DEPARTMENT OF
Anesthesiology
UNIVERSITY OF WISCONSIN
SCHOOL OF MEDICINE AND PUBLIC HEALTH

Dear Dr. Merino,

18th of May, 2019

Manuscript ID BMJ-2019-049203R1 "Effect of Major Surgical and Medical Admissions on Cognitive Trajectory: 19-year follow-up of the Whitehall II Longitudinal Prospective Cohort Study"

We would like to thank you for the opportunity to respond to the comments of the committee and reviewers and your interest in our paper. Please see below a point by point response to all comments.

We would like to thank you for your consideration and look forward to hearing from you.

Yours sincerely,

Dr. Robert D. Sanders BSc MBBS PhD FRCA (on behalf of all co-authors)
Assistant Professor,
Department of Anesthesiology,
University of Wisconsin, Madison

COMMENTS FROM THE EDITORS

1. We were interested in the question. The possible effect of general anesthesia on cognition has been widely studied and remains controversial. Can the authors say something about anesthesia in their 'major surgery' category ? Is that the essential factor, as minor surgery did not show any cognitive decline?

OUR RESPONSE: This is a reasonable question as major surgery is typically conducted under general anesthesia. Unfortunately, HES does not contain sufficient data on anesthesia (or type of anesthesia) to address this question further. It is possible that anesthesia is not a major factor, based on results from prior meta-analyses and recent randomized controlled trials (ENGAGES[1] and STRIDE[2]). However, we acknowledge that none of these data are definitive.

We have added the following to the discussion as a limitation:
"HES also lacks data on the type of anesthesia administered limiting our ability to comment on the role of anaesthesia in any long-term cognitive change."

2. The statistical analysis is complex but very well described. Our statistician did not have major concerns but did have two questions that you may address in the methods: why two constants in the regression equation? Why include diabetes and smoking in the models?

OUR RESPONSE: The two constant terms represent the overall intercept for the population (interpreted as the cognitive score at the median age in this study sample), and the subject-specific intercept (interpreted as a constant offset from the population for each subject, modelled as a random effect). Diabetes and smoking were included because of their known associations with cognitive decline, though as we report the inclusion of these (or other) covariates had little impact on the estimates of the primary outcomes, likely because these effects are already captured by the random effects structure of our primary model.

3. One of our editors wanted additional clarifications. Were the patients planned surgical admissions? Did they undergo emergency surgery?

OUR RESPONSE: Thank you. We included all surgeries in this analysis but we have now included a supplemental analysis excluding surgeries identified as emergencies based on the OPCS codes. This sensitivity analysis did not affect the conclusions (please see Table 4).

4. The strength of the association is difficult to appraise, as these psychometric scores are not easily translated into an intuitive QoL or disability score. Could you provide more information on the magnitude and clinical importance of the decline?

OUR RESPONSE: We agree, decline in cognitive function is not straightforward to interpret as there are no clinical criteria to judge accelerated cognitive decline. We addressed this issue in two ways: by converting cognitive decline to an aging metric ("years of cognitive aging") and providing the observed decline on the cognitive score ("0.17 fewer words in the combined verbal/semantic fluency test"). As QoL or disability were not the focus of our analyses we prefer not to convert the change scores against another metric. In the revised manuscript we now give an example

explaining: “Hence we estimate a subject who incurred surgery at the median 67 years and 5 months of age would incur a cognitive hit that on average would result in a cognitive age of 67 years and 10 months of age.” We also explain that the magnitude of the effect is within one standard deviation of the annualized age-related decline. Previously we have argued that deviations less that 1.5 standard deviations lack clinical importance[3] and hence the changes associated with surgery are likely to be subtle.

In response to Reviewer 2 we have also added an analysis that identifies those at risk of substantial cognitive decline postoperatively. We have used a threshold of 1.96 standard deviations for congruence with the postoperative cognitive decline literature.[4]

Reviewer: 1

1.While this manuscript addresses a clinically important issue, I worry that a sizeable proportion of the “major” surgeries likely represent relatively minor procedures, which would not be expected to have much effect on cognition. Also, the population is relatively young, so susceptibility to the deleterious effects of major surgery will be diminished. The statistical modeling is very complex, so the manuscript should be carefully reviewed by a biostatistician with expertise in longitudinal analyses.

OUR RESPONSE: Thank you. We discuss your concerns about the classification of procedures and age of the participants below.

2. This is a well-written manuscript that presents results from a longitudinal analysis of data from the Whitehall II study to determine whether incident major surgical admissions induce long-term changes in cognitive trajectory. The authors found that major surgery is associated with a small, long-term change in the cognitive trajectory that is less pronounced than that for major medical admissions. The manuscript has several strengths, including: its large size and duration of follow-up, assessments of cognition before and after hospitalizations, and translation of effect estimates to equivalent years of cognitive aging.

OUR RESPONSE: We thank the reviewer for these positive comments.

3. The authors used a very low bar for defining a major surgery, namely an overnight hospital stay coupled to an ICD-10 code. There is no requirement that the surgical procedure occurred in an operating room. It is highly likely that a sizeable proportion of the “major” surgeries represented relatively minor procedures. The authors should use a more conservative definition of major surgery that includes only procedures that occurred in the operating room, as recommended by: Kwok AC, Semel ME, Lipsitz SR, Bader AM, Barnato AE, Gawande AA, et al. The intensity and variation of surgical care at the end of life: a retrospective cohort study. Lancet. 2011;378(9800):1408-13. The authors should also provide a complete listing of the ICD-10 codes used to identify major surgery.

OUR RESPONSE: Thank you for raising this important issue and apologies that our methods here were unclear. In our submitted manuscript major surgery was defined based on the OPCS codes (which are analogous to CPT codes in US healthcare) not ICD-10 codes and the requirement that an

OPCS code was classified as surgery (after review by two authors of this manuscript: RDS and HJM). The length of stay also required at least a two-night stay (apologies this was unclear). The requirement for a two-night stay is equivalent to that used in several major randomized controlled trials and prospective cohort studies[5-8]. Please note a code for "conducted in an operating room" is not available in HES.

In order to further address the reviewer's concern, two authors (RDS and HJM) rated the OPCS codes based on the BUPA classification of major surgical procedures as implemented in the Surgical Risk Scale[9]. These authors have previously used this BUPA procedure scale to rank procedures and showed that major surgeries incurred higher postoperative mortality than more minor procedures[10] supporting both the classification and our interpretation. BUPA major surgeries must be conducted in an operating room. We then undertook sensitivity analysis using this refined definition of major surgery as well as a minimum two-night length of stay (referred to as "BUPA major" in the revised manuscript) and our results are virtually unchanged as shown in Table 4.

- Primary outcome (years equivalent of cognitive impact): 0.35 (0.0077 to 0.73)
- BUPA Major (years equivalent of cognitive impact): 0.32 (-0.096 to 0.72)

We have also included a list of OPCS codes as requested in Appendix 1.

4. It is very odd that the number of hospitalizations for "major" surgery is so much larger than the number of hospitalizations for other medical conditions (2932 vs. 2114). The opposite is true in the US. This raises serious concerns about the authors' ascertainment of hospitalizations and/or use of ICD-10 codes.

OUR RESPONSE: We are confident that we have classified major surgery appropriately and have used a similar classification to "match" major medical events. Some of the discrepancy might be related to linkage (we treated non-surgical admissions in proximity to a surgical admission as part of the surgical event, to avoid underestimating the effect of surgery) and exclusion of short stays. In the original database prior to linkage and exclusion of brief stays, only 34% of the admissions involved surgery.

5. The study population was relatively young, with a mean age of only 55.5. Susceptibility to the adverse consequences of major surgery and other hospitalizations will be much greater in an older population. At the very least, the authors should repeat their analyses in the subgroup of participants who were 65 years or older.

OUR RESPONSE: We based our analyses on evidence of cognitive decline even in participants aged 45-49 years[11] and an emerging consensus on cognitive decline leading to dementia being a process that unfolds over 15 to 20 years. Participants were 45 to 69 (mean 55) years at the first cognitive wave of data collection; they then underwent up to 19 years of cognitive follow up. Nonetheless there are significant prior data suggesting the elderly may be most vulnerable to the cognitive effects of surgery. A possible major limitation of those prior data is that they did not account for previous cognitive trajectory. Our approach is novel in that participants were followed starting in midlife.

B6/319 Clinical Science Center 600 Highland Avenue Madison, WI 53792-3272
608/263-8100 FAX 608/263-0575

In order to address the reviewer's concern we now include an analysis where we excluded individuals who had surgeries prior to age 65. Interestingly this exclusion reduced our estimate of the impact of surgery rather than increasing it. We've added some discussion of this finding in the revised manuscript, as described below.

Results
"A sensitivity analysis based on the age of exposure to surgery was conducted based on the hypothesis that older persons would be more vulnerable to the cognitive effects of surgery. We excluded subjects who had surgeries prior to age 65 such that all surgeries remaining in the analysis occurred after age 65. These data surprisingly showed a reduced impact of surgery at older ages (Table 4)."

Discussion
"One surprising result from the sensitivity analyses was that in people who had no surgeries prior to age 65, subsequent surgeries were not associated with any cognitive decline. This result may reflect a selection bias in type of surgery at different ages, or because longer hospital stays in younger patients are associated with surgical complications, or other factors that contribute to a measurable impact. Alternatively, it could support the hypothesis that surgery at younger ages is a marker poorer health or more aggressive disease that are associated with subsequent cognitive decline. Our study lacks data to resolve these alternatives, however, we emphasize that the mean effects in both age groups were smaller than those seen to be clinically significant in prior studies"

6. The statistical modeling is very complex, so the manuscript should be carefully reviewed by a biostatistician with expertise in longitudinal analyses.

OUR RESPONSE: Thank you, we are confident that the statistical analyses are sound as the authors have considerable expertise in this domain.

7. Table 1 is very difficult to understand, and the right column of results is particularly unclear. The categories are rather nonspecific, with the vast majority of surgeries listed as "Other".

OUR RESPONSE: We had access to hospital/surgery information from national health registers that included surgeries that occurred without patients necessarily returning for cognitive follow-up afterwards (as part of the Whitehall II study). We intended Table 1 to demonstrate that the surgeries for which we have subsequent cognitive data (right column) were of a similar makeup to those which occurred at any time. We've added text to explain Table 1, as follows.

"In Table 1, major surgical events for the entire group (left column) are compared to those for which cognitive follow-up was available (right column), organized according to higher risk surgical categories that have plausible associations with cognitive outcomes (cardiac, thoracic, vascular, and intracranial surgeries)[4,12]. The proportions were similar, suggesting that cognitive follow-up was available for a representative cross section of surgeries in the study population."

8. Figure 1 does not provide sufficient details. Specifically, the authors do not indicate the timing of the hospital admissions relative to the cognitive assessment intervals. It is likely that some of these intervals included more than 1 hospitalization and potentially both surgical and medical hospitalizations. How were the authors able to distinguish the effects of these different hospitalizations within a single interval?

OUR RESPONSE: We show the relative timing of admissions and cognitive assessments in Figure 2B. We agree that these intervals could include medical and surgical admissions and we have modelled the cumulative number of admissions to account for this. Our analysis presumes that the impacts of admissions on cognition are cumulative: that is, that the impact of two surgical admissions during the same interval would have twice the impact of one. Most intervals from most subjects contained zero or one admission (96%).

9. Figure 3 is also very difficult to understand, with rather unusual color patterns and discontinuity between the left and right sides of the figure.

OUR RESPONSE: Please find a modified version of figure 3 below. Apologies it was unclear. Specifically we have modified the color scheme (using a palette suggested by Wong, B. (2011). Color blindness. *Nature Methods*, *8*(6), 441.[13]) and the discontinuity. We hope that this is clearer.

Minor Comments

10. Table 2: p-values should not be reported < 0.001.

OUR RESPONSE: Apologies, this has been corrected.

12. More complete details should be provided for the Framingham cardiovascular disease risk score.

OUR RESPONSE: The Framingham cardiovascular disease risk score was calculated based on the paper by D'Agostino et al., Circulation 2018[14]. This is now appropriately referenced. Apologies for this oversight.

13. What do the authors mean by "baseline": "… baseline adjustments for numbers of surgical admissions, medical admissions, and stroke admissions (including events that occurred after cognitive follow-up but during the range of years analyzed), …"

OUR RESPONSE: In this context "baseline" refers to differences in cognition that are modelled as constant throughout the study (and therefore apply prior to any admissions). These adjustments are important because without them we could erroneously attribute differences in starting cognition to a subsequent admission. Biologically, these adjustments could reflect correlations of overall health with cognition prior to the interval we study. Our approach is analogous to an experimental pre-post design that includes a regression term for pre-intervention differences.

This has been clarified as follows in the manuscript:

"Since our focus was on changes after surgery, we also included baseline adjustments for numbers of surgical admissions, medical admissions, and stroke admissions (including events that occurred after cognitive follow-up but during the range of years analyzed), and for two-way interactions between them in order to ensure that the association observed with surgery is not attributable to differences by subgroups in preoperative cognitive function."

14. Information should be provided about the amount of missing data

OUR RESPONSE: We have now clarified this by changing the title of the multiple imputation section to "Missing Data". Apologies, that this was unclear. Missing data only occurred for certain categories of patient characteristics listed as missing in Table 2 (race, education, smoking status), for which we imputed data when those variables were used, and lack of cognitive follow-up (mean # of assessments is also noted in Table 1). Additionally, Table 1 shows how many surgical admissions had cognitive data after the admission.

"Missing Data
In order to account for occasional missing demographic data, we generated 100 imputed datasets using the R package MICE[15], fit MCMC models to each, and computed credible intervals for each fixed effect across all of the imputed models[16]."

Reviewer: 2

GENERAL COMMENTS:
1.The authors are to be congratulated on applying their well-honed expertise in the area of Perioperative Cognitive Disorders (PND) to the Whitehall II database, one that has been validated for reliability of the cognitive assessment because "556 subjects underwent retesting within 3 months of their initial assessment with good test-retest reliability." In particular the authors seem to have overcome some of the major methodological issues that have plagued the reports from previous observational studies of defined databases. In particular the very large cohort size, the use of "stroke" as a positive control, and the prolonged duration of follow-up overcome limitations in previous studies.

The use of the "step-change" in their choice of a Bayesian model is appropriate because the events that precipitate and follow a hospital admission for a major medical condition or major surgical procedure do impose a "new condition" (possibly inflammation-based) capable of impairing function.

OUR RESPONSE: Thank you for these positive comments.

2. The choice to quantify the disorder in terms of an enhanced age-related cognitive trajectory decline has been used in past Whitehall II database interrogations. However, as the "additional months/years" of decline apply to an average patient at age 67.4 years, I am not sure how meaningful using these data would be in explaining possible surgical risk to an 80 year-old. As this is not the typical manner in which prevalence of the PND conditions (± interventions) are reported, I would ask the authors to also cater to the existing PND field in which a binary outcome is used that describes the likelihood of crossing a threshold to having the new condition. However, the

problem with this "usual PND investigational approach" is that there is lack of uniformity regarding the diagnosis of the PND conditions apart from delirium whose features are listed in the DSM V. While I understand the authors' choice to consider the entire spectrum of cognitive assessment and not reduce it to a binary outcome, I would have benefitted further with an additional description of the outcome using the binary approach. By also reporting on the binary outcome the authors' may discover that they have sufficient granularity to comment on risk factors that modulate the incidence of the PND disorders. While I appreciate that additional work that will be needed, I do think that it will make this report more meaningful to those working in the field.

OUR RESPONSE: Thank you for this insightful comment. We have approached this request for a binary outcome using the reviewer's suggestion (establishing a predicted cognitive trajectory and then calculating the odds of ratio of being greater than 1.96 standard deviation from the predicted value after surgery). This also helps address the editors' questions above regarding clinical importance[3]. The binary outcome approach shows an approximate doubling in the odds of an extreme decline with this effecting 5.5% of participants. We agree that this is important to report and we have modified our conclusions to express while the mean effect is modest, there is a small increase in the risk of severe decline following surgery.

SPECIFIC COMMENTS
Abstract/Participants:
3. While it is stated that one could have up to 5 cognitive assessment, was there a minimum # of cognitive assessments that an enrolled pat could have?

OUR RESPONSE: No. In the whole cohort a patient was included as long as they had at least one cognitive assessment. They were excluded if they lacked any cognitive data. We also conducted a sensitivity analysis requiring at least 4 cognitive assessments (n=4916) and find similar results to the whole cohort.

4. Abstract/Participants: "to" missing following "linkage" and prior to "Health Episode…"

OUR RESPONSE: Thank you, this has been corrected.

5. Abstract/Results: Define "Minor surgery"

OUR RESPONSE: In the original manuscript, "minor surgery" referred to surgery that did not involve at least a two-night hospital stay. In response to concerns from reviewer 1 (comment 3), we added a sensitivity analysis focusing on major surgeries as defined according to BUPA guidelines and as implemented in the Surgical Risk Scale. We have now removed mention of minor surgery from the abstract. To avoid confusion between these groupings, we have clearly defined the categories of "minor surgery" and of "BUPA major" surgery in the Methods section.

6. Introduction: For those concerned about Perioperative Neurocognitive Decline, it is definitely the "step-change" in cognitive trajectory (worsening of the expected cognitive decline either through an early inflection point or a change in the trajectory). In the Introduction, there needs to

be more about the "binary" nature of the condition (presence of a > 2SD change of the cognitive score).

OUR RESPONSE: We have now added an explanation to the introduction to facilitate understanding of both the mean effect and also the binary (extreme decliner) analyses, as described below. Many thanks for pointing this out.

Introduction
"As a secondary outcome we developed a binary outcome of substantial cognitive decline, more analogous methodologically to prior studies of postoperative cognitive decline[10] and consistent with clinically important deviations from the age-related cognitive trajectory[20]. It allows some further correspondence to the prior literature and emphasizes cognitive changes that may impact quality of life."

7. I appreciated the discussion on limitations of existing observational studies that report on the interrogation of an already collected database to elicit the longitudinal changes that are then linked to a medical or surgical "incident." When flaws of previous studies are highlighted it is important to demonstrate how the existing study avoided these flaws without creating any new ones.

OUR RESPONSE: We now highlight more clearly the strengths of our study, as follows.
"We aim to address these concerns using cognitive data on 7,532 persons, investigating whether incident major surgical admissions induce long-term changes in cognitive trajectory, using five waves of cognitive assessments spanning approximately 20 years with adjustment for major medical admissions."

8. In the following sentence there is a lack of punctuation between the main and dependent clauses.
"Longer life expectancy implies increasing number of surgeries in older adults hence better understanding of the extent of any cognitive "hit" following surgery is urgently required."

OUR RESPONSE: Thank you, this has been corrected.
"Longer life expectancy implies increasing number of surgeries in older adults, hence better understanding of the extent of any cognitive "hit" following surgery is urgently required."

9. Methods/Exposures: It is stated that "To limit effects of intrahospital transfers, we linked together any admissions within 14 days."
How would the data be handled if a patient had a major surgical procedure but was discharged and then re-admitted within the 14-day period with a postop complication, e.g., sepsis and/or stroke? As stated above it appears that the effect of the medical complication of a surgical complication is now lumped together with the major surgical procedure.

OUR RESPONSE: We have now clarified our methods using this example to ensure the reader understands we weighted our approach to detect an effect of surgery. The reviewer is correct that an admission within 14 days that resulted from surgery would be considered part of the cognitive burden of surgery (assuming that this admission (e.g. for sepsis) is a complication of surgery). As

such, our analyses were made sensitive to finding a cognitive effect of surgery. We consider that this makes identification of only a small mean effect of surgery more important.

In the discussion we now state:
"We have not sought to underestimate the impact of these surgery-related events on the cognitive trajectory, however: a stroke occurring within two weeks of surgery in this study would have been labelled as a surgical event and hence contributed to any decline associated with surgery."

10. From the following sentence, I presume that every subject had a minimum of one cognitive assessment.
"Admissions during the study period but without cognitive follow-up were retained for use in adjusting baseline cognitive scores."

OUR RESPONSE: Yes.

11. Are data captured from Whitehall II subject that are admitted to hospitals outside of the NHS for England and Wales i.e., private or Scotland, Ireland or rest of EU?

OUR RESPONSE: HES covers public hospitals in England, Scotland, and Wales but not other countries.

12. To whom did the discarded HES linkages refer if not to the 7532 subjects?
"Out of 43,692 HES entries during the study period (for 7,532 subjects), 35,099 remained after linkage."

OUR RESPONSE: This statement refers to linking admissions within 14 days (mentioned in the previous paragraph) rather than association with study subjects (43,692 is the accurate number of HES entries for 7,532 subjects). The reasons for linking admissions this way were to: a) Ensure that surgical admissions followed by a non-surgical inpatient state were accounted for as surgeries, and b) to not count multiple admissions that were associated with the same event.

13. Based upon 5,110 HES entries occurred after the first cognitive assessment and prior to the last cognitive assessment in the study (2,932 surgeries, 2,114 medical admissions, and 64 strokes). I assume that 3,872 (i.e. 8982-5110) HES admissions were in subjects that had NO cognitive assessments and had no global cognitive score. If so, were these included in the analysis?

OUR RESPONSE: Persons without cognitive data were not included in any descriptions in the paper. The 3872 additional admissions occurred after the last cognitive assessment available for that subject (including any subjects who had only a single assessment). We were of course not able to use these to estimate cognitive changes after admission, but we did account for them in adjusting baseline cognitive scores. Those that are surgeries are also mentioned in the left column of Table 1.

14. How many HES entries did you link to subjects that had 2 or more cognitive assessments prior to the hospital admission?

OUR RESPONSE: Of the 5,110 admissions, 2092 were in subjects who had at least 2 cognitive assessments prior to any admissions.

15. Methods/Outcomes
The Whitehall database has in the past used an "offset in age-related cognitive trajectory" to provide context to their outcome findings (see Ref #22). This is not the way that most PND studies have been explored and reported. They set up a threshold effect describing whether subject had/had not experience sufficient decline to indicate that they had developed a new condition of postoperative cognitive decline. This binary outcome is the most common method of describing the impact of an intervention. Surely, you can use a cohort (n = 4,916) that had at least 4 cog assessments and was used for sensitivity analysis in order to determine the binary of effect/no effect null hypothesis.

OUR RESPONSE: See response above to comment 2. Thank you for this suggestion.

16. Methods/Covariates
From the sentence below it is still not clear whether there are patients with NO cognitive assessments.
"We also included the number of cognitive assessments for each participant as a covariate"

OUR RESPONSE: Apologies this was unclear. Every person included in the analyses had at least one cognitive assessment.

17. Statistical Model
In defending their use of Bayesian statistics, the authors state that the binary methods are "…not most useful or informative statistic in the context we study.."
Almost all the results that are reported from prospective RCT for interventions to prevent one or other aspect of the spectrum of PNDs are binary. Therefore, a "throw-away" line stating binary methods is not the most useful or informative requires further elaboration. Note that a binary outcome has been used in studies involving the Whitehall II database and cognitive assessment (for MMSE in Ozawa et al PMID: 26874911).

OUR RESPONSE: Apologies for the confusion here. We have reworded this section. We actually were critiquing statistical approaches that declare an absence or presence of effect of a predictor based on an arbitrary statistical threshold ("$p<0.05$") rather than interpreting the magnitude of effect according to uncertainty (confidence intervals) and clinically relevant interpretations of the outcome variable. We did not intend to criticize binary outcome measures, however we recognize how poorly this was worded. We removed the word binary and now state:

"We chose a Bayesian approach because we feel that an "effect/no effect" judgment based on a null hypothesis test is not the most useful or informative statistic in the context we study, and related p-values are often misinterpreted[17]"

18. Reference needed for "We separated medical admissions for stroke events due to the expected substantial cognitive impact following stroke."

OUR RESPONSE: Now added, thank you.

19. Results
Including subjects with only one cognitive assessment appears meaningless because one could not establish their individual cognitive trajectory except through the use of population statistics.

OUR RESPONSE: Linear mixed effects models as we employ are adept at handling missing outcome data. Nonetheless we agree with the reviewer's logic and hence we conducted a sensitivity analyses, with a minimum of four cognitive assessments. Importantly this sensitivity analysis showed similar effects.

20. Discussion: Concerning the following sentences it seems that there are more than one "Schulte et al" studies. In this case please provide the citation for the other.
OUR RESPONSE: The two cohort analyses were contained within the same publication. We have reworded the sentence to add clarity and referenced appropriately, as follows:

"More recently, Schulte et al. found that incident surgery during 8 years of cognitive follow up was not associated with decline using 4 waves of data (n=431)[18]. In a larger cohort, the same authors suggested that surgery in the prior 20 years was associated with cognitive decline over the subsequent 8 year period[18], however this design cannot exclude that this effect may be attributable to lower cognitive performance prior to surgery (**Figure 3**)."

21. Regarding the following statement
"Fourth, all events occurred over the same period as the cognitive assessments allowing us to test how incident admissions affect cognition."

I understand that this can apply to patients that have bot a pre and postop assessments but from the methods it is not clear that this was the case. Can you please resolve the confusion for me.

OUR RESPONSE: Our point here is that in the vast majority of persons (98%) there was a preoperative cognitive data point. Some prior studies included surgery data from before any period of cognitive follow up. Hence compared to those prior studies we have a much stronger baseline. We have now conducted a sensitivity analysis just including subjects with at least one cognitive assessment prior to surgery. Our results did not change based on these findings (mean effect of surgery = 0.34 yrs). We have added these results to Table 4.

22. Figure 3: This was difficult to interpret. I cannot view the label on the ordinate because the line marker numbering occludes this.

OUR RESPONSE: Apologies this occurred due to the upload process to BMJ. We will ensure this does not happen with this revision.

B6/319 Clinical Science Center 600 Highland Avenue Madison, WI 53792-3272
608/263-8100 FAX 608/263-0575

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. Sieber FE, Neufeld KJ, Gottschalk A, et al. Effect of Depth of Sedation in Older Patients Undergoing Hip Fracture Repair on Postoperative Delirium: The STRIDE Randomized Clinical Trial. JAMA Surg 2018;153:987-95.
3. Singh-Manoux A, Kivimaki M. The importance of cognitive aging for understanding dementia. Age (Dordr) 2010;32:509-12.
4. Moller JT, Cluitmans P, Rasmussen LS, et al. Long-term postoperative cognitive dysfunction in the elderly ISPOCD1 study. ISPOCD investigators. International Study of Post-Operative Cognitive Dysfunction. Lancet 1998;351:857-61.
5. Devereaux PJ, Yang H, Yusuf S, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. Lancet 2008;371:1839-47.
6. Devereaux PJ, Mrkobrada M, Sessler DI, et al. Aspirin in patients undergoing noncardiac surgery. N Engl J Med 2014;370:1494-503.
7. Devereaux PJ, Sessler DI, Leslie K, et al. Clonidine in patients undergoing noncardiac surgery. N Engl J Med 2014;370:1504-13.
8. Mrkobrada M, Hill MD, Chan MT, et al. Covert stroke after non-cardiac surgery: a prospective cohort study. Br J Anaesth 2016;117:191-7.
9. Sutton R, Bann S, Brooks M, Sarin S. The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. The British journal of surgery 2002;89:763-8.
10. Venkatesan S, Myles PR, Manning HJ, et al. Cohort study of preoperative blood pressure and risk of 30-day mortality after elective non-cardiac surgery. Br J Anaesth 2017.
11. Singh-Manoux A, Kivimaki M, Glymour MM, et al. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. BMJ 2012;344:d7622.
12. Newman MF, Kirchner JL, Phillips-Bute B, et al. Longitudinal assessment of neurocognitive function after coronary-artery bypass surgery. N Engl J Med 2001;344:395-402.
13. Wong B. Color blindness. Nat Methods 2011;8:441.
14. D'Agostino RB, Sr., Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008;117:743-53.
15. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of Statistical Software 2011;45:1-68.
16. Zhou X, Reiter JP. A note on Bayesian inference after multiple imputation. The American Statistician 2010;64:159-63.
17. Ioannidis JP. The proposal to lower p value thresholds to. 005. Jama 2018;319:1429-30.
18. Schulte PJ, Roberts RO, Knopman DS, et al. Association between exposure to anaesthesia and surgery and long-term cognitive trajectories in older adults: report from the Mayo Clinic Study of Aging. Br J Anaesth 2018;121:398-405.

Information for submitting a revision

Deadline: Your revised manuscript should be returned within one month.

How to submit your revised article: Log into http://mc.manuscriptcentral.com/bmj and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Once the revised manuscript is prepared, you can upload it and submit it through your Author Center. When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) and Committee in the space provided. You can use this space to document any changes you make to the original manuscript and to explain your responses. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s). As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this as a supplemental file with file designation 'Revised Manuscript Marked copy'. Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision. Please include these items in the revised manuscript to comply with BMJ style (see: http://www.bmj.com/about-bmj/resources-authors/article-submission/article-requirements and http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists).

Items to include with your revision (see http://www.bmj.com/about-bmj/resources-authors/article-types/research):

1. What this paper adds/what is already known box (as described at http://resources.bmj.com/bmj/authors/types-of-article/research)

2. Name of the ethics committee or IRB, ID# of the approval, and a statement that participants gave informed consent before taking part. If ethics committee approval was not required, please state so clearly and explain the reasons why (see http://resources.bmj.com/bmj/authors/editorial-policies/guidelines.)

3. Patient confidentiality forms when appropriate (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality).

4. Competing interests statement (see http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests)

5. Contributorship statement+ guarantor (see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)

6. Transparency statement: (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/transparency-policy)

7. Copyright statement/licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

8. Data sharing statement (see http://www.bmj.com/about-bmj/resources-authors/article-types/research)

9. Funding statement and statement of the independence of researchers from funders (see http://resources.bmj.com/bmj/authors/article-submission/article-requirements).

10. Patient and public involvement statement

https://docs.google.com/document/d/1djgVLEUFtPQzLpf5HyuiFcMgrNJ_o7Yb8z73zgXxXYM/e

11. Dissemination plans: At the end of the paper please state how the results of your study have been (or will be) sent to patients and the public under the heading “Dissemination plans”. If you have prepared a lay summary eg for your funders, please include it in a supplementary file.

12. Patient confidentiality forms when appropriate

11. Please ensure the paper complies with The BMJ’s style, as detailed below:

a. Title: this should include the study design eg "systematic review and meta-analysis.”

b. Abstract: Please include a structured abstract with key summary statistics, as explained below (also see http://resources.bmj.com/bmj/authors/types-of-article/research). For every clinical trial - and for any other registered study- the last line of the abstract must list the study registration number and the name of the register.

Please report all outcomes that were listed in the trial registry, or explain that you will publish them elsewhere. Please clearly identify each outcome as primary, secondary, or post-hoc in the text, abstract, and any tables or figures. We expect authors to report prespecified outcomes. If outcomes in the trial registry have later been changed, please explain the reasons for the change and the dates of the change in the paper. You may report the changed outcomes, but we will

expect you to also report on the originally specified outcomes unless otherwise agreed with the handling editor for your paper.

Occasionally the outcomes that are prespecified in a trial registry do not match up with those included in the trial protocol. When there are discrepancies between protocol and registry specified outcomes, we expect the paper to report and interpret the registry specified outcomes. You may also report any protocol specified outcomes, but if you do please be sure to include the date of the protocol and the point at which each outcome was added to the protocol, and explain why the registry entry differed from the protocol and why the registry was not updated to reflect any protocol changes.

c. Introduction: This should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now.

d. Methods: For an intervention study the manuscript should include enough information about the intervention(s) and comparator(s) (even if this was usual care) for reviewers and readers to understand fully what happened in the study. To enable readers to replicate your work or implement the interventions in their own practice please also provide (uploaded as one or more supplemental files, including video and audio files where appropriate) any relevant detailed descriptions and materials. Alternatively, please provide in the manuscript urls to openly accessible websites where these materials can be found.

e. Results: Please report statistical aspects of the study in line with the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines http://www.equator-network.org/reporting-guidelines/sampl/. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:

i. For a clinical trial: Absolute event rates among experimental and control groups; RRR (relative risk reduction); NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000.)
ii. For a cohort study: Absolute event rates over time (eg 10 years) among exposed and non-exposed groups; RRR (relative risk reduction.)
iii. For a case control study:OR (odds ratio) for strength of association between exposure and outcome.
iv. For a study of a diagnostic test: Sensitivity and specificity; PPV and NPV (positive and negative predictive values.)
v. For a systematic review and/or meta-analysis: Point estimates and confidence intervals for the main results; one or more references for the statistical package(s) used to analyse the data, eg RevMan for a systematic review. There is no need to provide a formal reference for a very widely used package that will be very familiar to general readers eg STATA, but please say in the text which version you used. For articles that include explicit statements of the quality of evidence and strength of recommendations, we prefer reporting using the GRADE system.

Please report all outcomes that were listed in the trial registry, or explain that you will publish them elsewhere. Please clearly identify each outcome as primary, secondary, or post-hoc in the text, abstract, and any tables or figures. We expect authors to report prespecified outcomes. If outcomes in the trial registry have later been changed, please explain the reasons for the change and the dates of the change in the paper. You may report the changed outcomes, but we will expect you to also report on the originally specified outcomes unless otherwise agreed with the handling editor for your paper.

Occasionally the outcomes that are prespecified in a trial registry do not match up with those included in the trial protocol. When there are discrepancies between protocol and registry specified outcomes, we expect the paper to report and interpret the registry specified outcomes. You may also report any protocol specified outcomes, but if you do please be sure to include the date of the protocol and the point at which each outcome was added to the protocol, and explain why the registry entry differed from the protocol and why the registry was not updated to reflect any protocol changes.

f. Discussion: To minimise the risk of careful explanation giving way to polemic, please write the discussion section of your paper in a structured way. Please follow this structure: i) statement of principal findings of the study; ii) strengths and weaknesses of the study; iii) strengths and weaknesses in relation to other studies, discussing important differences in results; iv) what your study adds (whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses); v) meaning of the study, including possible explanations and implications for clinicians and policymakers and other researchers; vi) how your study could promote better decisions; vi) unanswered questions and future research

g. Footnotes and statements

Online and print publication: All original research in The BMJ is published with open access. Our open access policy is detailed here: http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse. The full text online version of your article, if accepted after revision, will be the indexed citable version (full details are at http://resources.bmj.com/bmj/about-bmj/the-bmjs-publishing-model).