**Letter to the Editor**

Title: In search of Delayed Neurocognitive Recovery

Author Information:

J Robert Sneyd, MD FRCA, University of Plymouth, UK

Lisbeth A Evered, PhD, Weill Cornell Medicine, New York, USA; Melbourne Medical School, Australia

Corresponding Author:

Professor J Robert Sneyd, Coldrenick, Menheniot, LISKEARD PL14 3RQ, UK

+44 7870 271531

robert.sneyd@pms.ac.uk

Clinical trial number and registry URL, if applicable; N/A

Prior Presentations: None

Acknowledgments: None

Word and Element Counts: Number of words 557; number of Figures 0; number of Tables 0; number of Appendices 0, number of Supplementary Digital Files 0

Abbreviated Title (Running Head): N/A

Summary Statement: N/A

Funding Statement: “Support was provided solely from institutional and/or departmental sources.”

Conflicts of Interest: None

Joosten and colleagues are to be congratulated on their deployment of technically complex closed-loop systems to support patients during anesthesia and surgery.1 The possibility of experiencing impaired neurocognitive function in association with a surgical episode is a concern to patients and those who care for them. It makes sense to establish whether changes in clinical technologies might diminish or abolish these unwelcome syndromes.

Nevertheless, we have concerns about the Primary Outcome Measure and its analysis.

*On clinical trials.gov*:  https://clinicaltrials.gov/ct2/show/NCT03148730 the Primary Outcome Measure is “**Incidence** of postoperative cognitive dysfunction”. This implies a definition of postoperative cognitive dysfunction. The authors’ chose the Montreal Cognitive Assessment score (maximum 30) so we need to consider what is a meaningful change. Reduction by a single point is very unlikely to be clinically significant and certainly does not represent a reduction of >1 SD below normative published data.2 A decline of two points? Five points? Falling from above 26 to below 26? Then each patient either does or does not have postoperative cognitive dysfunction and the incidence would be the proportion of patients with the condition in each of the 2 treatment groups. Then we can compare the incidence following control and closed-loop treatments.

*In the paper* the primary outcome was the **change** of the cognition score. This is an important alteration because incidences, group differences and individual **change** are not the same thing… The authors have concluded there is a difference in cognitive outcome based on a screening test (the Montreal Cognitive Assessment) and ignored the results of the more robust cognitive assessment tools which returned no difference between groups.

Not all the primary outcome data is shown. The Montreal Cognitive Assessment has been treated as parametric (normally distributed) in the power calculation but is (partially) reported as non-parametric in the results. Baseline scores are set out in Table1. We looked for but cannot find any summary of the data after baseline, the values at one week and 90 days are not reported. Instead we are given what is probably the median (it is not defined) and the interquartile ranges of the change from baseline. It would be helpful to see the raw data perhaps presented as a scattergram with lines to show the individual trajectories. In addition, summary statistics i.e. the median and interquartile ranges of the scores at one week and 90 days for each of the treatment groups.

 Regarding the analysis and statistical significance we note the 95% confidence intervals of the differences include zero at both one week and at 3 months? How can these results be significant?

In addition the post-hoc sensitivityanalysisshowed no difference between the treatment groups when the absolute values were analysed. What was the reason for the decision to use change in scores from baseline as the primary analysis? Was any statistical correction made for multiple testing? Bonferroni correction or similar?

 The abstract will be widely read. The conclusion is not based on the pre-specified primary outcome measure. In addition, none of eighteen secondary outcome measures defined on clinicaltrials.gov were reported (understandable because of space constraints). However, the authors include three additional measures of which two are similar to, but not the same as, those pre-specified.

Overall we are worried that Joosten et al. have overstated their findings. An alternative interpretation is that the high-tec technique made no difference to outcome.

References

1. Joosten A, Rinehart J, Bardaji A, Van der Linden P, Jame V, Van Obbergh L, Alexander B, Cannesson M, Vacas S, Liu N, Slama H, Barvais L: Anesthetic Management Using Multiple Closed-loop Systems and Delayed Neurocognitive Recovery: A Randomized Controlled Trial. Anesthesiology: The Journal of the American Society of Anesthesiologists 2020; 132: 253-266

2. Evered L, Silbert B, Knopman DS, Scott DA, DeKosky ST, Rasmussen LS, Oh ES, Crosby G, Berger M, Eckenhoff RG, Nomenclature Consensus Working G: Recommendations for the Nomenclature of Cognitive Change Associated with Anaesthesia and Surgery-2018. Anesthesiology 2018; 129: 872-879