknown causal factors (so-called "confounders") are balanced by design. Data from nonrandomized studies, such as those discussed above, can be used to investigate causal associations—but, using multivariable statistical analysis, one can adjust only for the influence of the known and measured confounders. The influence of unknown and unmeasured confounders on observed associations (between anesthetic factors and mortality) may always be present and obscure inferences.

Given the challenges in studying causal associations, candidate predictors in a multivariable statistical model are preferably selected on theoretical rather than statistical (*P* value) grounds to avoid the chance that important measured confounders are being overlooked.[12] Often, one may summarize confounders in a single propensity score.[12] A covariable selection approach based on *P* values may, indeed, be useful in studies focusing purely on predictive associations because they allow one to select the smallest set of covariables that best predicts the outcome of interest.[13–15] However, one should then refrain from drawing conclusions that are of a causal nature.

The authors of the studies noted above selected their confounders based on *P* values of the univariable association of each confounder and the outcome. Accordingly, given the causal aim of the studies, the number of covariables in each of the final models was likely too low. Consequently, relevant confounders might have been missed—and the effects of low BIS values on long-term outcomes is likely overestimated.

A second issue is the expected size of the association. Although the outcome is of utmost importance, it is theoretically not likely that relatively small differences in depth of anesthesia, as indicated by a lower processed electroencephalogram index, will have a large influence on long-term outcomes. The effect will, therefore, be much smaller than that of known important factors such as comorbidity. If such factors are poorly accounted for in the model, the analysis leaves more room for smaller associations to be detected. This effect was clearly illustrated by Lindholm *et al.*,[9] who found that "low BIS" was no longer statistically significant after malignancy was included as covariable in the model. The complex associations among individual predictors and with outcomes (*i.e.*, not just the raw association of the predictors and outcome) is also important. These so-called "partial" (or semipartial) correlations are the driving force behind the attenuation of the anesthesia depth effect in the context of other potential causes. For example, it is likely that advanced comorbidities and electroencephalogram indices of anesthetic depth are related.

The same principle holds true when defining outcomes. If overall mortality is dominated by one disease-specific cause (*e.g.*, 61% of mortalities attributed to cancer[1]), a small association will not easily be detected—even if "comorbid malignancy at surgery" is a model covariable.

Third, all five studies suffer from the same problem, namely arbitrary dichotomization of the continuous determinant, BIS.[1,2,9–11] The chosen thresholds were based on manufacturer recommendations to prevent awareness. Dichotomization of continuous variables, although tempting, is often unwise because it may cause spurious findings.[16] Future analyses should deal with this aspect more appropriately.

Finally, Monk *et al.*[2] and Lindholm *et al.*[9] studied cohorts of elective noncardiac surgery patients, Leslie *et al.*[10] and Kertai *et al.*[1,11] studied patients with a high probability of intraoperative awareness who participated in randomized controlled trials. Not surprisingly, mortality was higher in the awareness trial patients. Indeed, many of the predictors of intraoperative awareness are also predictors of poor out-

comes, including mortality. Although three studies[2,10,11] suggest that the association between low BIS values and mortality is real, the extensive data collected by Kertai *et al.*[1,11] also indicate that this relationship may be an epiphenomenon, reflecting a poor preoperative condition that predisposes patients to late postoperative mortality. A strong argument for this reasoning is that entering malignancy status in the multivariable models removed the association. A second argument is the absence of a relationship between low BIS values and higher mean or cumulative doses of volatile anesthetics.[1,11]

Can we now be confident that "deep anesthesia," as indicated by electroencephalogram monitors, is just a marker of a poor preoperative condition and not in itself harmful? Unfortunately, we cannot. Retrospective or *post hoc* analyses can generate hypotheses and create new study questions, but, ultimately, these hypotheses need to be addressed in carefully designed prospective observational or interventional studies. Anesthesiologists have "let the genie out of the bottle" when they began to acknowledge the possibility that anesthesia may not be as reversible as previously believed. Experimental data are rapidly accumulating, indicating that exposure to clinically relevant concentrations of GABAergic (γ-aminobutyric acid–mediated) anesthetics, such as propofol and volatile agents, in vulnerable periods of rapid brain development is harmful.[17,18] Ongoing observational and randomized studies should answer the question as to whether experimental anesthetic neurotoxicity translates into clinical effects in vulnerable patient groups.

To summarize, the data by Kertai *et al.*[1,11] are sufficiently reassuring that there is no imminent reason to change our practice and strive to run patients as "light" as possible using electroencephalogram monitoring—trying to tightrope the thin line between "deep anesthesia" and risk of awareness. Nonetheless, we must investigate the mechanisms by which patients with cancer respond to standard anesthetic doses with more pronounced cortical electrical depression and how pharmacokinetics and dynamics are altered by preexisting disease states. Also, the modulating effect of various anesthesia techniques on the inflammatory response to surgery in various patient groups is amenable to prospective studies. But, in the end, only adequately powered randomized trials can answer the question of whether management aimed at minimizing anesthetic exposure will improve outcomes in vulnerable patients.

Cor J. Kalkman, M.D., Ph.D.,* Linda M. Peelen, Ph.D.,† Karel G. Moons, Ph.D.† *Division of Perioperative and Emergency Medicine, †Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, the Netherlands. c.j.kalkman@umcutrecht.nl

References

1. Kertai MD, Palanca BJA, Pal N, Burnside BA, Zhang L, Sadiq F, Finkel KJ, Avidan MS, B-Unaware Study Group: Bispectral index monitoring, duration of bispectral index below 45, patient risk factors, and intermediate-term mortality after noncardiac surgery in the B-Unaware Trial. ANESTHESIOLOGY 2011; 114:545–56
2. Monk TG, Saini V, Weldon BC, Sigl JC: Anesthetic management and one-year mortality after noncardiac surgery. Anesth Analg 2005; 100:4–10
3. Liu SS: Effects of Bispectral Index monitoring on ambulatory anesthesia: A meta-analysis of randomized controlled trials and a cost analysis. ANESTHESIOLOGY 2004; 101:311–5
4. Myles PS, Leslie K, McNeil J, Forbes A, Chan MT: Bispectral index monitoring to prevent awareness during anaesthesia: The B-Aware randomised controlled trial. Lancet 2004; 363: 1757–63
5. Avidan MS, Zhang L, Burnside BA, Finkel KJ, Searleman AC, Selvidge JA, Saager L, Turner MS, Rao S, Bottros M, Hantler C, Jacobsohn E, Evers AS: Anesthesia awareness and the bispectral index. N Engl J Med 2008; 358:1097–108
6. Levy WJ: Is anesthetic-related mortality a statistical illness? (letter) Anesth Analg 2005; 101:1238
7. Berry AJ: Observational studies identify associations, not causality (letter). Anesth Analg 2005; 101:1238
8. Drummond JC, Patel PM: Editorial Board reproached for publication of BIS-mortality correlation (letter). Anesth Analg 2005;101:1238–9
9. Lindholm ML, Träff S, Granath F, Greenwald SD, Ekbom A, Lennmarken C, Sandin RH: Mortality within 2 years after surgery in relation to low intraoperative bispectral index values and preexisting malignant disease. Anesth Analg 2009; 108:508–12
10. Leslie K, Myles PS, Forbes A, Chan MT: The effect of bispectral index monitoring on long-term survival in the B-aware trial. Anesth Analg 2010; 110:816–22
11. Kertai MD, Pal N, Palanca BJ, Lin N, Searleman SA, Zhang L, Burnside BA, Finkel KJ, Avidan MS, B-Unaware Study Group: Association of perioperative risk factors and cumulative duration of low bispectral index with intermediate-term mortality after cardiac surgery in the B-Unaware Trial. ANESTHESIOLOGY 2010; 112: 1116–27
12. Rothman KJ, Greenland S, Lash TL: Modern Epidemiology. Philadelphia, Lippincott & Williams, 2008
13. Royston P, Moons KG, Altman DG, Vergouwe Y: Prognosis and prognostic research: Developing a prognostic model. BMJ 2009; 338:b604
14. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG: Prognosis and prognostic research: What, why, and how? BMJ 2009; 338:b375
15. Moons KG, Altman DG, Vergouwe Y, Royston P: Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. BMJ 2009; 338:b606
16. Royston P, Altman DG, Sauerbrei W: Dichotomizing continuous predictors in multiple regression: A bad idea. Stat Med 2006; 25:127–41
17. Culley DJ, Xie Z, Crosby G: General anesthetic-induced neurotoxicity: An emerging problem for the young and old? Curr Opin Anaesthesiol 2007; 20:408–13
18. Jevtovic-Todorovic V, Hartman RE, Izumi Y, Benshoff ND, Dikranian K, Zorumski CF, Olney JW, Wozniak DF: Early exposure to common anesthetic agents causes widespread neurodegeneration in the developing rat brain and persistent learning deficits. J Neurosci 2003; 23:876–82