Critical appraisal of ENGAGES: cognitive dissonance and anesthesia research

Troy S. Wildes, Michael S. Avidan

Department of Anesthesiology, Washington University School of Medicine, St. Louis, MO, USA
Correspondence to: Michael S. Avidan, MBBCh. Department of Anesthesiology, Washington University School of Medicine, Campus Box 8054, 660 S. Euclid Ave., St. Louis, MO 63110, USA. Email: avidanm@wustl.edu.
Provenance: This is an article commissioned by the Academic Editor Dr. Hao Zhang (Department of Anesthesiology, Rocket Force Characteristic Medical Center of PLA, Beijing, China).

Thank you for the opportunity to respond to the editorial by Koch *et al.* in which several criticisms of the ENGAGES trial (1) are described and the contribution to the cumulative literature regarding delirium prevention using EEG guidance is questioned. We address these points here.

First, the ENGAGES study is questioned because of conflict with results from prior randomized trials (2-4). Science is an iterative process where initial studies commonly produce false positive or exaggerated findings (5), particularly for outcomes they are not primarily designed to address. Such findings are commonly refined or even contradicted by more focused or rigorous subsequent trials (6). The three previous trials mentioned here have limitations regarding evidence informing any causal relationship between delirium outcomes and the modifiable factors of EEG-guidance, anesthetic dose, and suppression (2-4). Two of these—CODA (2) and BAG-RECALL (4)—studied delirium as a secondary outcome. BAG-RECALL did not show a statistically significant reduction in delirium with bispectral index (BIS)-guidance. The third trial, the SuDoCo trial (3), was led by Radtke and Spies, was partially funded by the BIS manufacturer, and had an important design weakness: there were separate groups of anesthesiologists managing anesthetic care for patients in the BIS-guided and in the usual care groups. There are several methodological concerns regarding the CODA trial, which are summarized in *Table 1*.

Secondly, delirium screening in ENGAGES is described as inadequately sensitive. If Koch *et al.* are correct in their assertion, we would expect a lower delirium incidence in ENGAGES than in similar trials; yet the delirium incidence in ENGAGES was higher than in the CODA (2),

SuDoCo (3) and PODCAST (9) trials. The ENGAGES investigators included world experts in delirium research, who designed and scrupulously oversaw the detection methodology employed in this trial for the primary outcome of delirium. Specific reservations are raised regarding using the Confusion Assessment Method (CAM). However, it is the most widely used method to detect delirium, and has excellent psychometric properties for hypoactive and hyperactive delirium (10). It is also incorrectly stated that delirium was not assessed on the day of surgery in ENGAGES. It was, and was reported in Table 3: delirium incidence on day of surgery was 19.2% in the EEG-guided and 21.9% in the usual care group [difference =2.7% (95% CI, –7.6% to 2.2%; P=0.26)] (1). The utility of delirium chart review is also questioned, but this depends on local clinical delirium detection practices and extent to which clinical staff members are trained. In contrast to the 7% figure quoted by Koch *et al.*, 42% of delirium episodes in the ENGAGES trial were detected by chart review.

It is asserted that EEG-guidance in the ENGAGES study was ineffective. We agree with Koch *et al.* regarding the shortcomings of the BIS monitor, especially in relation to older adults. For this reason, in ENGAGES, we expended time and energy teaching anesthesia clinicians to interpret the EEG waveform during anesthesia, and not to rely on BIS readings. This is in contrast to BAG-RECALL, CODA and SuDoCo, where clinicians relied on BIS readings. In the ENGAGES trial, clinicians were specifically taught to recognize suppression from raw waveforms (11), and did not depend on the BIS-derived burst-suppression ratio. Koch *et al.* highlight that in the EEG-guided group in ENGAGES,

Page 2 of 3

Wildes and Avidan. Cognitive dissonance and anesthesia research

Table 1 Questions related to methods and results of the CODA trial

Different calculated result for the primary outcome POCD by Fisher's exact test	We calculated the odds ratio for the primary outcome as 0.66 (0.44–1.003, P=0.059) by Fisher's exact test. This is different from the odds ratio for POCD reported in the manuscript [either 0.62 (0.39–0.97, P=0.02) or 0.67 (0.32–0.98, P=0.025)] (2)
Retrospective registration	CODA was retrospectively registered in September 2007, while enrollment began in January 2007
Study to justify sample size was not yet published	The study referenced to justify the sample size (7) was published a year after enrollment to the CODA trial began.
Power calculation did not use actual referenced incidence	In the study referenced to justify the sample size, the incidence of POCD at 3 months was 12.7% (7); in the power calculation a 30% incidence was assumed
Low MAC	The mean age- and N_2O-adjusted MAC was 0.57 (SD =0.29) in the BIS-guided group (2). Based on this, it is likely that many patients received low MAC for the majority of the surgery
Nitrous oxide (N_2O) as sole agent	N_2O was administered to 63% of the patients who received inhaled agents with an average concentration of 63% (2). In a 60-year-old, 63% N_2O is 0.7 MAC (8), which implies some patients might have received N_2O as their sole hypnotic agent
Duration of BIS <40	The median BIS in the control group was 36, with a mean duration of anaesthesia of 2 hours (2). This is difficult to reconcile with a median duration of BIS <40 as only 26 minutes
Age distribution	The age range for the trial was 60 to 90. The mean age was 68 (SD =8) (2). It is hard to understand the age distribution of patients based on these results
Duration of surgery	The median duration of surgery was <2 hours, based on a mean duration of anesthesia of 2.05 hours (2). Thus, fewer than half likely met the target duration of surgery

MAC, minimum alveolar concentration; POCD, postoperative cognitive dysfunction; BIS, bispectral index.

the proportion of time with BIS <40 was higher than it was in the BIS-guided group in CODA (2). However, several relevant results in the CODA (2) trial are difficult to understand (*Table 1*). Furthermore, in SuDoCo (3), the proportion of time with BIS <40 in the BIS-guided group was substantially higher than it was in ENGAGES (1). Thus, lower proportion of time with BIS<40 does not appear to provide a satisfactory explanation for decrease in delirium incidence. To address the question of whether only partial effectiveness of EEG guidance might have masked the benefit of the intervention, the ENGAGES investigators conducted three sensitivity analyses to determine if the results would change when discarding those patients in the EEG-guided group with the most EEG suppression, lengthiest periods with BIS <40, or the highest anesthetic concentrations; this still yielded no significant difference in delirium incidence between the guided and usual care groups (1).

It is intimated that important results were not prominently reported in the ENGAGES manuscript. Koch *et al.* note that in ENGAGES, EEG suppression and duration of BIS <40 were associated with delirium (1). We

therefore assume that Koch *et al.* do actually believe that our assessments of delirium were reliable. But association does not establish causation. Indeed, this was the core question ENGAGES was designed to study. It remains possible that avoiding EEG suppression and BIS <40 reduces postoperative delirium. However, ENGAGES suggests that this can only be a small effect. ENGAGES clarifies that EEG suppression in response to volatile anesthetics primarily reflects pre-existing brain vulnerability, rather than representing a physiologic state that has a strong causal link to postoperative delirium. As an analogy, ST depression with exercise testing reveals pre-existing myocardial vulnerability. ST depression during the test is unlikely to be the cause of subsequent myocardial infarctions.

Finally, ENGAGES investigators are berated for not emphasizing the post-hoc finding that 30-day mortality was lower in the EEG-guided group (1). Accepting this criticism requires cognitive dissonance, considering that previously Koch *et al.* suggested that EEG guidance was ineffective. However, on the basis of this criticism, we must now assume that Koch *et al.* believe that EEG guidance in the ENGAGES

trial was in fact effective, that EEG guidance did reduce delirium [but that it was unrecognized because the European Society of Anaesthesiology detection guidelines (12) were not expressly followed], and that the delirium detection deficiency was exposed by the higher 30-day mortality in the usual care group. The much likelier and more parsimonious explanation is that the difference in 30-day mortality was a chance 'significant' finding, which might be refined when we report 1-year outcomes.

Koch *et al.* express grave concern that ENGAGES obscures "the tremendous benefits of intraoperative EEG guidance in anesthesia, several of which were also confirmed in this [ENGAGES] study." We are also proponents of monitoring the target organ of general anesthesia with EEG. But in contrast to Koch *et al.*, we believe that astute readers will consider us disingenuous if we trumpet "tremendous benefits" that remain hypothetical. Clinical trials provide evidence for or against hypotheses, and should not be portrayed as definitive: it is wrong to consider trials as providing confirmation. When authoritative organizations prematurely produce strong recommendations in guidelines, and when they falsely claim in "best practice" documents that "level 1 evidence" for a practice (such as EEG-guidance of anesthesia to prevent delirium) exists (13), they undermine their authority, and instead appear merely authoritarian.

Acknowledgments

Footnote

References

1. Wildes TS, Mickle AM, Ben Abdallah A, et al. Effect of Electroencephalography-Guided Anesthetic Administration on Postoperative Delirium Among Older Adults Undergoing Major Surgery: The ENGAGES Randomized Clinical Trial. JAMA 2019;321:473-83.
2. Chan MT, Cheng BC, Lee TM, et al. BIS-guided anesthesia decreases postoperative delirium and cognitive decline. J Neurosurg Anesthesiol 2013;25:33-42.
3. Radtke FM, Franck M, Lendner J, et al. Monitoring depth of anaesthesia in a randomized trial decreases the rate of postoperative delirium but not postoperative cognitive dysfunction. Br J Anaesth 2013;110 Suppl 1:i98-105.
4. Whitlock EL, Torres BA, Lin N, et al. Postoperative delirium in a substudy of cardiothoracic surgical patients in the BAG-RECALL clinical trial. Anesth Analg 2014;118:809-17.
5. Grimes DR, Bauch CT, Ioannidis JPA. Modelling science trustworthiness under publish or perish pressure. R Soc Open Sci 2018;5:171511.
6. Prasad V, Vandross A, Toomey C, et al. A decade of reversal: an analysis of 146 contradicted medical practices. Mayo Clin Proc 2013;88:790-8.
7. Monk TG, Weldon BC, Garvan CW, et al. Predictors of cognitive dysfunction after major noncardiac surgery. Anesthesiology 2008;108:18-30.
8. Lerou JG. Nomogram to estimate age-related MAC. Br J Anaesth 2004;93:288-91.
9. Avidan MS, Maybrier HR, Abdallah AB, et al. Intraoperative ketamine for prevention of postoperative delirium or pain after major surgery in older adults: an international, multicentre, double-blind, randomised clinical trial. Lancet 2017;390:267-75.
10. van Velthuijsen EL, Zwakhalen SM, Warnier RM, et al. Psychometric properties and feasibility of instruments for the detection of delirium in older hospitalized patients: a systematic review. Int J Geriatr Psychiatry 2016;31:974-89.
11. Wildes TS, Winter AC, Maybrier HR, et al. Protocol for the Electroencephalography Guidance of Anesthesia to Alleviate Geriatric Syndromes (ENGAGES) study: a pragmatic, randomised clinical trial. BMJ Open 2016;6:e011505.
12. Aldecoa C, Bettelli G, Bilotta F, et al. European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium. Eur J Anaesthesiol 2017;34:192-214.
13. Berger M, Schenning KJ, Brown CHt, et al. Best Practices for Postoperative Brain Health: Recommendations From the Fifth International Perioperative Neurotoxicity Working Group. Anesth Analg 2018;127:1406-13.