In Reply:

We thank Dr. Silverstein *et al*., Dr. Nainar *et al*., and Dr. Silbert *et al*. for their thoughtful comments about our article.[1] We agree entirely that the issue at hand is whether postoperative cognitive decline (POCD) is clinically relevant. For years, elderly people and their families have dreaded the prospect of surgery because they have been led to believe that there is a causal association between surgery and permanent cognitive decline. Dr. Nainar *et al*. contend that it is unacceptable for us to conclude that the decision to proceed with surgery in the elderly may currently be made without factoring in the specter of persistent POCD. An alternative perspective is that it is unacceptable that, based on scant evidence, the impression has been created that persistent POCD is a major public health problem. To date, there are no studies that convincingly demonstrate persistent POCD that is attributable to a surgical event. Currently, POCD does not have any definitive standard diagnostic criteria, neuroimaging or biomarker correlates, or pathognomonic anatomical or histologic findings. This is in contrast to dementia, for which all these criteria are satisfied. POCD is currently a hypothetical phenomenon; it is not a recognized disease.

In a similar vein, we concur with Dr. Silbert *et al*. that our study has not provided level 1 evidence refuting persistent POCD attributable to a surgical event. Perhaps, a more appropriate approach is to first examine whether there is any level 1 evidence in support of either early or persistent POCD attributable to a surgical event. To the best of our knowledge, such evidence is lacking. It is curious that for some of our colleagues, the burden of proof seems to have shifted from verification of a hypothetical condition (POCD) toward its refutation. The null hypothesis should be that there is no cognitive decline attributable to a surgical event; our study could not reject this null hypothesis.

Dr. Silverstein *et al*. agree that the evidence for lasting POCD after noncardiac surgery is not compelling but note that even POCD lasting weeks to months may have a major effect on elderly people's quality of life. In the study cited by Dr. Silverstein *et al*. showing POCD at 6 months, decline could not be attributed to surgery because there was no nonsurgical control group. The lead author of this study cautioned, "the observed incidence of long-term cognitive deterioration may not be a complication but instead may represent the basal rate to be expected in a population with a mean age of 70 yr and comorbid medical disease."[2] Even the evidence for POCD at 3 months attributable to surgery[3,4] is unconvincing for several reasons. (i) The patients with POCD at 3 months overlapped only moderately with the patients with POCD at 1 week.[3,4] (ii) The diagnosis of POCD relied on a correction factor for learning[3,4] in control groups that might not have been appropriately matched with the surgical groups.[5] (iii) The cognitive trajectories before surgery were unknown. (iv) Dementia was excluded in the surgical groups by a mini-mental status examination cutoff score[3,4] that may not detect mild dementia.

Dr. Silbert *et al*. point out that vascular disease and other conditions are known risk factors for cognitive decline. Considering this, it is surprising that they suggest that it is established that cardiac surgery is associated with persistent POCD. The study they referenced to corroborate this[6] did demonstrate persistent POCD after heart surgery but could not attribute cognitive decline to the surgical event because there was no matched nonsurgical control group. Because patients undergoing cardiac surgery usually have vascular disease and other risk factors, it is unsurprising that they experience long-term cognitive decline, regardless of whether or not they have cardiac surgery.[5]

Although vascular disease and vascular risk factors may be associated with both cognitive impairment and higher rates of surgery, the contention that evidence for POCD is found in the higher rate of decline in participants with mild dementia (clinical dementia rating [CDR], 0.5) than those who were nondemented (CDR, 0) is flawed. Within the CDR of 0.5 group, participants who had a surgical event exhibited a similar rate of cognitive decline to those who had neither surgical nor major illness events. Furthermore, participants in the CDR of 0.5 group who did not have surgery or illness declined markedly faster than participants in the CDR of 0 group who did have surgery.

All three letters seek clarification about the psychometric tests used in our study. The primary analyses of cognitive change in our study (table 4)[1] used a composite of scores from a 90-min comprehensive battery of cognitive tests that includes the same or similar measures used in other studies of POCD, including tests of attention, executive function, episodic memory, semantic memory, and visuospatial skills. Composite scores tend to be more reliable than scores from individual measures, and our preliminary analyses did not detect any differences for any individual measure. The CDR was obtained independently of the composite psychometric score mainly to stage dementia severity; failure to rigorously assess surgical patients for dementia misses a potent confounder in studies focusing on POCD. In fact, as shown in table 4 of our article, the initial CDR was the strongest predictor of a given participant's cognitive trajectory in our study; predominantly, people with dementia declined markedly on psychometric evaluations.[1] POCD may also manifest as an increased rate of development of dementia in initially nondemented participants, but we did not find any evidence for this possibility in our study.

Dr. Silbert *et al*. and Dr. Nainar *et al*. stress the limitations of our study, which we detailed in our Discussion section.[1] However, our approach also had important methodological strengths in relation to existing studies, which also bear mentioning. Most of these strengths derive from the fact that participants typically enrolled in the study long before any surgical events. Thus, participants in both the surgical and nonsur-

Supported by grants P01 AG03991 and P50 AG05681 from the National Institutes of Health (Bethesda, Maryland) and a grant from the University of Missouri Alzheimer's Disease and Related Disorders Program (Columbia, Missouri).

gical groups would be more likely to have similar baseline characteristics. Furthermore, the majority of the learning effect, which is typically pronounced early during repeated testing, was absorbed by the preevent cognitive trajectories, essentially uncoupling the learning effect from postevent cognitive trajectories. Although not without faults, this is a much more robust approach than assuming that the mean and variance of the learning effect in control patients is the same as in surgical patients. Finally, it allowed each participant to be his or her own control by providing preevent cognitive trends.

We agree with Dr. Silverstein *et al.* that determining preoperative trajectory can be helpful for several reasons, including establishing a cognitive course. As mentioned earlier, enrolling participants long before potential surgery or illness increases the possibility of including groups that are well matched at baseline. We further statistically adjusted for the effects of several covariates to account for the potential group differences at baseline. Certainly, we did not suggest that control groups are unnecessary in prospective studies. However, we strongly believe that control groups must be well matched for comorbidities, as Selnes *et al.*[5] have shown, and particularly for dementia, as we have shown.[1] Unfortunately, it is difficult to match groups appropriately at the time of surgery. The concern about a biasing effect through exclusion of those who died and those who were lost to follow-up is important, and one that we acknowledged as a limitation of our approach.[1]

Power calculations are useful for study design rather than for analysis of results. The clinical significance of the results of a study is best appreciated from the magnitude and precision of the findings. In relation to our study, the changes in slopes in table 4[1] convey the rate of change in cognitive decline, and the SEs and corresponding CIs reflect the precision of the results. The power calculation suggested by Dr. Silverstein *et al.* based on a 1% difference in a binary outcome (*i.e.*, POCD *vs.* no POCD at 1 yr) is likely to be inappropriate. A difference of this magnitude is probably within the margin of error of the diagnostic tests for POCD and is therefore of dubious clinical significance. Furthermore, POCD is imprecisely defined based on an arbitrary statistical change in psychometric scores and an unvalidated correction for learning.[3,4] Our cognitive assessment methods had two potential advantages over previous approaches. First, we evaluated participants for the development of dementia through standard Cox's proportional hazards models on the time to the onset of dementia. Second, we assessed longitudinal trends in continuous psychometric scores to compare the rates of cognitive change among subject groups using general linear mixed models. These approaches have the advantage of enhancing statistical power compared with a binary determination based on incidence and are generally more informative.

The discomfort about the simulated event is understandable. It is reasonable to expect that cohorts of elderly people, especially those with dementia, would experience accelerating cognitive decline over time. In this scenario, if the rate of cognitive decline was analyzed before and after any arbitrary temporal event, we would expect a steeper decline after the event. To test the hypothesis that the change in slope was not exacerbated by medical or surgical events (*i.e.*, the null hypothesis), we needed to determine the expected change in slope after events that are known not to affect the cognitive trajectory. To this end, we simulated events for the control group based on the observed events in the surgical and illness groups and then averaged over our uncertainty about when these events occurred. In an earlier approach, we disregarded the time of the event and modeled the complete longitudinal course of each of the three groups. With either approach, we could not detect exacerbation of cognitive decline in the surgical and medical groups. We would welcome alternative suggestions for dealing with this difficult statistical problem.

Because our study is observational, it is not surprising that the number of observations before and after the event for the surgical and illness groups might not be the same. However, such variability does not bias the slope estimates if we are willing to assume a piecewise linear longitudinal trend over time pre- and postevent, and our analyses indicate that a linear trend fits reasonably well. Moreover, during our modeling process, we noted that only using observations within 1 or 2 yr of the event (or simulated event) did not substantially alter our conclusions. Fewer longitudinal follow-ups do imply less accurate estimates of slopes, which are reflected in our reported SEs of the estimated slopes (table 4).[1]

Dr. Silbert *et al.* make an important point about seeking vulnerable subgroups or individuals who may have POCD attributable to surgery. Our study suggests that POCD is not a universal phenomenon, but we cannot exclude that some patients with specific characteristics or undergoing particular surgeries are susceptible. Our cognitive trends were roughly normally distributed without any apparent difference in the number and distribution of outliers among the surgical, illness, and control groups, but we had limited power to detect POCD in potentially vulnerable subgroups. However, without knowing an individual's cognitive trajectory before surgery, we do not see how an informed judgment can be made as to whether cognitive decline is exacerbated by surgery. Therefore, an appropriate way to identify vulnerable subgroups would be to expand our initial study and tap into the information provided by several additional Alzheimer Disease Research Centers.

We concur with Dr. Silverstein *et al.* that POCD attributable to surgery and persisting for years is, at most, an infrequent problem. We further contend that the jury is out regarding early POCD attributable to surgery and lasting beyond a week. In response to concerns of Dr. Nainar *et al.*, we readily acknowledge that our study has important limitations and caution that our findings should be interpreted in the context of a single-center, retrospective study focusing on a heterogeneous cohort of participants with potential confounders and missing data. Dr. Silbert *et al.* are correct that the results of our study, similar to those of any study, should

be treated with caution.[7] However, randomized controlled trials are not always practical or ethical,[7] and there are instances when reliable evidence has been obtained from observational studies. A striking example is research demonstrating the efficacy of condoms in preventing human immunodeficiency virus transmission.[8] In the case of POCD, without collaborating with existing longitudinal studies of aging such as Alzheimer Disease Research Centers, it would be impractical to enroll a large cohort and conduct frequent neurocognitive assessments prospectively on subjects who may have surgery in the future. Moreover, for most surgeries, it would be unethical to randomize patients to surgery or control solely based on the soft evidence for POCD. We applaud Dr. Nainar *et al.* for their conclusion that our study should not hinder research; it should encourage scrupulous follow-up studies designed to support, refute, or refine our preliminary findings. To this end, we are pursuing a multicenter Alzheimer Disease Research Center-linked study and would welcome interested collaborators. Future studies should rigorously reexamine whether elderly patients experience early POCD and, further, whether there are specific patients undergoing specific surgeries who are susceptible to persistent POCD that is attributable to surgical events.

Michael S. Avidan, M.B.B.Ch., F.C.A.S.A.,* Adam C. Searleman, B.S., Martha Storandt, Ph.D., Chengjie Xiong, Ph.D., John C. Morris, M.D., Alex S. Evers, M.D. *Washington University in St. Louis, St. Louis, Missouri. avidanm@wustl.edu

References

1. Avidan MS, Searleman AC, Storandt M, Barnett K, Vannucci A, Saager L, Xiong C, Grant EA, Kaiser D, Morris JC, Evers AS: Long-term cognitive decline in older subjects was not attributable to noncardiac surgery or major illness. ANESTHESIOLOGY 2009; 111:964–70
2. Williams-Russo P: Cognitive impairment: Yes or no? ANESTHESIOLOGY 2000; 92:1848–50
3. Moller JT, Cluitmans P, Rasmussen LS, Houx P, Rasmussen H, Canet J, Rabbitt P, Jolles J, Larsen K, Hanning CD, Langeron O, Johnson T, Lauven PM, Kristensen PA, Biedler A, van Beem H, Fraidakis O, Silverstein JH, Beneken JE, Gravenstein JS: Long-term postoperative cognitive dysfunction in the elderly ISPOCD1 study. ISPOCD investigators international study of post-operative cognitive dysfunction. Lancet 1998; 351:857–61
4. Monk TG, Weldon BC, Garvan CW, Dede DE, van der Aa MT, Heilman KM, Gravenstein JS: Predictors of cognitive dysfunction after major noncardiac surgery. ANESTHESIOLOGY 2008; 108:18–30
5. Selnes OA, Grega MA, Bailey MM, Pham LD, Zeger SL, Baumgartner WA, McKhann GM: Do management strategies for coronary artery disease influence 6-year cognitive outcomes? Ann Thorac Surg 2009; 88:445–54
6. Newman MF, Kirchner JL, Phillips-Bute B, Gaver V, Grocott H, Jones RH, Mark DB, Reves JG, Blumenthal JA: Longitudinal assessment of neurocognitive function after coronary-artery bypass surgery. N Engl J Med 2001; 344:395–402
7. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, Petticrew M, Altman DG: Evaluating non-randomised intervention studies. Health Technol Assess 2003; 7: iii–x, 1–173
8. de Vincenzi I: A longitudinal study of human immunodeficiency virus transmission by heterosexual partners. European Study Group on Heterosexual Transmission of HIV. N Engl J Med 1994; 331:341–6

(Accepted for publication January 6, 2010.)

High-dose Phenylephrine and Ephedrine in Obstetrics: Proving the Holy Grail Requires More Data

To the Editor:

Dr. Smiley's supportive editorial, Burden of Proof, prompted me to eagerly read the obstetric article by Dr. Ngan Kee *et al.*, hoping to find the promised "Holy Grail" of obstetric anesthetic blood pressure control.[1,2] Instead, I found a questionable clinical methodology of high-dose vasopressor therapy for American Society of Anesthesiologists 1–2 delivery under bupivacaine spinals with limited scientific disclosure of data, burdening me to ask, where is the proof and how is it pertinent to standard clinical methods?

1. Intravenous prehydration was withheld, and total fluid was limited to 2 l during the 27 min to delivery.
2. Immediately, preoperatively obtained oscillotonometric systolic pressures were chosen as therapeutic goals (but nowhere are these group systolic values or statistical analysis provided).
3. Infusions of phenylephrine (100 μg/ml—totaling 960–1,690 μg) *versus* ephedrine (8 mg/ml—totaling 44.8–79.2 mg) were initiated at 1 ml/min to sustain the therapeutic goal "baseline" pressure.
4. Rescue treatment occurred using phenylephrine 100 μg bolus in both groups without reporting amounts or intervals administered.
5. Supplemental oxygen was withheld unless saturation decreased below 95% (no indication of number of patients given oxygen in any group is provided).
6. The vasopressor was apparently simply "stopped if systolic pressure was greater than 120% baseline," occurring in approximately 40% of all patients.

We find that

7. While the E *versus* P group had more hypotension (defined as systolic drop to $<$ 80% of baseline) and phenylephrine rescue (22% *vs.* 2%), respectively, the mean minimum recorded systolic blood pressures were similar and under classic parameters, not typically requiring treatment in either group (101 *vs.* 104 Torr, $P = 0.33$).
8. Almost double the "equipotent volume" of P *versus* E was infused (13 *vs.* 7.7 ml), respectively, without information of phenylephrine "rescue" injections/amounts.
9. Hypertension was found in 41% *versus* 47% of P *versus* E patients, respectively, and these maximum absolute pressures reported differed significantly statistically ($P = 0.044$).