PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A randomised controlled trial to investigate the relationship between mild hypercapnia and cerebral oxygen saturation in patients undergoing major surgery
AUTHORS	Wong, Clarence; Churilov, Leonid; Cowie, Dean; Tan, Chong; Hu, Raymond; Tremewen, David; Pearce, Brett; Pillai, Param; Karalipillai, Dharshi; Bellomo, Rinaldo; Weinberg, Laurence

VERSION 1 – REVIEW

REVIEWER	Ozan Akca, MD, FCCM University of Louisville Department of Anesthesiology & Perioperative Medicine Comprehensive Stroke Center & Neuroscience ICU
REVIEW RETURNED	24-Jan-2019

GENERAL COMMENTS	In this small RCT, Wong and colleagues assessed the effects of mild hypercapnia on intraoperative cerebral oximeter saturation in adult patients. Study subject is interesting. If the hypotheses are validated, they may impact future anesthesia practice. • It appears as the investigators accomplished the goal ETCO2 level by managing mechanical ventilation (by decreasing minute volume). This approach would result in differences in intrathorocic pressures between the two study groups. Such difference might have influenced the venous return and might have resulted in CO differences. Although such effect could be small, it needs to be discussed in the manuscript. • Did the investigators measure CO/CI and SV in the study? Or in any subgroup of patients? Were any intravascular volume monitors utilized in the study? • Intraoperative mean blood pressure results were not provided in the manuscript (only a statistically not significant result was provided). Please report the BP data. BP is both a perfusion marker and also can be considered as a marker of CO (i.e. BP=CO x SVR). • Did the investigators include any neuroaxial anesthesia/analgesia patients? If so, please report the nature of neuroaxial anesthesia/analgesia, and the type and amount of medications given through epidural/spinal routes. Neuroaxial anesthetic and analgesics may both alter intravascular volume status and hemodynamic responses. Additionally, these methods would provide superior analgesic approach, and may impact the tissue oxygenation levels due to better pain management. • Although, they were not statistically significant, there was about 10% difference in PaO2 levels and ~30% difference in narcotic analgesic doses between the groups. Both of these differences were in the favor of the hypercapnia group. Because both arterial oxygen and pain are established factors to influence tissue oxygenation, there needs to be discussion statements to address these

	differences between the groups. • How were the ETCO2 data reported? Did the authors first take the mean ETCO2 data per patient and then reported medians for the groups? Please describe in detail. • Blindness of anesthesiologists can be accomplished by blinding them only to the real time monitoring of cerebral oximeter data. In that case, the improved cerebral oxygen saturation's possible/potential effects on postoperative delirium would not bias anesthesiologists. If there will be a follow-up study focusing on postop delirium outcome, I strongly recommend the authors to keep the anesthesia team blinded to the cerebral oximeter data. • This reviewer suggests that those two "hypercapnia group" patients whom rSO2 data could not be collected, the authors may consider reporting the other outcomes' data (i.e. delirium, LOS, etc.) . • If one considers CO2 as a "drug" (i.e. medicine), the diminshing tissue oxygenation effect by time may suggest some type of tolerance. Although such tolerance can be due to metabolic contributions, this is a very interesting result and may mean further diminished effect size of hypercapnia in longer procedures. • How many times the arterial blood gas levels were measured throughout the study? If the answer is just one or two, then these results don't deserve a dedicated table, instead they can be included in table 1 or 2. • As far as this reviewer is aware, the effects of mild hypercapnia on cerebral oxygen saturation was reported twice before (by our group): 1) In a randomized controlled trial, [Ref: Effect of intra-operative end-tidal carbon dioxide partial pressure on tissue oxygenation. Anaesthesia. 2003 Jun;58(6):536-42.] the authors tested the effects of mild hypercapnia on tissue oxygenation in abdominal surgery patients. ; 2) In a crossover design a few years ago, [Ref: Manuscript reference #41 -] investigators used on-pump cardiac surgery setting to omit the contribution of inotropic effects (i.e. CO increase) of hypercapnia on tissue oxygenation. The authors may consider rephrasing their Introduction section to state the relationship between mild hypercapnia and rSO2 in patients undergoing surgery without pre-existing cerebral desaturation events was in fact studied before. • Please include duration of surgery data in the Abstract.

REVIEWER	Alex Fowler Queen Mary University of London London
REVIEW RETURNED	25-Jan-2019

GENERAL COMMENTS	BMJ Open Review A randomised controlled trial to investigate the relationship between mild hypercapnia and cerebral oxygen saturation in patients undergoing major surgery. Use of novel monitoring modalities, such as regional oxygen saturation or bispectral index are active areas of research in the care of patients undergoing surgery. This is particularly important given the recent acknowledgement that anaesthesia has wider impacts on neurophysiology and may contribute to long term cognitive impairment. It is well established that carbon dioxide (CO2) levels within blood causes changes to cerebral vasculature. This manuscript describes a well conducted randomised trial comparing normal CO2 to mildly raised CO2 on the change in regional oxygen saturation amongst adults having surgery.

	We hope the authors will find our comments valuable. Major comments 1. Introduction: This is concise and clearly states the rationale for the trial. It may be worth expanding on the recent focus on long term cognitive impact of anaesthesia and steps to minimise this. 2. Methods: The inclusion criteria differ between the ANZICTR entry and the manuscript [Age >18 years vs Age >50 years on ANZICTR]. Please clarify which of these were used as inclusion criteria, presumably >18 years as the range for age in table 1 includes patients aged 31 years. It is important that this discrepancy is explained in the manuscript (this is an important part of the CONSORT guideline). 3. Methods: How was a 15% change between control and intervention groups at the end of surgery judged to be clinically important? 4. Methods: These are otherwise very clear, and would enable study replication. The study has been reported broadly in line with CONSORT and SAMPL guidelines. Important, trial specific harms were considered (hyperkalemia particularly). 5. Results: The operations performed are not entirely clear, for example orthopaedic (commoner in TMH group) vs. HPB operations (commoner in TN group) may not be directly comparable in terms of risk for postoperative delirium and may have very different risk profiles. 6. Results: Two neurosurgical patients were included; was this for spinal surgery? Hypercapnia could be considered harmful by some during certain types of intra-cranial surgery. If patients undergoing intra-cranial surgery were included, it would be important to discuss this. 7. Results: What were the characteristics of patients who suffered postoperative delirium? It is difficult to unpick the association between TMH/TN and subsequent delirium with the data presented as they are currently, especially as the TN group seems to have higher risk patients (More ASA 3+, more HPB surgery). 8. Discussion: Important to highlight that a single time point CAM positive is not necessarily diagnostic of postoperative delirium. This is particularly important when considering that the length of stay between the two groups was the same, a surprising finding if 25% of patients in the TN group had postoperative delirium which is well known to extend length of stay. Use of a standard definition of postoperative delirium would provide a more reliable and interpretable outcome for any future work, as would baseline cognitive assessment to capture un-diagnosed cognitive impairment. 9. Discussion: Important to note that there was a change from baseline within both groups, and that this was a decrease in rS02 in the TN group, and an increase in rSO2 in the TMH group. This is especially important as the sample size was based on an assumption of no change in the control group, when there was an observed negative change.

	10. Results/Discussion: Looking at table 2; at 360mins after start of surgery, there was an average of 43.4% decrease in oxygen saturations in the TN group (sample size =2), while TMH stayed fairly static. This is an important finding; patients with maintained normocapnia have decreasing rSO2 over time, while targeted mild hypercapnia is associated with a stable rSO2 compared to baseline. Does this not meant that the interpretation should be subtly altered to state 'Mild hypercapnia is associated with a stable increase in rSO2 from baseline, while normocapnia is associated with a decrease in rSO2?'. To say that there is a larger increase over time in the hypercapnia group is unclear, as the change is actually the same but is statistically exaggerated by the sharp decrease in rSO2 in the normocapnia group. 11. Discussion: A recent publication in this area exploring the use of NIRS during cardiac surgery may be an important discussion point. (Rogers et al. Randomized trial of near-infrared spectroscopy for personalized optimization of cerebral tissue oxygenation during cardiac surgery. Br J Anaesth. 2017;119(3):384-393.) [I have no association with this group]. Minor comments 1. Consider consistent use of post-operative or postoperative; the tables use post-op while in the main text, there is no hyphenation. 2. Methods: In exclusion criteria there should be a comma between 'requiring one lung isolation' and 'liver transplantation'.

REVIEWER	David Blanco Universitat Politècnica de Catalunya
REVIEW RETURNED	30-Jan-2019

GENERAL COMMENTS	This report shows the results of an evaluation of the consistency between the CONSORT checklist you submitted and the information that was reported in the manuscript. Please, make the following revisions: • For CONSORT Item 6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"), to be consistent with the way you presented the study results for the primary outcome, please indicate in "Outcomes and Data Collection" section that the rSO2 was measured in the two hemispheres. I would also suggest that you include a subsection called Arterial blood gases (as you did in the protocol) where you indicate what gases were measured, when, which of these were considered secondary outcomes and which ones not. In this case, explain what the role of these variables was in the analysis – for example, you mention that Hb is a confounding variable in the discussion but you do not explain anything else in the Methods section. • For CONSORT Item 17a ("For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)"), please do not only report in the "Results / Secondary outcomes" section the results for each group, but also the estimated effect size and its precision (using 95% confidence intervals). Moreover, please do not include here the results for total haemoglobin if this variable is not listed as one of the secondary outcomes in the Methods section.

REVIEWER	Nousjka Vranken Maastricht University Medical Centre, Maastricht, the Netherlands
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. The authors conducted a very interesting study with robust methodology and submitted a well written manuscript. I do have some questions and comments that I listed below. 1. Why did the authors solely focus on postoperative delirium? Other postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation. 2. The authors suggest that mild hypercapnia results in higher intraoperative rSO2 values, resulting in a decreased occurrence of postoperative delirium. 3. The authors should provide further elucidation why an increase in cerebral rSO2 intraoperatively would be a wanted effect. 4. Only can simply not state that there is an absolute lack of data misinterpretation throughout the manuscript. It is essentially impossible to be certain about the absence of any potential misattributions. 5. Introduction, line 29 '' Numerous studies have shown that NIRS can be applied clinically in the resuscitation and cardiac surgery settings where cerebral desaturation events can be both effectively monitored and managed.10-13'': the authors imply that there are clear protocols for interpreting desaturations. This is to date not the case. Absolute and relative thresholds theoretically requiring prompt intervention have been proposed, however these have not been validated nor is there any consensus on which interventions are indicated and when. 6. Methods section, line 30: '' age greater than 18 years'' should be ''age above 18 years''. 7. Methods section, line 39: there is a comma missing after the word ''isolation''. Do the authors mean intraoperative single lung ventilation? 8. Methods section: was the anesthesiologist or any of the staff members blinded to the current rSO2 readings? 9. Methods section: did the authors encounter any problem in positioning the cerebral rSO2 sensors, since the anesthesia protocol mentions routine use of BIS? 10. Statistical analyses: why did the authors use the Kruskal Wallis test for non-parametric data? The Mann-Whitney U test would suffice since only two independent groups are being tested. 11. Results section, line 17-18: how did the authors calculate the treatment effect? The mean difference between the TMH and TN group does not seem to be 19% in the left cerebral hemisphere. 12. Results section: rather than only describing the results from the longitudinal time-by-treatment interaction, it would be nice to visualize the relationship between delta rSO2 and time. 13. The authors elaborate on the potential mechanism behind the effect of hypercapnia on cerebral rSO2. An important part of the mechanism which is not included in the discussion section is cerebral autoregulation. This is extremely important, since this autoregulatory mechanism is influenced by the pCO2 state as well as other factors such as hemodilution. This has been described by Ševerdija and Vranken: Ševerdija EE, Vranken NP, Simons AP, Gommer ED, Heijmans JH, Maessen JG, Weerwind PW. Hemodilution combined with hypercapnia impairs cerebral autoregulation during normothermic cardiopulmonary bypass. J Cardiothorac Vasc Anesth

	2015;29:1194-1199. Vranken NPA, Weerwind PW, Sutedja NA, Ševerdija EE, Barenbrug PJC, Maessen JG. Cerebral oximetry and autoregulation during cardiopulmonary bypass: a review. J Extra Corpor Technol 2017;49:182-191. 14. Moreover, the authors did not include the fact that higher rSO2 values might in fact be representative of a rather disadvantageous scenario. This has been studied by Henriksen who described the so-called ''brain luxury perfusion'' (Henriksen L. Brain luxury perfusion during cardiopulmonary bypass in humans. A study of the cerebral blood flow response to changes in CO2, O2, and blood pressure. Cereb Blood Flow Metab 1986;6:366-378.).

REVIEWER	Lynda Cochrane Clinical Statistics Consultants UK
REVIEW RETURNED	23-Jun-2019

GENERAL COMMENTS	The principal weakness of the analysis is the lack of inclusion of independent variables. The authors assumed that, because no statistically significant between-treatment group differences in baseline values were found, the two groups were comparable. This would have been true only if the study had been powered to detect these differences. Details are described below. The principal weakness of the analysis is the lack of inclusion of independent variables. The authors assumed that, because no statistically significant between-treatment group differences in baseline values were found, the two groups were comparable. This would have been true only if the study had been powered to detect these differences. **Statistical methods** 1) The number of measurements of the primary outcome (ΔrSO_2) made throughout the duration of the surgery varies between individuals, therefore Generalised Estimating Equations (GEE) are more appropriate than Repeated Measures (RM) ANOVA. 2) The effect of time on the dependent variable is clearly not linear. 3) X^2 methods are not appropriate for ordinal data (Protocol Page 23). 4) A key to understanding the figures in [], {} in Table 2 would be helpful. 5) ±SD is not a meaningful statistical expression. It is more appropriate to use either (SD), (SE) or a 95% confidence interval. 6) **Significance of between-group differences (Tables 1, 3, 4)** 7) As mentioned above, the lack of statistically significant between-group differences in baseline and post-operative characteristics (Tables 1 and 3, respectively) could be due to low power. The duration of surgery is of particular note: themedian in the TN group is only 66% of that in the TMH group (Table 1) and this should be addressed. 8) Similarly, the total dose of intra-operative opioid in TN patients

	is 77% of that in the TMH group (Table 3). 9) Although the comparison of base excess in the two groups was not significant at the 5% level, there is some evidence of a difference (Table 4, P=0.069). Over-reliance on the 5% figure, particularly when a study is not specifically powered to evaluate differences, can be misleading. 10) There is a higher proportion of COPD cases in the TMH group (Table 1, P=0.047). 11) Due to the doubt about the reliability of the between-group comparisons, relevant independent variables should also be included in analyses (most notably of ΔrSO_2 and post-operative delirium). Univariate relationships between the dependent and independent variables, and inter-relationships between the independent variables, should first be investigated before being included in a multivariable model. Full details of the model(s) implemented should be presented. 12) Are the statistically significant difference noted in Table 4 clinically significant? 13) Given the between-hemisphere differences, it would be more appropriate for the figures in Table 4 to be presented separately for right and left. **Inclusion / exclusion criteria** 14) Inclusion criteria were stated in the protocol (Page 5) as adults aged > 65 years and exclusion criteria (Page 6) as adults aged < 65 years. Are adults aged 65 included or excluded? 15) In contrast, the manuscript inclusion criteria were stated (Page 5, Line 30) as adults aged > 18 years with no age-related exclusion criteria.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Professor Akca's queries

We thank expert Professor Akca for his valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** It appears as the investigators accomplished the goal $ETCO_2$ level by managing mechanical ventilation (by decreasing minute volume). This approach would result in differences in intrathorocic pressures between the two study groups. Such difference might have influenced the venous return and might have resulted in cardiac output differences. Although such effect could be small, it needs to be discussed in the manuscript.

   Authors' response: Thank you for raising this very interesting point, which created excellent discussion and debate amongst our authors. We agree that in principle, with positive pressure ventilation, the intrathoracic pressure increases during inspiration causing a decrease in venous return, right ventricular output, and pulmonary blood flow. On expiration, the converse is true.[1] However the understanding and application of changes in intrathoracic pressure during mechanical ventilation and cardiac output is more complex. Although positive-pressure ventilation increases lung volume and increases peak airway pressure, the degree to which both intrathoracic pressure (being oesophageal, pleural or pericardial) and lung volume increase is a

function of airway resistance as well as lung and chest wall compliance.[2] Several studies has shown that that in patients in which the application of PEEP determined effective alveolar recruitment, mean pulmonary artery pressure decreases with peak airway pressure application while cardiac output is not severely affected. [3-5] True transmural ventricular filling pressures are not clinical available and research in this is lacking. Predicting the impact on the effects of increasing lung volume, peak airway pressures, lung compliance and pressures transmitted to the pericardial space is difficult at best.[6]

In our study, in both groups we kept the PEEP consistent at 5 cm PEEP. This important point has been included in our revised manuscript. We acknowledge that any decrease in systemic venous return and, thus, right ventricular inflow must result in decreased pulmonary venous return and inflow to the left ventricle because the two ventricles pump in series. In addition to this passive coupling of the right and left ventricle, PEEP may have more direct mechanical effects on LV due to ventricular interdependence.[7,8] We feel assured that because PEEP was kept constant for both groups, and that the changes in lung volumes were small, the impact on cardiac output, as the reviewer has stated, is likely to be very small. Furthermore, we have reviewed each of the medical records and the use of vasoactive drugs and inotropes is similar. Finally, as presented graphically under point 3 below, blood pressure was similar across both groups. Unfortunately, we did not measure cardiac output as this was a consideration for our study.

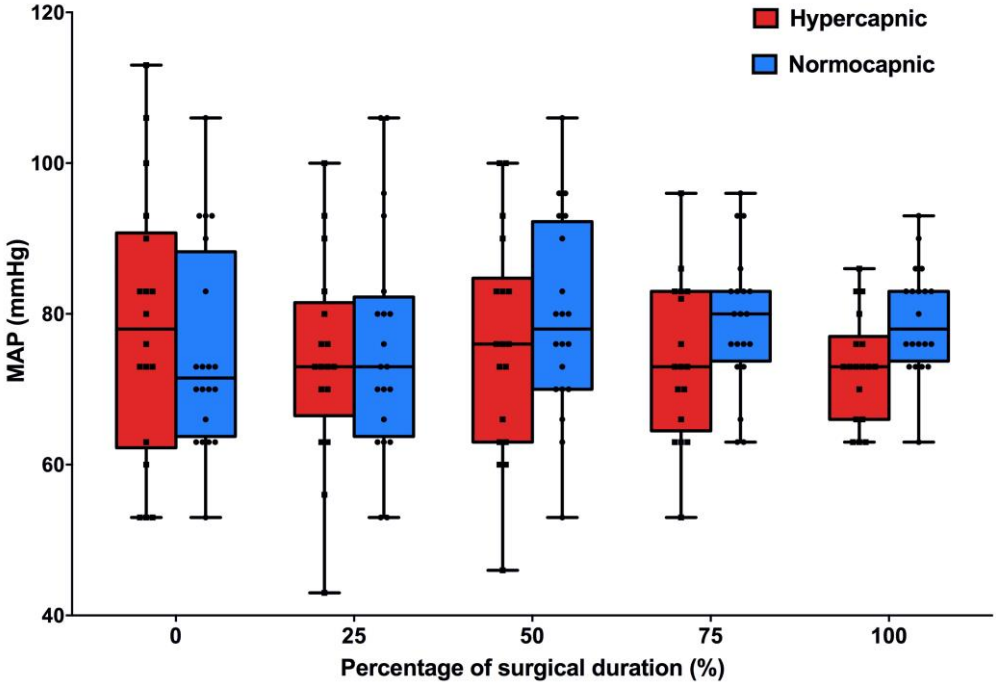
We have included this in the limitations. In the limitation section we state *"The effects on changes in intrathoracic pressures on cardiac output were not measured. Changes in intrathoracic pressure may have adversely impacted cardiac output, which may in turn have affected the $ETCO_2$. However, given that the PEEP was held constant in both groups, and the changes in lung tidal volumes were relatively small, the impact of intrathoracic pressure on cardiac output is likely to be small."*

2. **Expert reviewer:** Did the investigators measure CO/CI and SV in the study? Or in any subgroup of patients? Were any intravascular volume monitors utilized in the study?

Authors' response: That you for this excellent question, which we have also acknowledged as a limitation in the discussion section. We strongly agree that the measurement of cardiac output, stroke volume, and systemic vascular resistance (in patients where central venous pressure was measured) would have been important additional haemodynamic metrics to collect. These metrics may have provided more robust and insightful information about the determinants of oxygen delivery and its subsequent impact on rSO_2. We kept the study design as pragmatic as possible, and therefore use of advanced monitoring was at the discretion of the attending anesthetists and was not applied to any of the patients. All patients, however, had continuous blood pressure measured as part of the inclusion criteria, and blood pressure was defended consistently between both groups to within 20% of the patients' preoperative baseline value. We also accurately recorded mean arterial pressure at 5 specific time points, which we have presented below; we found no significant difference in the mean arterial pressure in both groups of patients. Comparing blood pressure between the groups was not directly relevant to the study outcomes, and as a result we have not included the Figure under point 3 below in the revised manuscript but would be happy to do if requested.

3. **Expert reviewer:** Intraoperative mean blood pressure results were not provided in the manuscript (only a statistically not significant result was provided). Please report the BP data. BP is both a perfusion marker and also can be considered as a marker of CO (i.e. BP=CO x SVR).

Authors' response: Thank you for raising this excellent important point. We have recorded intraoperative non-invasive blood pressure data as the mean blood pressure (MAP). As summarized in the Figure below, we have reviewed the MAP in each anaesthesia record in detail and observed no clinical or statistical difference between the groups (*P*=0.307)



4. **Expert reviewer:** Did the investigators include any neuroaxial anesthesia/analgesia patients? If so, please report the nature of neuroaxial anesthesia/analgesia, and the type and amount of medications given through epidural/spinal routes. Neuraxial anesthetic and analgesics may both alter intravascular volume status and hemodynamic responses. Additionally, these methods would provide superior analgesic approach, and may impact the tissue oxygenation levels due to better pain management.

Authors' response: Thank you for this important question. We have already outlined the standardisation of anaesthesia, which we report in the Methods section. We have briefly updated this to state *"Conduct of anaesthesia, including the use of additional invasive monitoring, intra-operative medications, fluids intervention, and use of vasoactive medications, regional anaesthesia and use of intraoperative opioids were entirely at the discretion of the attending anaesthetist."*

Five patients (25%) in the mild hypercapnia (TMH) group and two patients (10%) in targeted normocapnia (TN) group received single shot spinal analgesia with intrathecal morphine. No patients receive any epidural analgesia. These results have been updated in Table 3. We are not aware of any studies to date that report the effects of intrathecal morphine on tissue oxygenation levels due to better pain management. Importantly, no patients received epidural analgesia or anaesthesia.

Intrathecal morphine		
Number of patients	5	2

Mean dose (mcg)	220	350
Epidural analgesia		
Number of patients	0	0

5. **Expert reviewer:** Although, they were not statistically significant, there was about 10% difference in PaO_2 levels and ~30% difference in narcotic analgesic doses between the groups. Both of these differences were in the favor of the hypercapnia group. Because both arterial oxygen and pain are established factors to influence tissue oxygenation, there needs to be discussion statements to address these differences between the groups.

   Authors' response: Thank you for this excellent comment, which also has resulted in robust discussion amongst our authors. Thank you for pointing out this interesting association. As suggested, we have discussed this further in our discussion section.

   We have now stated *"Whilst our study was not designed to measure differences in analgesia and partial pressure of oxygen in arterial blood, we observed a 10% higher median PaO_2 level in the TMH group and found that the median intraoperative analgesia requirements were also approximately 30% higher. Both arterial oxygen levels and pain have been reported to influence tissue oxygenation [9], which was not directly measured in our study. The effect of pain on cerebral oxygenation is unclear, and has not be borne out in clinical studies [10]; further studies exploring this association are needed.*

6. **Expert reviewer:** How were the $ETCO_2$ data reported? Did the authors first take the mean $ETCO_2$ data per patient and then reported medians for the groups? Please describe in detail.

   Authors' response: Thank you for this excellent question. We agree that this is an extremely important metric to report. $EtCO_2$ data was collected was collected throughout the study. The mean $EtCO_2$ was taken and the medians were reported for the group. The corresponding result is now presented for completeness in Table 4. In keeping with our intervention, $EtCO_2$ was higher in the TMH group compared to the normocapnic group. Thank you once again for this comment.

Table 4. Average arterial blood gas values and corresponding end-tidal carbon dioxide.

	TMH group[b] (*n*=20)	**TN group[b] (*n*=20)**	*P*-value
pH	7.31 [7.27 to 7.33]	7.46 [7.43 to 7.47]	<0.001
PaO_2 (mmHg)	156.8 [146.3 to 217.2]	142.5 [122.5 to 199.1]	0.380
PaCO_2 (mmHg)	51.50 [46.88 to 60.88]	34.75 [32.75 to 38.12]	<0.001
$EtCO_2$ (mmHg)	46.40 [39.80 to 50.20]	30.40 [28.50 to 32.00]	<0.001

Bicarbonate (mEq L^{-1})	25.00 [24.00 to 27.75]	24.00 [22.00 to 24.62]	0.020
Base excess (mmol L^{-1})	-1.00 [-2.50 to 0.25]	1.00 [-0.88 to 2.00]	0.069
Potassium (mEq L^{-1})	3.98 [3.73 to 4.38]	4.03 [3.58 to 4.31]	0.759
Total Hb (g L^{-1})	130.50 [118.12 to 140.62]	122.25 [106.88 to 131.25]	0.132

7. **Expert reviewer:** Blindness of anesthesiologists can be accomplished by blinding them only to the real time monitoring of cerebral oximeter data. In that case, the improved cerebral oxygen saturation's possible/potential effects on postoperative delirium would not bias anesthesiologists. If there will be a follow-up study focusing on postop delirium outcome, I strongly recommend the authors to keep the anesthesia team blinded to the cerebral oximeter data.

   Authors' response: Thank you for this excellent suggestion that will definitely consider in similar studies that we are currently designing. We also agree that blinding the anaesthetist to the outcomes being collected will also increase the internal validity of the study.

   For the present study, data collection for all the trial outcomes, including the assessment of delirium was collected by an independent researcher blinded to treatment allocation. Analyses were also performed by a statistician who was also blinded to all intraoperative events and to treatment allocation. This has already been outlined in our manuscript. The anaesthetist delivering the intervention did not participate in the assessment of postoperative delirium. This has also been included in the revised manuscript.

   In the manuscript we have stated *"Data collection for all the trial outcomes was collected by an independent researcher blinded to treatment allocation. The sequence was decoded after the data was analysed. The anaesthetist delivering the intervention did not participate in the assessment of postoperative delirium".*

   "A single trained interviewer, blinded to randomisation, and proficient and trained in the Confusion Assessment Method, conducted all the assessments pre-operatively when patient arrived at the hospital and at 8am on the next day after surgery in the ward (within 18-24 hours postoperatively). The baseline cognitive function was not formally assessed with collateral history from family or carers."

8. **Expert reviewer:** This reviewer suggests that those two "hypercapnia group" patients whom rSO_2 data could not be collected, the authors may consider reporting the other outcomes' data (i.e. delirium, LOS, etc.).

   Authors' response: Thank you for the excellent suggestion. As the expert reviewer correctly pointed out, the other outcomes' data can be presented for the two patients in the TMH group. In fact, the other outcomes' data were included in the analysis of the secondary endpoints and were presented in Table 1 and Table 3. This is also stated clearly in the results section *'two participants in the hypercapnic group had failure of bilateral probe attachment and lead connection problem that were unable to be rescued. These patients were subsequently excluded from the analyses of oxygenation as no rSO_2 data were captured and were included in the analysis of all other variables and endpoints'*

9. **Expert reviewer:** If one considers CO_2 as a "drug" (i.e. medicine), the diminishing tissue oxygenation effect by time may suggest some type of tolerance. Although such tolerance can be due to metabolic contributions, this is a very interesting result and may mean further diminished effect size of hypercapnia in longer procedures.

Authors' response: Thank you for this very thought-provoking question which is beyond the scope of our study objectives and would be extremely difficult to quantify. We have reviewed all our raw data again to see if there are any signals to suggest tolerance. We have performed a time by treatment analysis. On the longitudinal time-by-treatment interaction analysis, the difference in $\%\Delta rSO_2$ on both left and right between the two groups diverged with time with the intervention group exhibiting smaller percentage decrease over time compared to the control group (time-by-treatment interaction $P<0.001$ for both left and right hemispheres).

We obtained very similar results on robustness analyses when the above model was adjusted for age, baseline oximetry and preoperative haemoglobin levels, as well as when percentage of total duration of surgery instead of minutes from the start of surgery were included. In the TMH group, nearly all participants recorded cerebral oxygen saturation levels above the baseline valve throughout the intra-operative period, with the effect most pronounced between 90 to 240 minutes into surgery. However, we have found no conclusive evidence to suggest tolerance over time. We also think that any renal compensatory mechanisms to compensate for acid-base homeostasis would not have become operational within the study time-frame.

On review of the literature, there is very little data available on this subject. Glatte et al. reviewed the effects of CO_2 tolerance on human subjects and found that up to approximately 6 mmHg inspired partial pressure of carbon dioxide ($PICO_2$), there was no physiologic differences. Studies performed at 7 to 8 mmHg carbon dioxide have shown only minor respiratory changes including slight increases (up to 24%) in respiratory minute volume and minimal alveolar PCO_2 differences. No significant pH changes are recorded. After 10 to 15 days, respiratory parameters all returned to control values except the slightly elevated $PaCO_2$ levels.

Short and long term studies carried out at carbon dioxide partial pressures of 8 to 13 mmHg have revealed increasing but still minimal physiologic shifts. Subjects are asymptomatic with small increases in respiratory minute volume, tidal volume, and $PaCO_2$. Questionable changes in central nervous system "excitability" have been noted and are probably of little importance. At 14 to 15 mmHg ambient PCO_2, more prominent changes are seen in both respiratory and metabolic parameters. Pulmonary changes include an increase in the alveolar partial pressures of carbon dioxide of 2 to 3 mmHg with an increase in tidal volume and respiratory rate. It was also found that initial decreases in serum pH would be compensated during the experiment. At these levels of $PICO_2$, some subjects become aware of increased rate and depth of breathing while others do not.[11]

Brakett et al. studied seven volunteers who tolerated inspired partial pressure of carbon dioxide of 53 mmHg for 40 to 90 minutes and inspired partial pressure of carbon dioxide of 76 mmHg for 15 to 25 minutes. Acid-base studies were performed to describe quantitatively how pH is defended with acute increases in partial pressure of carbon dioxide. $PaCO_2$ showed an increase from a control value of 39 to 53 mm. Hg while breathing the first gas mixture and an increase to 78 mm. Hg during the second phase of the experiment. The pH of the serum went from a control of 7.42 down to 7.31 and finally 7.17. Analysis of data revealed modest buffering by endogenous

bicarbonate stores with a linear relationship of hydrogen ions concentration and $PaCO_2$. It was found that for every millimeter of mercury increase in $PaCO_2$, hydrogen ion concentration increased by 0.77 nM.[12]Schwartz et al. performed similar studies on dogs have shown that this hydrogen ion concentration to $PaCO_2$ relationship increased only by 0.33 nM of hydrogen ion concentration for every increment of 1 mm of $PaCO_2$. This is thought to be secondary to renal mechanisms and acid excretion which become operational after several days of hypercapnia.[13]

10. **Expert reviewer:** How many times the arterial blood gas levels were measured throughout the study? If the answer is just one or two, then these results don't deserve a dedicated table, instead they can be included in table 1 or 2.

    Authors' response: Thank you for the suggestion. All patients had a blood gas performed at the start of surgery and at completion. Other blood gases were performed at the discretion of the attending Anaesthetists. The maximum number of gasses performed in a patient was 3. Only 3 patients in the TMH group and 3 patients in the TN group has 3 gases performed and for brevity have not reported these in the manuscript. The median (IQR) number of blood gases in the TNH group was 2 (2:2) which was identical to the TN group.

    In the manuscript we have outlined the blood gas sampling. We have stated "*Immediately after tracheal minute ventilation was adjusted to achieve an $EtCO_2$ concentration of 45-55 mmHg in the TMH group or 35-40 mmHg in the TN group. Due to presence of alveolar dead space, $EtCO_2$ can be lower than $PaCO_2$ by up to 5 mmHg. Therefore, an arterial blood gas was obtained to check $PaCO_2$ and ventilation was further adjusted accordingly to achieve the desired $PaCO_2$ target ranges. The $PaCO_2$-$EtCO_2$ gradient was then maintained throughout the surgery, with the assumption that the $PaCO_2$ would remain constant. Additional ABG were sampled at the discretion of the anaesthetist if the gradient required re-evaluation e.g. requirements for adjustment of ventilation setting. Finally, at completion of surgery, an ABG was sampled to accurately document the $PaCO_2$ value, and to assess whether $PaCO_2$ was being maintained within target values.*"

    Table 1 is a summary table of the baseline patient characteristics whilst Table 4 is a summary table of the arterial blood gas results and they represent some of the secondary outcomes of the study. They were separated to two tables for clarity of presentation.

11. **Expert reviewer:** As far as this reviewer is aware, the effects of mild hypercapnia on cerebral oxygen saturation was reported twice before (by our group): 1) In a randomized controlled trial, [Ref: Effect of intra-operative end-tidal carbon dioxide partial pressure on tissue oxygenation. Anaesthesia. 2003 Jun;58(6):536-42.] the authors tested the effects of mild hypercapnia on tissue oxygenation in abdominal surgery patients. ; 2) In a crossover design a few years ago, [Ref: Manuscript reference #41 -] investigators used on-pump cardiac surgery setting to omit the contribution of inotropic effects (i.e. CO increase) of hypercapnia on tissue oxygenation. The authors may consider rephrasing their Introduction section to state the relationship between mild hypercapnia and rSO_2 in patients undergoing surgery without pre-existing cerebral desaturation events was in fact studied before.

    Authors' response: Thank you for this excellent comment. We have now made reference to the above two papers in our introduction. Similar to our study, the randomized controlled trial conducted by Professor Akca's research group delivered mild hypercapnia intra-operatively and cerebral oxygen saturation was found to be higher than that of the normocapnia group.[14] In contrast to our study, the main focus on the trial was to investigate tissue oxygenation and its

relationship with wound infection risk after surgery. Study participants were much younger and the majority of operations involved were abdominal surgeries.

The cross-over study conducted by Professor Akca's research group proved that dilatation of peripheral vessels has little effect on increased tissue oxygenation in the setting of mild hypercapnia in patients undergoing cardiopulmonary bypass. Bypass pump flow was kept constant throughout the measurement periods to negate any potential contributing effects of cerebral oxygenation from changes in cardiac output.[15] Importantly our study complements and builds on the above research as we explore the temporal relationship between mild hypercapnia and cerebral oxygen saturation, which was not the primary focus of the authors' study. Therefore, further to a 30 minute equilibrium period that the authors' study had used, we extended the monitoring period to the full duration of surgery.

We have modified the introduction section. We now state *"rSO_2 was reported to be higher in mild hypercapnia in patients undergoing surgery but the intra-operatively temporal relationship between rSO_2 and mild hypercapnia remains unclear.[14]"*

12. **Expert reviewer:** Please include duration of surgery data in the Abstract.

Authors' response: Thank you for this suggestion. Due to the limited word count, and the fact that duration of surgery is not a primary or secondary outcome, we respectfully have not reported this in the abstract. However, given that the design of our study included patients undergoing surgery of greater 2-hour duration, we believe that this is an important statement to include in the abstract. Accordingly, in the Abstract section under 'Design' we now state *"A prospective, randomised controlled trial in adult participants undergoing major surgery greater than 2 hours expected duration requiring at least one overnight admission."*

Reviewer 2: Dr Fowler and Professor Pearse's queries

We thank expert Dr Fowler and Professor Pearse for their valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** Introduction: This is concise and clearly states the rationale for the trial. It may be worth expanding on the recent focus on long term cognitive impact of anaesthesia and steps to minimise this.

Authors' response: Thank you for this excellent suggestion. Whilst the key focus of our study was to assess the changes in regional cerebral oxygenation in the setting of mild hypercapnia, we recognize the clinical importance of delirium and the significant impact of delirium on patients' recovery after major operations.

Postoperative cognitive changes have been reported in patients over 65 years for over a century, and anaesthesia has often been mentioned as a possible cause for these changes. The cognitive impact of anaesthesia is undoubtedly a growing area of research interest in recent years.

However, the topic of cognitive dysfunction is complex and our study was not designed to investigate for cognitive dysfunction in the setting of mild hypercapnia nor varying levels of regional cerebral oxygen saturation. As such, we have kept the focus of the introduction section on our primary outcome of the study, which is the use of NIRS cerebral oximetry to measure changes in rSO_2 in the setting of intra-operative mild hypercapnia. We have also expanded on this important Question in our response to Reviewer 4 below.

2. **Expert reviewer:** Methods: The inclusion criteria differ between the ANZICTR entry and the manuscript [Age >18 years vs Age >50 years on ANZICTR]. Please clarify which of these were used as inclusion criteria, presumably >18 years as the range for age in table 1 includes patients aged 31 years. It is important that this discrepancy is explained in the manuscript (this is an important part of the CONSORT guideline).

Authors' response: Thank you for this important and insightful question. We agree with the reviewer that the information we have presented on the Trials Registration website is confusing and we have now clarified this discrepancy by amending this section in the Trial Registry. The inclusion criterion is adult patients with age over 18 years.

This is stated in the manuscript under the methods section *"Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care."*

3. **Expert reviewer:** Methods: How was a 15% change between control and intervention groups at the end of surgery judged to be clinically important?

Authors' response: Thank you for this very important question. There is unfortunately no universally accepted threshold to identify pathological cerebral saturation. The threshold for identifying cerebral ischemia may be influenced by a number of patient-specific or technology-dependent variables.

Numerous studies have utilised threshold of approximately 20% decrease in rSO_2 as the threshold for abnormal rSO_2. Murphy et al. defined cerebral desaturation events as a greater than or equal to 20% decrease in rSO_2 values from baseline measures or an rSO_2 of less than or equal to 55%.[16] Denault et al. proposed a clinical algorithm with the use of NIRS intra-operatively to monitor for cerebral saturation level and defined abnormal rSO_2 as a 20% reduction from baseline or an absolute decrease below 50%.[17] Levy et al. compared changes in oxygen saturation with electroencephalographic evidence of cerebral ischaemia and estimated an ischaemic threshold of 47% in cerebral saturation.[18]

If we consider 65% as the baseline cerebral oxygen saturation, a 15% increase from baseline would approximate 75% and a 15% decrease from baseline would approximate 55%. If rSO_2 of

47% is the threshold for cerebral ischaemia, rSO_2 of 55% can be considered as a conservatively approach to define cerebral desaturation.

4. **Expert reviewer:** Methods: These are otherwise very clear, and would enable study replication. The study has been reported broadly in line with CONSORT and SAMPL guidelines. Important, trial specific harms were considered (hyperkalemia particularly).

   Authors' response: Thank you for the positive comment.

5. **Expert reviewer:** Results: The operations performed are not entirely clear, for example orthopaedic (commoner in TMH group) vs. HPB operations (commoner in TN group) may not be directly comparable in terms of risk for postoperative delirium and may have very different risk profiles.

   Authors' response: Thank you for this important comment. We recognize that our study participants had undergone a heterogeneous mix of different types of surgery. For example, hepatobiliary-pancreatic operations were commoner in TN group yet thoracic operations were commoner in the TMH group. Therefore, their risk profiles for developing postoperative delirium were difficult to quantify. Risk of postoperative delirium is dependent on multiple factors and whilst type of surgery can affect the risk profile, other factors such as intra-operative opioid, pre-medication, baseline cognitive function and other medical co-morbidities also play an important role in determining the risk. In our study, we have compared and presented these variables in Table 1 and Table 3. They are mostly similar between the groups. To address the limitation of the heterogeneity of type of surgery, we have stated in the limitation section *"our findings of a greater incidence of early postoperative delirium in the TN group need to be interpreted with caution as confounders of postoperative delirium were not controlled, and our study was not powered to investigation postoperative delirium,"*

6. **Expert reviewer:** Results: Two neurosurgical patients were included; was this for spinal surgery? Hypercapnia could be considered harmful by some during certain types of intra-cranial surgery. If patients undergoing intra-cranial surgery were included, it would be important to discuss this.

   Authors' response: Thank you for raising this important question. We made a comment in the footnote of the table stating neurosurgery only includes non-intracranial procedures, e.g. complex spinal surgery. However, we do acknowledge the term neurosurgery can cause confusion and as a result, we have modified the term neurosurgery to spinal surgery in Table 1

Surgical Characteristics		
Duration of surgery (mins)	219.0 [123.8 to 303.8]	144.0 [107.8 to 218.2]

Left baseline oximetry (%)	68.7 [63.9 to 72.2]	63.4 [57.3 to 69.6]
Right baseline oximetry (%)	67.9 [64.6 to 70.3]	64.0 [59.4 to 69.0]
O_2 finger saturations (%)[f]	98.5 [98.1 to 99.0]	98.5 [97.9 to 99.0]
LOS (days)[g]	5 [2.0 to 12.0]	5 [1.8 to 11.5]
Type of surgery		
colorectal	2 (11.1)	1 (5.0)
endocrine	2 (11.1)	2 (10.0)
ear nose & throat	0 (0.0)	1 (5.0)
hepatobiliary	6 (33.3)	9 (45.0)
spinal surgery[h]	1 (5.6)	1 (5.0)
orthopedic	2 (11.1)	1 (5.0)
thoracic[i]	4 (22.2)	1 (5.0)
urology	1 (5.6)	3 (15.0)
vascular	0 (0.0)	1 (5.0)

7. **Expert reviewer:** Results: What were the characteristics of patients who suffered postoperative delirium? It is difficult to unpick the association between TMH/TN and subsequent delirium with the data presented as they are currently, especially as the TN group seems to have higher risk patients (More ASA 3+, more HPB surgery).

Authors' response: Thank you for this important question. We acknowledge that heterogeneity of type of surgery is one of the limitations of our study. As a result, we have stated in the limitation section *"our findings of a greater incidence of early postoperative delirium in the TN group need to be interpreted with caution as confounders of postoperative delirium were not controlled, and our study was not powered to investigation postoperative delirium,"*

We have examined the characteristics of patients who suffered postoperative delirium. All 5 patients were in the TN group. The median [IQR] age was 72 [59.5 to 77]. Operations included right hepatectomy, Whipple's procedure, hemithyroidectomy and live donor renal transplantation.

The median [IQR] duration of surgery was 171 [83.5 to 254.5]. ASA were 3,2,1,4 and 4. 1 out of 5 suffers from diabetes, 0 out of 5 suffers from COPD.

We have included the above findings in the discussion section *"Patients who suffered from postoperative delirium were all in the TN group but they were also older (median [IQR] age 72 [59.5 to 77]) and had higher ASA scores (ASA scores of 3, 2, 1, 4 and 4). Their baseline medical co-morbidities and duration of surgery (median [IQR] duration of surgery 171 minutes [83.5 to 254.5]) were similar to other study participants."*

8. **Expert reviewer:** Discussion: Important to highlight that a single time point CAM positive is not necessarily diagnostic of postoperative delirium. This is particularly important when considering that the length of stay between the two groups was the same, a surprising finding if 25% of patients in the TN group had postoperative delirium which is well known to extend length of stay. Use of a standard definition of postoperative delirium would provide a more reliable and interpretable outcome for any future work, as would baseline cognitive assessment to capture un-diagnosed cognitive impairment.

   Authors' response: Thank you for this excellent comment. Delirium is often under-recognized and under-diagnosed potentially due to its fluctuating nature. Although there are various rapid bedside assessment tools used to diagnose delirium in hospitalized patients, the confusion assessment method is the best known and most frequently used by clinicians. (Inouye) CAM has been used in over 300 original articles to date, as either a process or outcome measure, and translated into over six languages world-wide. When validated against the reference standard ratings of geriatric psychiatrists based on comprehensive psychiatric assessment, CAM had a sensitivity of 94-100%, specificity of 90-95%, and high inter-observer reliability in the original study of 50 patients.[19] More recently, this work has been extended,[20] and in 7 high-quality validation studies on over 1,000 subjects, the CAM had a sensitivity of 94% (95% CI 91-97%) and specificity of 89% (95% CI 85-94%).

   We agree with the reviewer that baseline cognitive assessment would be helpful in capturing baseline cognitive function. As cognitive impairment is a known risk factor for delirium,[21] it would be useful to include the information when interpreting CAM scores. In our study, we performed CAM on our study participants immediate on arrival to hospital pre-operatively but we did not perform any formal baseline cognitive assessment. We acknowledge that this is part of the limitation of our study and therefore, we have stated under the limitation section *"mental state was only assessed by CAM, once pre-operatively and once postoperatively. Accordingly, our findings for delirium should be viewed as hypothesis generating"*

9. **Expert reviewer:** Discussion: Important to note that there was a change from baseline within both groups, and that this was a decrease in $rS0_2$ in the TN group, and an increase in rSO_2 in the TMH group. This is especially important as the sample size was based on an assumption of no change in the control group, when there was an observed negative change.

   Authors' response: Thank you for this comment. We agree with the reviewer and have now stated in the manuscript that there was a change from baseline within both groups, and that this was a decrease in $rS0_2$ in the TN group, and an increase in rSO_2 in the TMH group. In the discussion section we have further elaborated on this and we now state *"We found that the*

sustained difference in rSO_2 over time was a combined effect of stable increase in rSO_2 from baseline in the TMH group and a stable decrease in rSO_2 from baseline in the TN group. In the literature, the association between normocapnia and reduced CBF and lower levels of rSO_2 were reported briefly. The exact mechanism and associations between normocapnia and lower rSO_2 are not entirely clear. Whilst theoretical absolute and relative saturation thresholds requiring prompt interventions have been proposed, these thresholds have not been validated and there is a lack of consensus on the indication and timing of interventions. In our study, reduction in rSO2 from baseline was small in the majority of patients in the TN group and the attending anaesthetists had no rSO_2 target to titrate to."

This finding is important as the sample size was based on an assumption of no change in the control group, when there was an observed negative change. We have therefore included the following statement in the limitation section *"Sample size calculation was based on the assumption that there were no changes in rSO_2 values from baseline in the TN group. The observed negative change can therefore impact the calculation."*

10. **Expert reviewer:** Results/Discussion: Looking at table 2; at 360mins after start of surgery, there was an average of 43.4% decrease in oxygen saturations in the TN group (sample size =2), while TMH stayed fairly static. This is an important finding; patients with maintained normocapnia have decreasing rSO_2 over time, while targeted mild hypercapnia is associated with a stable rSO_2 compared to baseline. Does this not meant that the interpretation should be subtly altered to state 'Mild hypercapnia is associated with a stable increase in rSO_2 from baseline, while normocapnia is associated with a decrease in rSO_2?'. To say that there is a larger increase over time in the hypercapnia group is unclear, as the change is actually the same but is statistically exaggerated by the sharp decrease in rSO_2 in the normocapnia group.

Authors' response: Thank you for the very thoughtful comment. We agree with the reviewer that this is an extremely important finding. We have therefore modified the manuscript under the Introduction, Discussion and Conclusion sections to now state *"mild hypercapnia is associated with a stable increase in rSO_2 from baseline, while normocapnia is associated with a decrease in rSO_2"*

11. **Expert reviewer:** Discussion: A recent publication in this area exploring the use of NIRS during cardiac surgery may be an important discussion point. (Rogers et al. Randomized trial of near-infrared spectroscopy for personalized optimization of cerebral tissue oxygenation during cardiac surgery. Br J Anaesth. 2017;119(3):384-393.) [I have no association with this group].

Authors' response: Thank you for the suggestion. Rogers et al. performed a randomized controlled trial to investigate whether there are any cognitive function benefits in patients undergoing cardiopulmonary bypass when a NIRS-based optimization of cerebral oxygenation was implemented. [22] We have now incorporated this important study in the discussion section *"There has been conflicting evidence in the literature regarding the relationship between rSO2 and LOS or postoperative cognitive performance. Cognitive outcomes were similar in groups with or without NIRS-based rSO_2 optimisation in a recent randomised controlled trial."*

12. **Expert reviewer:** Consider consistent use of post-operative or postoperative; the tables use post-op while in the main text, there is no hyphenation.

    Authors' response: Thank you for this suggestion. The correction has been made in the revised manuscript. We have now used the term postoperative consistently throughout the manuscript. We have also expanded the abbreviation from post-op to postoperative in Table 3.

13. **Expert reviewer:** Methods: In exclusion criteria there should be a comma between 'requiring one lung isolation' and 'liver transplantation'

    Authors' response: Please accept our apologies for the oversight. The correction has been made in the revised manuscript under the methods section.

    "Exclusion criteria included patients undergoing cardiac surgery, procedures requiring one lung isolation, liver transplantation, intracranial surgery, GCS less than 15, known cognitive impairment, intellectual disability or a mental illness, moderate pulmonary hypertension (mean pulmonary arterial pressure greater than 40 mmHg), and American Society of Anesthesiology status V."

Reviewer 3: A/Professor Blanco's queries

We thank expert Associate Professor Blanco for his valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** For CONSORT Item 6a ("Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed"), to be consistent with the way you presented the study results for the primary outcome, please indicate in "Outcomes and Data Collection" section that the rSO_2 was measured in the two hemispheres. I would also suggest that you include a subsection called Arterial blood gases (as you did in the protocol) where you indicate what gases were measured, when, which of these were considered secondary outcomes and which ones not. In this case, explain what the role of these variables was in the analysis – for example, you mention that Hb is a confounding variable in the discussion but you do not explain anything else in the Methods section.

   Authors' response: Thank you for the suggestion. This has now been amended in the manuscript. In the revised manuscript we have maintained consistency in the way we present the study results and indicated in "Outcomes and Data Collection" section that the rSO_2 was measured in the two hemispheres. We now state *"rSO2 was measured in the two hemispheres separately. Following manufacturer instructions, two NIRS sensors were attached to patient's left and right forehead, recording both absolute and trend data bilaterally."*

We have also included a subsection title “Arterial blood gases” where we articulate what gases were measured, and when, and which of these are secondary outcomes. We have also briefly explained the role of these variables.

“Arterial blood gases
All arterial blood gas variables were collected by ABL80 FLEX Blood Gas Analyzer (Radiometer, Copenhagen, Denmark) with a fully automated micromode eliminating risk of user-induced bias or loss of accuracy with very small samples, and an interference-protected lactate analyses. ABG variables include partial pressure of oxygen, partial pressure of carbon dioxide, pH, bicarbonate concentration, base excess, lactate, haemoglobin and electrolyte concentrations such as sodium and potassium ion concentration. The machine calculates the bicarbonate concentration using the Henderson-Hasselbalch equation and the standard base excess (SBE) using the Van Slyke equation with the following reference points pH = 7.40, $PaCO_2$ = 40mmHg, and temperature = 37°C to determine changes in bicarbonate, protein anion, and phosphate concentrations, and therefore SBE. Two or more ABG samples were measured intra-operatively as described previously. The mean values of pH, bicarbonate concentration, base excess, and serum potassium concentration from the first and the last ABG sample were considered as some of the secondary outcomes for the study. Intra-operative pH, bicarbonate, and base excess are important variables that inform acid-base status of a patient, in particular, bicarbonate and base excess are useful when determining the extent of metabolic contributions or compensation. Potassium concentration is a key physiological parameter that affects cardiac action potential conduction, and its relevance in the study is paramount as hyperkalaemia from hypercapnic-induce acidosis is a potential complication of the intervention. Potential confounders to rSO_2 measurements such as Haemoglobin concentration and partial pressure of oxygen were recorded. Other variables such as lactate and sodium concentration were collected for routine clinical care and they were not considered as part of the outcome measures”

2. **Expert reviewer:** For CONSORT Item 17a (“For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)”), please do not only report in the “Results / Secondary outcomes” section the results for each group, but also the estimated effect size and its precision (using 95% confidence intervals). Moreover, please do not include here the results for total haemoglobin if this variable is not listed as one of the secondary outcomes in the Methods section.

Authors’ response: Thank you for this important comment.

In regard to reporting the primary outcome, we have reported the estimated effect size and precision along with the results. Under the results section, we have stated

“In the left hemisphere, the median [IQR] baseline oximetry was 68.7% [63.9 to 72.2] in the TMH group vs. 63.4% [57.3 to 69.6] in the TN group (P=0.233). On the right hemisphere the median [IQR] baseline oximetry was 67.9% [64.6 to 70.3] in the TMH group vs. 64.0% [59.4 to 69.9] TN group (P=0.286). On both sides, the %ΔrSO_2 was greater in the TMH group than the TN group throughout the duration of surgery (Figure 2). The average (standard deviation, SD) percentage changes in rSO_2 from the baseline to the conclusion of the surgery in TMH group were +8.56% (18.90%) on the left and +13.86% (18.17%) on the right, and in TN group they were -6.18% (17.24%) on the left and -5.48% (18.94%) on the right. The resulting treatment effects were 19%

(95% CI [9.2 to 28.8]; P<0.001) on the left and 19% (95% CI [10.9 to 27.0]; P<0.001) on the right (Table 2)."

In regard to reporting the results for total haemoglobin, we sincerely apologise for the misplacement of the results. They should be reported alongside with other baseline characteristics instead of other secondary outcomes. We have now modified the results section to rectify this error. We have stated under the results section *"Both groups also had similar mean arterial pressure intra-operatively (repeated measure ANOVA P=0.128), similar total haemoglobin (130.50 vs. 122.25 g L-1; P=0.132), and similar total dose of intra-operative opioid received, 21.67 mg in the TMH group [13.75 to 32.50] and 16.67 mg in the TN group [10.00 to 22.50] (P=0.22)."*

We have removed the results of total haemoglobin under the section secondary outcomes. *"No statistically significant differences in base excess (-1.00 vs. 1.00 mmol L-1; P=0.069) and potassium (3.98 vs. 4.03 mEq L-1; P=0.759) were observed intra-operatively"*

Reviewer 4: Dr Vranken's queries

We thank expert Dr Vranken for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** Why did the authors solely focus on postoperative delirium? Other postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation.

   Authors' response: Thank you for raising this important question, which has been discussed in detail among the authors. This question has also resulted in considerable discussion and debate. As discussed in detail in our manuscript, we have clearly stated that delirium is a secondary outcome and our findings of a greater incidence of early postoperative delirium in the targeted normocapnia group need to be interpreted with caution as confounders of postoperative delirium were not controlled, our study was not powered to investigation postoperative delirium, and mental state was only assessed by CAM, once pre-operatively and once postoperatively. Accordingly, our findings for delirium should be viewed as hypothesis generating.

   We chose delirium as a secondary end point as it is a well-defined complication of surgery, and clearly outlined in the Diagnostic and Statistical Manual of Mental Disorders fourth edition (DSM-IV–TR; www.dsmivtr.org/). There are key characteristics of delirium which include classic symptoms that patient may express in hypoactive, hyperactive, or mixed psychomotor behaviours and importantly tests have been developed and validated for use in diagnosis and grading of delirium, and as outlined in our manuscript where we used the Confusion Assessment Method (CAM), a validated tool to measure delirium.

   We agree with the Reviewer that postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation and certainly may be temporally associated with surgery. Postoperative cognitive dysfunction (POCD) is however difficult to

define. While the diagnosis of delirium requires a detection of symptoms, the diagnosis of POCD requires preoperative neuropsychological testing (baseline) and a determination that defines how much of a decline is called cognitive dysfunction. The spectrum of abilities referred to as cognition is diverse, including learning and memory, verbal abilities, perception, attention, executive functions, and abstract thinking. It is possible to have a decrement in one area without a deficit in another. Self-reporting of cognitive symptoms has been shown to correlate poorly with objective testing, so valid pre- and postoperative testing is essential to the diagnosis of POCD. [23] Unfortunately, we did not have the resources to accurately undertake the accurate assessment of POCD in our study.

Further, there has not been a standard methodology used in the multiple studies within the POCD literature. [24] Selection of neuropsychological test instruments, the amount of change considered to be significant, timing of testing, and inclusion and exclusion criteria have all varied. [25] Therefore, for the present study, whilst we acknowledge that postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation, given the logistical barriers to accurately measure POCD, we focussed on delirium as a more relevant secondary end point.

2. **Expert reviewer:** The authors suggest that mild hypercapnia results in higher intraoperative rSO_2 values, resulting in a decreased occurrence of postoperative delirium.

   Authors' response: Thank you for the comment. We have found the mild hypercapnia results in higher intraoperative rSO_2 values and that the effect was sustained throughout the intraoperative period. As a secondary outcome, we have also observed that in the TMH group, the incidence of postoperative delirium was less than that of TN group.

3. **Expert reviewer:** The authors should provide further elucidation why an increase in cerebral rSO_2 intraoperatively would be a wanted effect.

   Authors' response: Thank you for this important comment. We performed a randomized controlled trial to investigate the relationship between mild hypercapnia and regional cerebral oxygen saturation. $EtCO_2$ level was adjusted to achieve $PaCO_2$ levels corresponding to group allocation. The attending anaesthetist had no specific rSO_2 target to titrate to and as our expert reviewer correctly pointed out, evidence in optimizing rSO_2 is ambiguous. For example, Rogers et al. found no clinical benefit in optimizing NIRS directed management of rSO_2 in patients undergoing cardiopulmonary bypass.[26] On the other hand, there are clear evidence that low rSO_2 is linked to cerebral ischaemia and neurological complications.[16,27,28] The threshold for identifying cerebral ischemia may be influenced by a number of patient-specific or technology-dependent variables and unfortunately, there is no universally accepted threshold to identify pathological cerebral saturation. Numerous studies have utilised threshold of approximately 20% decrease in rSO_2 as the threshold for abnormal rSO_2. [16,17] Levy et al. compared changes in oxygen saturation with electroencephalographic evidence of cerebral ischaemia and estimated an ischaemic threshold of 47% in cerebral saturation.[18] As such, we defined a 15% decrease in rSO_2 a clinically important reduction in rSO_2 but we have not defined a clinically significant level of improvement in rSO_2 from baseline. Thank you for this very important comment once again.

4. **Expert reviewer:** Introduction, line 29 '' Numerous studies have shown that NIRS can be applied clinically in the resuscitation and cardiac surgery settings where cerebral desaturation events can be both effectively monitored and managed.10-13'': the authors imply that there are clear protocols for interpreting desaturations. This is to date not the case. Absolute and relative

thresholds theoretically requiring prompt intervention have been proposed, however these have not been validated nor is there any consensus on which interventions are indicated and when.

Authors' response: This is an excellent point and we agree with the Reviewer. We have accordingly modified the introduction section of our manuscript and state "*Numerous studies have suggested that NIRS can be applied clinically in the resuscitation and cardiac surgery settings where cerebral desaturation events can be both effectively monitored and managed. However, whilst absolute and relative saturation thresholds theoretically requiring prompt interventions have been proposed, these thresholds have not been validated and there is a lack of consensus on the indication and timing of interventions.*"

5. **Expert reviewer:** Methods section, line 30: '' age greater than 18 years'' should be ''age above 18 years''.

   Authors' response: Thank you for this excellent suggestion. The correction has been made in the revised manuscript under the methods section.

   "Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care."

6. **Expert reviewer:** Methods section, line 39: there is a comma missing after the word ''isolation''. Do the authors mean intraoperative single lung ventilation?

   Authors' response: Please accept our apologies for the oversight. The correction has been made in the revised manuscript under the methods section.

   "Exclusion criteria included patients undergoing cardiac surgery, procedures requiring one lung isolation, liver transplantation, intracranial surgery, GCS less than 15, known cognitive impairment, intellectual disability or a mental illness, moderate pulmonary hypertension (mean pulmonary arterial pressure greater than 40 mmHg), and American Society of Anesthesiology status V."

7. **Expert reviewer:** Methods section: was the anesthesiologist or any of the staff members blinded to the current rSO_2 readings?

   Authors' response: Thank you for this question. The attending anaesthetist was not blinded to the intervention nor rSO_2 readings. However, study participants, surgeons, all peri-operative staff and the researcher responsible for collecting outcome data were blinded to the treatment allocation. The above was stated under the methods section.

   "Study participants, surgeons, and all peri-operative staff were blinded to treatment allocation. However, it was not possible to blind the attending anaesthetist who was responsible for delivery of the intervention."

"Data collection for all the trial outcomes was collected by an independent researcher blinded to treatment allocation. The sequence was decoded after the data was analysed. The anaesthetist delivering the intervention did not participate in the assessment of postoperative delirium."

8. **Expert reviewer:** Methods section: did the authors encounter any problem in positioning the cerebral rSO_2 sensors, since the anesthesia protocol mentions routine use of BIS?

   Authors' response: Thank you for this important question. The cerebral oximeter probes were placed just adjacent to the BIS strip according to manufacturer's instructions. We have not encountered any problems related to the positioning of the probes. However, there were signal dropouts in a number of patients, and this could partly be related to sub-optimal probe attachment to the forehead and partly related to the sub-optimal connection to the SedLine® brain function monitor.

   We have stated in the manuscript under the results section "*Two participants in the hypercapnic group had failure of bilateral probe attachment and lead connection problem that were unable to be rescued. These patients were subsequently excluded from the analyses of oxygenation as no rSO2 data were captured and were included in the analysis of all other variables and endpoints. In the hypercapnic group, three participants had unilateral discontinuous oximetry readings due to intermittent signal dropout. In the normocapnic group, signal dropout occurred in two patients on the left side. The corresponding data were excluded.*"

9. **Expert reviewer:** Statistical analyses: why did the authors use the Kruskal Wallis test for non-parametric data? The Mann-Whitney U test would suffice since only two independent groups are being tested.

   Authors' response: Thank you for raising this important question. As our expert reviewer correctly pointed out, even though the Kruskal Wallis test and the Mann-Whitney U test are both non-parametric statistical tests, the Kruskal Wallis test is used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. It can be considered an extension to the Mann-Whitney U test which only allows comparison between two variables.

   We have reviewed our data analysis process and have confirmed that all non-parametric continuous data in our study were compared by the Mann-Whitney U test. We have now modified the statement in the methods section to *"Parametric continuous data were compared by the Student's t-test, and non-parametric continuous data were compared by the Mann-Whitney U test."*

10. **Expert reviewer:** Results section, line 17-18: how did the authors calculate the treatment effect? The mean difference between the TMH and TN group does not seem to be 19% in the left cerebral hemisphere.

    Authors' response: Thank you for this important question. As we have stated under the methods section in the manuscript, for the primary outcome we compared the absolute difference between the TMH and TN groups in percentage change in rSO2 from baseline to completion of surgery using an unpaired, two-tailed t-test. A more detailed longitudinal analysis of time-by-treatment interaction was also conducted using a random effect generalized least squares regression model due to the repeated measures nature of the data. The mean difference between the TMH and TN

group was calculated to be 19%. This figure represents the mean percentage difference between the rSO_2 values at the start and on completion of surgery. Important to note that the figure is presented as mean (standard deviation), which means exaggerated difference, which occurred in small number of patients with duration of surgery up to 480 minutes, could make a sizable contribution to the mean percentage difference.

11. **Expert reviewer:** Results section: rather than only describing the results from the longitudinal time-by-treatment interaction, it would be nice to visualize the relationship between delta rSO_2 and time.

    Authors' response: Thank you for the valuable suggestion. We agree with our expert reviewer that a graphical representation of percentage change in rSO_2 over time could assist with visualization of the relationship. This suggestion was discussed amongst our authors in a number of meetings, which also included a consultation with the Professor of Biostatistics at the Melbourne Medical School. In order to preserve the robustness of the analysis, as well as to prevent misinterpretation of the data, the consensus from our author group is to adhere to the original methodology and present percentage change in rSO_2 from baseline of the two groups in two separate line series. This would enable readers to differentiate whether the delta rSO_2 originated from an increase in rSO_2 or a decrease in rSO_2 from baseline. As evident in our study, the sustained difference in rSO_2 between TMN and TN was a combined effect of an increase in rSO_2 from baseline in the TMH group and a decrease in rSO_2 from baseline in the TN group.

    We have modified the Introduction, Result, and Conclusion sections and we now state *"targeted mild hypercapnia was associated with a stable increase in regional cerebral oxygen saturation from baseline while targeted normocapnia was associated with a decrease in regional cerebral oxygen saturation from baseline in both hemispheres"*

12. **Expert reviewer:** The authors elaborate on the potential mechanism behind the effect of hypercapnia on cerebral rSO_2. An important part of the mechanism which is not included in the discussion section is cerebral autoregulation. This is extremely important, since this autoregulatory mechanism is influenced by the pCO_2 state as well as other factors such as hemodilution. This has been described by Ševerdija and Vranken.
    (Ševerdija EE et al. Hemodilution combined with hypercapnia impairs cerebral autoregulation during normothermic cardiopulmonary bypass. J Cardiothorac Vasc Anesth 2015;29:1194-1199 & Vranken NPA et al. Cerebral oximetry and autoregulation during cardiopulmonary bypass: a review. J Extra Corpor Technol 2017;49:182-191.)

    Authors' response: Thank you for this excellent comment. Ševerdija et al. performed a prospective interventional study on patients during normothermic cardiopulmonary bypass and found that haemodilution combined with hypercapnia impaired cerebral autoregulation. On the other hand, in the normocapnia and hypocapnia groups, there were no differences in autoregulation index.

    Dr Vranken and colleague performed a comprehensive systematic review on the topic of cerebral autoregulation and its significance in determining postoperative neurological complications following cardiopulmonary bypass. It was found that maintaining an intact cerebral autoregulatory activity can contribute to decrease in occurrence of post operative neurological complications.

    Autoregulation of cerebral blood flow is the ability of the brain to maintain relatively constant blood flow despite changes in perfusion pressure. [29] The mechanism behind dynamic cerebral autoregulation is complex and is not completely understood. Our study focused on measuring

rSO_2 rather than cerebral blood flow, and therefore, we have not collected important metrics to inform the effects of mild hypercapnia on cerebral blood flow or cerebral autoregulation. Nevertheless, we agree completely with Dr Vranken that cerebral autoregulation can potentially affect rSO_2 values. We have now modified the manuscript and included the two excellent references to support our statements.

"Numerous factors, for instance, cardiac output, oxygen affinity of haemoglobin, cerebral autoregulation, and the ratio of cerebral arterial to venous blood volume, affect rSO_2 in the setting of hypercapnia, but changes in PaCO2 and CBF, in turn, have direct influence on these factors"

"In the literature, the association between normocapnia and reduced CBF and lower levels of rSO_2 were reported briefly. Normocapnia was also found to be superior in preserving cerebral autoregulation, however, the exact mechanism and associations between normocapnia and variations in rSO_2 values are not entirely clear."

13. **Expert reviewer:** Moreover, the authors did not include the fact that higher rSO_2 values might in fact be representative of a rather disadvantageous scenario. This has been studied by Henriksen who described the so-called ''brain luxury perfusion'' (Henriksen L. Brain luxury perfusion during cardiopulmonary bypass in humans. A study of the cerebral blood flow response to changes in CO_2, O_2, and blood pressure. Cereb Blood Flow Metab 1986;6:366-378.).

    Authors' response: Thank you for this excellent suggestion, which has resulted in interesting debate amongst our authors. Henriksen performed a study to investigate cerebral blood flow response to changes in $PaCO_2$, PaO_2 and mean arterial blood pressure on patients undergoing cardiopulmonary bypass. It is very difficult to comment on whether higher rSO_2 presents a more disadvantageous scenario as rSO_2 was not measured in Henriksen's study. While brain tissue pO2 and regional cerebral oxygenation are correlated, cerebral blood flow and rSO_2 are not directly related.[30] Furthermore, neurological outcomes were not measured in Henriksen's study. The proposed cerebral blood flow heterogeneity in the setting of hypercapnia was not formally investigated as part of the primary or secondary endpoints in Henriksen's study, and therefore, whether increased cerebral blood flow has any impact on rSO_2 and neurological outcomes remain ambiguous.

Reviewer 5: Dr Cochrane's queries

We thank expert Dr Cochrane for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** The number of measurements of the primary outcome (ΔrSO_2) made throughout the duration of the surgery varies between individuals, therefore Generalised Estimating Equations (GEE) are more appropriate than Repeated Measures (RM) ANOVA.

    Authors' response: Thank you for this important comment. We completely agree with the expert comment from Dr Cochrane that the number of measurements of the primary outcome made throughout the duration of the surgery varies between individuals. As state under the methods section in the manuscript, for the primary outcome, we compared the absolute difference between the targeted normocapnia and targeted mild hypercapnia groups in percentage change in rSO_2 from baseline to completion of surgery using an unpaired, two-tailed t-test. Furthermore, a more

detailed longitudinal analysis of time-by-treatment interaction was also conducted using a random effect generalised least squares regression model.

2. **Expert reviewer:** The effect of time on the dependent variable is clearly not linear

   Authors' response: Thank you for the excellent comment. The primary outcome of this trial has been prespecified as the absolute difference between the targeted normocapnia and targeted mild hypercapnia groups in percentage change in rSO_2 from baseline to completion of surgery. The nature of the effect of time on the outcome measure is not relevant for any outcome that is measured at a prespecified time point, as both groups are compared at the same timepoint on either a value of an outcome of interest or the change from the baseline. The same argument applies to all the prespecified secondary outcomes that are measured at prespecified single timepoints, which include the incidence of postoperative delirium, intra-operative pH, bicarbonate, base excess, serum potassium, and length of hospital stay.

   The nature of the relationship between the time and rSO_2 is only relevant for the exploratory analyses of longitudinal changes (time by treatment interaction), using multiple timepoints that was conducted with random effect regression modeling. We agree with the reviewer that at some time points, there are clear and abrupt increases/decreases in the rSO_2 values, but linear trend fits the data better than any other analytical trend typically used in such situations (polynomial or exponential). Since the objective of this particular analysis is to investigate the statistical interaction, i.e. whether the trends over time are different between the treatment groups, we believe that the linear approximation of the trends would be more valid than any other alternatives, despite the potentially approximate nature of this finding. We have added a cautionary statement into the interpretation of the interaction test results as follows "*the strong nature of interaction between treatment and time for RSO2 outcome should be treated with caution due to the potential minor departures of the data from the linear trend*".

3. **Expert reviewer:** X^2 methods are not appropriate for ordinal data (Protocol Page 23).

   Authors' response: Thank you for raising this important error. The protocol has now been corrected and we acknowledge this oversight. We have changed the protocol to state that for ordinal data the Mann-Whitney test will the difference between two treatments using data from an independent-measures design; that is, two separate samples, or Wilcoxon test will be used to evaluate the difference between two treatment conditions using data from a repeated-measures design; that is, the same sample is tested/measured in both treatment conditions.

4. **Expert reviewer:** A key to understanding the figures in [], {} in Table 2 would be helpful.

   Authors' response: Thank you for this suggestion. We have included this information in the footnote under Table 2 for brevity of the Table title. We are more than happy to move the statement *"Data are reported as mean (standard deviation) {sample size}"* to the title of the table if requested

5. **Expert reviewer:** ±SD is not a meaningful statistical expression. It is more appropriate to use either (SD), (SE) or a 95% confidence interval.

   Authors' response: Thank you for this excellent comment. We agree that expressing the results in terms of standard deviation is more appropriate and have made the modification in the revised manuscript.

   We have now stated in the Results section

   "The average (standard deviation, SD) percentage changes in rSO_2 from the baseline to the conclusion of the surgery in TMH group were +8.56% (18.90%) on the left and +13.86% (18.17%) on the right, and in TN group they were -6.18% (17.24%) on the left and -5.48% (18.94%) on the right. The resulting treatment effects were 19% (95% CI [9.2 to 28.8]; P<0.001) on the left and 19% (95% CI [10.9 to 27.0]; P<0.001) on the right (Table 2). "

6. **Expert reviewer:** As mentioned above, the lack of statistically significant between-group differences in baseline and post-operative characteristics (Tables 1 and 3, respectively) could be due to low power. The duration of surgery is of particular note: the median in the TN group is only 66% of that in the TMH group (Table 1) and this should be addressed.

   Authors' response: Thank you for raising this important point. We completely agree with the reviewer that the observed difference in the duration of surgery between groups needs to be acknowledged despite the lack of statistical significance. We have therefore modified the statement in the results section to the following *"With regards to surgical characteristics, median duration of surgery was longer in the TMN group with median [IQR] duration of 219 min [124 to 304] versus 144 min [108 to 218] in the TN group (P=0.121)."*

7. **Expert reviewer:** Similarly, the total dose of intra-operative opioid in TN patients is 77% of that in the TMH group (Table 3).

   Authors' response: Thank you for this excellent comment. There are many factors that can potentially influence opioid requirements, such as age of patient, type of surgery and co-morbidities of patient etcetera. On further review of the literature, the relationship between pain and cerebral oxygenation has not been fully investigated. We are grateful for Dr Cochrane to point out this interesting observation. As suggested, we have discussed this further in our discussion section. We have now stated *"Whilst our study was not designed to measure differences in analgesia and partial pressure of oxygen in arterial blood, we observed a 10% higher median PaO_2 level in the TMH group and found that the median intraoperative analgesia requirements were also approximately 30% higher. Both arterial oxygen levels and pain have been reported to influence tissue oxygenation [9], which was not directly measured in our study. The effect of pain on cerebral oxygenation is unclear, and has not be borne out in clinical studies [10]; further studies exploring this association are needed.*

8. **Expert reviewer:** Although the comparison of base excess in the two groups was not significant at the 5% level, there is some evidence of a difference (Table 4, P=0.069). Over-reliance on the 5% figure, particularly when a study is not specifically powered to evaluate differences, can be misleading.

Authors' response: Thank you for this excellent observation. Base excess is defined as the amount of acid required to be added to each liter of fully oxygenated blood to return to pH of 7.4 at temperature of 37 degrees Celsius with $PaCO_2$ of 40. The normal range of base excess is between -3 to +3. It can be a useful marker to differentiate between a respiratory cause and a metabolic cause of acid base disturbance. Delivering TMH induces mild respiratory acidosis. However, due to the acute nature of the delivery of intervention, metabolic compensation is unlikely to have taken place. This is confirmed by the normal bicarbonate levels in both TMH and TN groups. In our study, base excess was observed to be -1.00 [-2.50 to 0.25] in the TMH group and 1.00 [-0.88 to 2.00] in the TN group. They both lie within the normal limits

9. **Expert reviewer:** There is a higher proportion of COPD cases in the TMH group (Table 1, P=0.047).

   Authors' response: Thank you for this excellent comment. This is an important observation because patients with COPD could have chronically elevated $PaCO_2$ levels. In our study, there was 100% compliance to the group allocation, that is, all patients in the TMN group had $PaCO_2$ level between 44 and 55mmHg. Accordingly, we have stated in the results section *"In terms of co-morbidities, both groups were similar except for the presence of chronic obstructive pulmonary disease. There was 100% compliance to the designated PaCO2 intra-operative targets."*

10. **Expert reviewer:** Due to the doubt about the reliability of the between-group comparisons, relevant independent variables should also be included in analyses (most notably of ΔrSO_2 and post-operative delirium). Univariate relationships between the dependent and independent variables, and inter-relationships between the independent variables, should first be investigated before being included in a multivariable model. Full details of the model(s) implemented should be presented.

    Authors' response: Thank you for this thoughtful comment. This comment raises two separate points. First there may be a potential confusion in regards to the nature of the ΔrSO_2 as the reviewer refers to it as an independent variable, while it is a prespecified primary outcome in our study.

    Second, due to the prospective randomized nature of the study all the outcomes and their analysis are prospective prespecified. Relevant covariate adjustments, if needed, were included in the analysis planning prior to data unblinding. The hypothesis generating epidemiological approach of identifying potential relevant covariates using univariate analysis with the aim to include these into a final analysis is not appropriate for a hypothesis testing randomised study where both known and unknown confounders can only appear imbalanced due to a chance. These principles are fully described in both FDA guidelines for covariate adjustment in clinical trials (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adjusting-covariates-randomized-clinical-trials-drugs-and-biologics-continuous-outcomes-guidance) and (https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf).

11. **Expert reviewer:** Are the statistically significant difference noted in Table 4 clinically significant?

    Authors' response: Thank you for this important question. The statistically significant difference in $PaCO_2$ confirms that there was 100% compliance to the designated $PaCO_2$ intra-operative targets, indicating that clinically the intended intervention was delivered. As a result of the delivery

of TMH, lower intra-operative pH was observed in the TMH group. Clinical concern of hypercapnia-induced acidosis is potential development of hyperkalaemia. Importantly, we observed no statistically significant difference in the potassium level between the groups. The bicarbonate was marginally higher in TMH signifying a potentially small metabolic compensation process in the setting of the mild acute hypercapnia-induced acidosis.

The above was discussed in the Results section and in the Discussion section.
"There was 100% compliance to the designated $PaCO_2$ intra-operative targets. The median [IQR] $PaCO_2$ in the TMH group and TN groups were 51.5 mmHg [46.9 to 60.9] and 34.8 mmHg [32.8 to 38.1] respectively (P<0.001)"

"A clinical concern of mild hypercapnia is hypercapnic-induced acidosis and the subsequent development of hyperkalaemia. Whilst a linear correlation between arterial carbon dioxide and plasma pH is well reported, the relationship between acute hypercarbia, respiratory acidosis and plasma potassium is also poorly understood. In the present study, we found no association between hypercarbia and serum potassium concentrations, a finding also supported by others."

12. **Expert reviewer:** Given the between-hemisphere differences, it would be more appropriate for the figures in Table 4 to be presented separately for right and left.

    Authors' response: Thank you for this suggestion. Clinically, the arterial blood gas values are systemic measurements of the acid base status and partial pressure of oxygen and carbon dioxide status. Unless there is a significant arterial blood flow restriction to one of the hemispheres, these measurements are homogenous between the hemispheres.

13. **Expert reviewer:** Inclusion criteria were stated in the protocol (Page 5) as adults aged > 65 years and exclusion criteria (Page 6) as adults aged < 65 years. Are adults aged 65 included or excluded?

    Authors' response: Thank you for this important question. We have amended the inclusion criteria in the Trial Registry as previously the information presented was not consistent. The inclusion criterion is adult patients with age over 18 years.

    This is stated in the manuscript under the methods section *"Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care."*

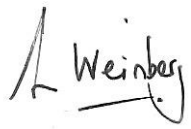
14. **Expert reviewer:** In contrast, the manuscript inclusion criteria were stated (Page 5, Line 30) as adults aged > 18 years with no age-related exclusion criteria.

    Authors' response: Please accept our apologies for the confusion and thank you once again for this important question. We have amended the inclusion criteria in the Trial Registry as previously the information presented was not consistent. The inclusion criterion is adult patients with age over 18 years.

    This is stated in the manuscript under the methods section *"Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure*

monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care."

Thank you for taking the time to review and consider the above manuscript for publication in BMJ Open



A/Prof Laurence Weinberg

BSc, MBBCh,MRCP,DPCritCareEcho,FANZCA,MD

Director, Department of Anesthesia, Austin Hospital

Associate Professor, Department of Surgery, University of Melbourne

Associate Professor, Perioperative Pain and Medicine Unit, Department of Surgery, University of Melbourne

References

1. Chapin JC, Downs JB, Douglas ME, Murphy EJ, Ruiz BC. Lung Expansion, Airway Pressure Transmission, and Positive End-Expiratory Pressure. *JAMA Surgery.* 1979;114(10):1193-1197.
2. Kingma I, Smiseth OA, Frais MA, Smith ER, Tyberg JV. Left ventricular external constraint: relationship between pericardial, pleural and esophageal pressures during positive end-expiratory pressure and volume loading in dogs. *Annals of biomedical engineering.* 1987;15(3-4):331-346.
3. Gattinoni L, Pesenti A, Baglioni S, Vitale G, Rivolta M, Pelosi P. Inflammatory pulmonary edema and positive end-expiratory pressure: correlations between imaging and physiologic studies. *Journal of thoracic imaging.* 1988;3(3):59-64.
4. Mirro R, Busija D, Green R, Leffler C. Relationship between mean airway pressure, cardiac output, and organ blood flow with normal and decreased respiratory compliance. *The Journal of pediatrics.* 1987;111(1):101-106.
5. O'Quin R, Marini JJ. Pulmonary artery occlusion pressure: clinical physiology, measurement, and interpretation. *The American review of respiratory disease.* 1983;128(2):319-326.
6. Luecke T, Pelosi P. Clinical review: Positive end-expiratory pressure and cardiac output. *Critical care (London, England).* 2005;9(6):607-621.
7. Klinger JR. Hemodynamics and positive end-expiratory pressure in critically ill patients. *Critical care clinics.* 1996;12(4):841-864.
8. Pinsky M, Vincent JL, De Smet JM. Estimating left ventricular filling pressure during positive end-expiratory pressure in humans. *The American review of respiratory disease.* 1991;143(1):25-31.
9. Akca O. Pain and Tissue Oxygenation. *Critical care medicine.* 2015;43:e462-463.
10. Hoiseth LO, Hisdal J, Hoff IE, Hagen OA, Landsverk SA, Kirkeboen KA. Tissue oxygen saturation and finger perfusion index in central hypovolemia: influence of pain. *Crit Care Med.* 2015;43(4):747-756.
11. Glatte HA, Jr., Welch BE. Carbon dioxide tolerance: a review. *Aeromedical reviews.* 1967;5:1-28.
12. Brackett NC, Jr., Cohen JJ, Schwartz WB. CARBON DIOXIDE TITRATION CURVE OF NORMAL MAN. EFFECT OF INCREASING DEGREES OF ACUTE HYPERCAPNIA ON ACID-BASE EQUILIBRIUM. *The New England journal of medicine.* 1965;272:6-12.
13. Schwartz WB, Brackett NC, Jr., Cohen JJ. THE RESPONSE OF EXTRACELLULAR HYDROGEN ION CONCENTRATION TO GRADED DEGREES OF CHRONIC HYPERCAPNIA: THE PHYSIOLOGIC LIMITS OF THE DEFENSE OF PH. *The Journal of clinical investigation.* 1965;44:291-301.
14. Akca O, Liem E, Suleman MI, Doufas AG, Galandiuk S, Sessler DI. Effect of intra-operative end-tidal carbon dioxide partial pressure on tissue oxygenation. *Anaesthesia.* 2003;58(6):536-542.
15. Akça O, Sessler DI, Delong D, Keijner R, Ganzel B, Doufas AG. Tissue oxygenation response to mild hypercapnia during cardiopulmonary bypass with constant pump output. *Br J Anaesth.* 2006;96(6):708-714.
16. Murphy GS, Szokol JW, Marymont JH, et al. Cerebral oxygen desaturation events assessed by near-infrared spectroscopy during shoulder arthroscopy in the beach chair and lateral decubitus positions. *Anesthesia and analgesia.* 2010;111(2):496-505.
17. Denault A, Deschamps A, Murkin JM. A proposed algorithm for the intraoperative use of cerebral near-infrared spectroscopy. *Seminars in cardiothoracic and vascular anesthesia.* 2007;11(4):274-281.
18. Levy WJ, MD, Levin S, MD, Chance B, PhD. Near-infrared Measurement of Cerebral Oxygenation: Correlation with Electroencephalographic Ischemia during Ventricular Fibrillation. *Anesthesiology: The Journal of the American Society of Anesthesiologists.* 1995;83(4):738-746.
19. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Annals of internal medicine.* 1990;113(12):941-948.

20. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. *Journal of the American Geriatrics Society.* 2008;56(5):823-830.
21. Cunningham J, Kim LD. Post-operative delirium: a review of diagnosis and treatment strategies. *Journal of Xiangya Medicine.* 2018;3(2).
22. Rogers CA, Stoica S, Ellis L, et al. Randomized trial of near-infrared spectroscopy for personalized optimization of cerebral tissue oxygenation during cardiac surgery. *Br J Anaesth.* 2017;119(3):384-393.
23. Deiner S, Silverstein JH. Postoperative delirium and cognitive dysfunction. *BJA: British Journal of Anaesthesia.* 2009;103(suppl_1):i41-i46.
24. Newman S, Stygall J, Hirani S, Shaefi S, Maze M. Postoperative cognitive dysfunction after noncardiac surgery: a systematic review. *Anesthesiology.* 2007;106(3):572-590.
25. Rasmussen LS, Larsen K, Houx P, Skovgaard LT, Hanning CD, Moller JT. The assessment of postoperative cognitive function. *Acta anaesthesiologica Scandinavica.* 2001;45(3):275-289.
26. Rogers CA, Stoica S, Ellis L, et al. Randomized trial of near-infrared spectroscopy for personalized optimization of cerebral tissue oxygenation during cardiac surgery. *Br J Anaesth.* 2017;119(3):384-393.
27. Slater JP, Guarino T, Stack J, et al. Cerebral oxygen desaturation predicts cognitive decline and longer hospital stay after cardiac surgery. *The Annals of thoracic surgery.* 2009;87(1):36-44; discussion 44-35.
28. Gale SD, Hopkins RO. Effects of hypoxia on the brain: neuroimaging and neuropsychological findings following carbon monoxide poisoning and obstructive sleep apnea. *Journal of the International Neuropsychological Society : JINS.* 2004;10(1):60-71.
29. Paulson OB, Strandgaard S, Edvinsson L. Cerebral autoregulation. *Cerebrovascular and brain metabolism reviews.* 1990;2(2):161-192.
30. Saldien V, Schepens T, Vanlinthout L, et al. Real-Time Monitoring of Cerebral Blood Flow and Cerebral Oxygenation During Rapid Ventricular Pacing in Neurovascular Surgery: A Pilot Study. *Journal of neurosurgical anesthesiology.* 2019.
31. Benfratello L. Random Effects Regression for Panel Data. In: Michalos AC, ed. *Encyclopedia of Quality of Life and Well-Being Research.* Dordrecht: Springer Netherlands; 2014:5387-5389.
32. Diggle P, Diggle P. *Analysis of longitudinal data.* 2nd ed. Oxford ; New York: Oxford University Press; 2002.

VERSION 2 – REVIEW

REVIEWER	Ozan Akca, MD, FCCM University of Louisville, Kentucky, USA
REVIEW RETURNED	08-Sep-2019

GENERAL COMMENTS	Title - A randomised controlled trial to investigate the relationship between mild hypercapnia and cerebral oxygen saturation in patients undergoing major surgery Manuscript ID: bmjopen-2019-029159 Thanks much for submitting the revised version of the manuscript to the Journal, and allowing me to review it one more time. Please see my further and minor comments below: • Abstract – Participants – please correct “TMN” to “TMH”. • Please report the “age>65y” (elderly patient percentage for each group) in Table 1. • If the data available, there is some value in reporting the intraoperative positioning of the patient. Lateral and trendelenburg positioning may impact the cerebral oxygenation data. Gravitational effect may contribute to further venous pooling favoring the

	dominant side.

REVIEWER	Nousjka Vranken Maastricht University Medical Centre, the Netherlands
REVIEW RETURNED	18-Sep-2019

GENERAL COMMENTS	Abstract: the sentence ‘’ The difference in %ΔrSO2 on both hemispheres between the two groups diverged with time with TMH exhibiting smaller percentage decrease over time compared to TN’’ is difficult to read and easy to misinterpret, please rephrase. Last sentence in the abstract: ‘’of TMH’’ should be ‘’on TMH’’. Page 4, line 23: the last part of the sentence ‘’ rSO2 measurements rely on the assumption that rSO2 is homogenous in the brain’’ is not correctly formulated (grammar). Page 5, line 15: increase of cerebral blood flow is not necessarily part of cerebral protection. Cerebral hyperperfusion, for example, is rather harmful. The references to support this statement are lacking. The last paragraph of the introduction lacks justification for the proposed link with postoperative delirium. Do the authors hypothesize that delirium is (generally) a complication caused by cerebral ischemia? And that cerebral blood flow, oxygen delivery, cerebral glucose utilisation and oxidative metabolism, and activation of ATP-sensitive potassium channels are all factorsd affecting the risk of postoperative delirium? Page 7, lines 35-36; ‘’ trend oximetry data is defined as the change in regional oxygen saturation value measured by the oximetry probes’’. The difference between which time points? Page 7, line 38: ‘’tested against’’ is not correct English. Page 7, line 42-45: the sentence suggests that in total four sensors were placed on the patients’ forehead. What was the reason to only subtract absolute values? are trend values only meaningful during surgery in the context of patient monitoring? Please specify. Page 7 line 48: what do the authors mean by ‘’ Regional cerebral oxygen saturation was collected’’? ‘’collecting’’ rSO2 is not correct English. Page 8, line 12: use the CAM abbreviation. Page 8, line 17: ‘’carers’’ is not the correct term, ‘’caretakers’’ is. Page 8, line 38: an arterial blood gas samples at the end of surgery does not prove that a certain arterial pCO2 was maintained throughout surgery. Page 8, line 50-51: use the beforementioned abbreviations for pCO2, pO2, etc. Page 9, line 19: use the beforementioned abbreviation for hb. Page 9, line 42: the abbrevation ECG is used without previous

	clarification. Page 9, line 45: '' BIS reading'' should be ''BIS readings''. Page 10, line 12: first, the authors dicuss a relative change of 15% to be relevant, which translates into 12% absolute change in case of a baseline reading of 80%. What do the authors mean by ''common'' standarddeviation of 10%? Page 10. The sentence '' For the primary outcome we compared the absolute difference between the TMH and TN groups in percentage change in rSO2 from baseline to completion of surgery using an unpaired, two-tailed t-test.'' is unnecessary since it is a repitition of previous statements on the statistical analyses. Page 11. The sentence '' The study was designed to investigation...'' is not correct English. Page 11. The paragraph '' Patient and Public Involvement'' is unnecessary and does not add valuable information to the manuscript. Page 12. the sentence ''These patients were subsequently excluded from the analyses of oxygenation as no rSO2 data were captured and were included in the analysis of all other variables and endpoints.'' is difficult to read, please rephrase. Using the word 'were'' three times should be avoided. Page 12: in the results section, Table 1 should be mentioned/referred to in the text together with the statement on statistical significant differences between the two groups. Table 1: what are ''other comorbidities''? please specify in the legend. Results section: tables: there is no need to use letters in superscript for the abbreviations used in the table, just specifiy them in alphabetical order in the legend. Methods and results section: the exact meaning (and calculation of) delta rSO% is unclear. Did the authors calculate the difference between baseline rSO2% and the mean rSO2 (throughout surgery) per patient? Methods section: what is the justification for data sampling every 15 minutes / every 2 hours? Results section, Table 2: the abbreviation ''NA'' is not explained. Depicting rSO2% every 15 minutes or every 2 hours is probably not relevant and difficult to read in table form. Graphical representation using a line graph would be more illustrative. Page 17, line 35: '' was also similar between the two groups without statistically significant difference'' contains a lot of repition of the same concept (no differences were found). ''Similar'' and ''without difference'' mean the same. Discussion (page 19): do not reintroduce the beforementioned abbreviations. Page 19 ''oxygen affinity of haemoglobin'' is incorrect. Do the authors mean haemoglobin affinity for oxygen?

	Page 19, line 26-31: this sentence is poorly formulated, please rephrase. Page 20. What do the authors mean by uding the sentence ‘’ however, the exact mechanism and associations between normocapnia and variations in rSO2 values are not entirely clear.’’? Normocapnia specifically enables variations in rSO2 compared to hypercapnia, since carbon dioxide forces carebral vasodilatation and therefore loss of autoregulatory function. Discussion section, page 20, second paragraph to page 21. From the text it appears that the authors descibe the differences between patients who did and patients who did not have a delirium postoperatively. If this is the case, then results should be first described in the results section before being mentioned in the discussion. Page 21, line 35: the term ‘’hypercarbia’’ is less commonly used in medical literature. Please stick to the term initially chosed to desribe increased levels of partial gas pressure for carbon dioxide. Page 22, line 56. The authors correctly identify the assumption that monitoring regional oxygen saturation in the prefrontal cortex is representative of global cerebral oxygen saturation. Why do the authors specifically mention the posterior circulation? Page 23, line 26. The effect size of 0.13 is mentioned in the discussion section, while it is not mentioned earlier in the manuscript text. The conclusion is too long. The authors already provided a summary, which makes repitition in the conclusion section an unnecessary. Author contributions: ‘’writing up" is not correct English. General comments: Very interesting study, the authors succeeded to describe and justify the study design into great detail. There are however a few methodological, but mostly textual comments I would like to share. In general, I would like to advise the authors to consult a native American speaker to proof-read the manuscript to increase readabiliuty and to correct grammatical errors. Spaces between references and text in the text body, double spaces. Avoid starting a sentence with an abbreviation.

REVIEWER	Lynda Cochrane Clinical Statistics Consultants UK
REVIEW RETURNED	10-Sep-2019

GENERAL COMMENTS	Thank you to the authors for their positive and appropriate responses to the issues raised. I would encourage them to carry out sensitivity analyses to confirm the robustness of their findings. My detailed comments are below.

	Reviewer 5: Dr Cochrane's queries We thank expert Dr Cochrane for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided. 1. Expert reviewer: The number of measurements of the primary outcome (ΔrSO2) made throughout the duration of the surgery varies between individuals, therefore Generalised Estimating Equations (GEE) are more appropriate than Repeated Measures (RM) ANOVA. Authors' response: Thank you for this important comment. We completely agree with the expert comment from Dr Cochrane that the number of measurements of the primary outcome made throughout the duration of the surgery varies between individuals. As state under the methods section in the manuscript, for the primary outcome, we compared the absolute difference between the targeted normocapnia and targeted mild hypercapnia groups in percentage change in rSO2 from baseline to completion of surgery using an unpaired, two-tailed t-test. Furthermore, a more detailed longitudinal analysis of time-by-treatment interaction was also conducted using a random effect generalised least squares regression model. t tests do not allow for the inclusion of independent variables which can affect the primary outcome (ΔrSO2). This is important because there are between-group differences at play in this study, as shown in Table 1. Carrying out additional analyses would add weight to the authors' findings. The authors state that GLS modelling was carried out and it would be helpful if details of this were reported. 2. Expert reviewer: The effect of time on the dependent variable is clearly not linear Authors' response: Thank you for the excellent comment. The primary outcome of this trial has been prespecified as the absolute difference between the targeted normocapnia and targeted mild hypercapnia groups in percentage change in rSO2 from baseline to completion of surgery. The nature of the effect of time on the outcome measure is not relevant for any outcome that is measured at a prespecified time point, as both groups are compared at the same time point on either a value of an outcome of interest or the change from the baseline. The same argument applies to all the prespecified secondary outcomes that are measured at prespecified single time points, which include the incidence of postoperative delirium, intra-operative pH, bicarbonate, base excess, serum potassium, and length of hospital stay. The nature of the relationship between the time and rSO2 is only relevant for the exploratory analyses of longitudinal changes (time by treatment interaction), using multiple time points that was conducted with random effect regression modeling. We agree with the reviewer that at some time points, there are clear and abrupt increases/decreases in the rSO2 values, but linear trend fits the data better than any other analytical trend typically used in such situations (polynomial or exponential). Since the objective of this

	particular analysis is to investigate the statistical interaction, i.e. whether the trends over time are different between the treatment groups, we believe that the linear approximation of the trends would be more valid than any other alternatives, despite the potentially approximate nature of this finding. We have added a cautionary statement into the interpretation of the interaction test results as follows “the strong nature of interaction between treatment and time for RSO2 outcome should be treated with caution due to the potential minor departures of the data from the linear trend”. That is helpful. 3. Expert reviewer: X2 methods are not appropriate for ordinal data (Protocol Page 23). Authors’ response: Thank you for raising this important error. The protocol has now been corrected and we acknowledge this oversight. We have changed the protocol to state that for ordinal data the Mann Whitney test will the difference between two treatments using data from an independent-measures design; that is, two separate samples, or Wilcoxon test will be used to evaluate the difference between two treatment conditions using data from a repeated-measures design; that is, the same sample is tested/measured in both treatment conditions. That is helpful. 4. Expert reviewer: A key to understanding the figures in [], {} in Table 2 would be helpful. Authors’ response: Thank you for this suggestion. We have included this information in the footnote under Table 2 for brevity of the Table title. We are more than happy to move the statement “Data are reported as mean (standard deviation) {sample size}” to the title of the table if requested. That is helpful. Under Table 1, should “age, which is reported as median [range]” read “age, which is reported as mean [range]”? 5. Expert reviewer: ±SD is not a meaningful statistical expression. It is more appropriate to use either (SD), (SE) or a 95% confidence interval. Authors’ response: Thank you for this excellent comment. We agree that expressing the results in terms of standard deviation is more appropriate and have made the modification in the revised manuscript. We have now stated in the Results section. That is helpful. “The average (standard deviation, SD) percentage changes in rSO2 from the baseline to the conclusion of the surgery in TMH group were +8.56% (18.90%) on the left and +13.86% (18.17%) on the right, and in TN group they were -6.18% (17.24%) on the left and -

	5.48% (18.94%) on the right. The resulting treatment effects were 19% (95% CI [9.2 to 28.8]; P<0.001) on the left and 19% (95% CI [10.9 to 27.0]; P<0.001) on the right (Table 2). There are several measures of average (mean, median, mode) and the specific metric reported should be stated; in this case, the mean. 6. Expert reviewer: As mentioned above, the lack of statistically significant between-group differences in baseline and post-operative characteristics (Tables 1 and 3, respectively) could be due to low power. The duration of surgery is of particular note: the median in the TN group is only 66% of that in the TMH group (Table 1) and this should be addressed. Authors' response: Thank you for raising this important point. We completely agree with the reviewer that the observed difference in the duration of surgery between groups needs to be acknowledged despite the lack of statistical significance. We have therefore modified the statement in the results section to the following "With regards to surgical characteristics, median duration of surgery was longer in the TMN group with median [IQR] duration of 219 min [124 to 304] versus 144 min [108 to 218] in the TN group (P=0.121)." Perhaps this could be better expressed as 'With regards to surgical characteristics, the duration of surgery was longer in the TMN group with median [IQR] duration of 219 min [124 to 304] versus 144 min [108 to 218] in the TN group, although this was not significant the 5% level (P=0.121). This should be followed in the discussion by mentioning power. 7. Expert reviewer: Similarly, the total dose of intra-operative opioid in TN patients is 77% of that in the TMH group (Table 3). Authors' response: Thank you for this excellent comment. There are many factors that can potentially influence opioid requirements, such as age of patient, type of surgery and co-morbidities of patient etcetera. On further review of the literature, the relationship between pain and cerebral oxygenation has not been fully investigated. We are grateful for Dr Cochrane to point out this interesting observation. As suggested, we have discussed this further in our discussion section. We have now stated "Whilst our study was not designed to measure differences in analgesia and partial pressure of oxygen in arterial blood, we observed a 10% higher median PaO2 level in the TMH group and found that the median intraoperative analgesia requirements were also approximately 30% higher. Both arterial oxygen levels and pain have been reported to influence tissue oxygenation 9, which was not directly measured in our study. The effect of pain on cerebral oxygenation is unclear, and has not be borne out in clinical studies 10; further studies exploring this association are needed. That is helpful. Should "has not be borne out in clinical studies" read 'has not been borne out in clinical studies'? 8. Expert reviewer: Although the comparison of base excess in the two groups was not significant at the 5% level, there is some

	evidence of a difference (Table 4, P=0.069). Over-reliance on the 5% figure, particularly when a study is not specifically powered to evaluate differences, can be misleading. Authors' response: Thank you for this excellent observation. Base excess is defined as the amount of acid required to be added to each liter of fully oxygenated blood to return to pH of 7.4 at temperature of 37 degrees Celsius with PaCO2 of 40. The normal range of base excess is between -3 to +3. It can be a useful marker to differentiate between a respiratory cause and a metabolic cause of acid base disturbance. Delivering TMH induces mild respiratory acidosis. However, due to the acute nature of the delivery of intervention, metabolic compensation is unlikely to have taken place. This is confirmed by the normal bicarbonate levels in both TMH and TN groups. In our study, base excess was observed to be -1.00 [-2.50 to 0.25] in the TMH group and 1.00 [-0.88 to 2.00] in the TN group. They both lie within the normal limits That is helpful 9. Expert reviewer: There is a higher proportion of COPD cases in the TMH group (Table 1, P=0.047). Authors' response: Thank you for this excellent comment. This is an important observation because patients with COPD could have chronically elevated PaCO2 levels. In our study, there was 100% compliance to the group allocation, that is, all patients in the TMN group had PaCO2 level between 44 and 55mmHg. Accordingly, we have stated in the results section "In terms of co-morbidities, both groups were similar except for the presence of chronic obstructive pulmonary disease. There was 100% compliance to the designated PaCO2 intra-operative targets." That is helpful 10. Expert reviewer: Due to the doubt about the reliability of the between-group comparisons, relevant independent variables should also be included in analyses (most notably of ΔrSO2 and post-operative delirium). Univariate relationships between the dependent and independent variables, and inter-relationships between the independent variables, should first be investigated before being included in a multivariable model. Full details of the model(s) implemented should be presented. Authors' response: Thank you for this thoughtful comment. This comment raises two separate points. First there may be a potential confusion in regards to the nature of the ΔrSO2 as the reviewer refers to it as an independent variable, while it is a prespecified primary outcome in our study. Sorry, this was a typographical error. Second, due to the prospective randomized nature of the study all the outcomes and their analysis are prospective prespecified. Relevant covariate adjustments, if needed, were included in the analysis planning prior to data unblinding. I haven't been able to find any evidence of covariate adjustments in the analysis plan.

	The hypothesis generating epidemiological approach of identifying potential relevant covariates using univariate analysis with the aim to include these into a final analysis is not appropriate for a hypothesis testing randomised study where both known and unknown confounders can only appear imbalanced due to a chance. These principles are fully described in both FDA guidelines for covariate adjustment in clinical trials (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adjusting-covariates-randomized-clinical-trials-drugs-and-biologics-continuous-outcomes-guidance) and (https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf). I am aware of these guidelines which, in addition to the advice on hypothesis testing in randomised trials also state: Baseline imbalance observed post hoc should not be considered an appropriate reason for including a variable as a covariate in the primary analysis. However, conducting exploratory analyses including such variables when large baseline imbalances are observed might be helpful to assess the robustness of the primary analysis. 11. Expert reviewer: Are the statistically significant difference noted in Table 4 clinically significant? Authors' response: Thank you for this important question. The statistically significant difference in PaCO2 confirms that there was 100% compliance to the designated PaCO2 intra-operative targets, indicating that clinically the intended intervention was delivered. As a result of the delivery of TMH, lower intra-operative pH was observed in the TMH group. Clinical concern of hypercapnia-induced acidosis is potential development of hyperkalaemia. Importantly, we observed no statistically significant difference in the potassium level between the groups. The bicarbonate was marginally higher in TMH signifying a potentially small metabolic compensation process in the setting of the mild acute hypercapnia-induced acidosis. That is helpful The above was discussed in the Results section and in the Discussion section. "There was 100% compliance to the designated PaCO2 intra-operative targets. The median [IQR] PaCO2 in the TMH group and TN groups were 51.5 mmHg [46.9 to 60.9] and 34.8 mmHg [32.8 to 38.1] respectively (P<0.001)" "A clinical concern of mild hypercapnia is hypercapnic-induced acidosis and the subsequent development of hyperkalaemia. Whilst a linear correlation between arterial carbon dioxide and plasma pH is well reported, the relationship between acute hypercarbia, respiratory acidosis and plasma potassium is also poorly understood. In the present study, we found no association between hypercarbia and serum potassium concentrations, a finding also supported by others." That is helpful

	12. Expert reviewer: Given the between-hemisphere differences, it would be more appropriate for the figures in Table 4 to be presented separately for right and left. Authors' response: Thank you for this suggestion. Clinically, the arterial blood gas values are systemic measurements of the acid base status and partial pressure of oxygen and carbon dioxide status. Unless there is a significant arterial blood flow restriction to one of the hemispheres, these measurements are homogenous between the hemispheres. As a statistician, I find it difficult to reconcile the two statements "there are the between-hemisphere differences" and "these measurements are homogenous between the hemispheres" but I take on trust other readers will. 13. Expert reviewer: Inclusion criteria were stated in the protocol (Page 5) as adults aged > 65 years and exclusion criteria (Page 6) as adults aged < 65 years. Are adults aged 65 included or excluded? Authors' response: Thank you for this important question. We have amended the inclusion criteria in the Trial Registry as previously the information presented was not consistent. The inclusion criterion is adult patients with age over 18 years. This is stated in the manuscript under the methods section "Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care." That is helpfull 14. Expert reviewer: In contrast, the manuscript inclusion criteria were stated (Page 5, Line 30) as adults aged > 18 years with no age-related exclusion criteria. Authors' response: Please accept our apologies for the confusion and thank you once again for this important question. We have amended the inclusion criteria in the Trial Registry as previously the information presented was not consistent. The inclusion criterion is adult patients with age over 18 years. This is stated in the manuscript under the methods section "Inclusion criteria included the following: adult patients (age over 18 years), surgery of greater than 2 hours expected duration requiring at least one overnight admission, a clinical indication for continuous blood pressure monitoring via an invasive arterial line, and intermittent positive pressure ventilation via an endotracheal tube as part of standard anaesthesia care." That is helpful

VERSION 2 – AUTHOR RESPONSE

Reviewer 1: Professor Akca's queries

We thank expert Professor Akca for his valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** Abstract – Participants – please correct "TMN" to "TMH".

   Authors' response: Thank you for pointing out this critical error, which has now been corrected. As stated above, the manuscript has also been professionally proofread with corrections to sentence structure rectified where appropriate. In the Abstract, we now state *"40 participants were randomised to either a TMH or TN group (20 to each)"*

2. **Expert reviewer:** Please report the "age>65y" (elderly patient percentage for each group) in Table 1.

   Authors' response: Thank you for the excellent suggestion. We have reviewed the original data and included this information in Table 1. Nine out of twenty patients (45%) in the TNH group were older than 65 years of age, and eleven out of twenty patients (55%) in the TN group were older than 65 years of age.

3. **Expert reviewer:** If the data available, there is some value in reporting the intraoperative positioning of the patient. Lateral and trendelenburg positioning may impact the cerebral oxygenation data. Gravitational effect may contribute to further venous pooling favoring the dominant side.

   Authors' response: Thank you for this excellent comment. We have revisited the original surgical and anaesthesia notes where the positioning of the patient is presented as part of the mandatory medical and anaesthesia records. We agree with the Reviewer that there is value in the reporting of the intra-operative positioning of the patient and agree that lateral and Trendelenburg positioning may impact the cerebral oxygenation data. Gravitational effect may contribute to further venous pooling favoring the dominant side. There was one patient from each group who underwent laparoscopic colonic surgery. The positioning for these patients was steep reverse Trendelenburg with minimal tilt. We have reviewed the haemodynamic data from this patient and there were no noticeable changes observed in any of the outcomes we reported. All other patients were positioned in the supine position with ensuring neutral position of the head. There were no patients positioned in the lateral position where gravity may have contributed to venous pooling on the ipsilateral side. These important points have been included in the Results section of the revised manuscript. *"In terms of intra-operative positioning of patients, one patient from each group was positioned in steep reverse Trendelenburg with minimal tilt. All other patients were positioned in the supine position with a neutral head position."* In the discussion section, we have also included the intra-operative position in the discussion of confounding variables. We now state *"In our study, confounding variables, such as MAP, PaO_2, Hb, and intra-operative position, were similar between the TMH and TN groups"*

   Of interest, the effects of patients positioning on cerebral oximetry has been studied previously and the evidence is conflicting. Closhen *et al.* reported that the Trendelenburg position does not impair cerebral oxygenation measured using two different monitors.[1] These authors reported that there was no clinically important decrease in cerebral oxygen saturation in a steep Trendelenburg position combined with CO_2 pneumoperitoneum in patients undergoing robotic assisted prostatic surgery and that extreme positioning seems to be acceptable with regard to cerebral oxygenation. Other authors have reported similar findings. Park *etl.* evaluated the effect of pneumoperitoneum in a 30 degrees Trendelenburg position on cerebral oxygenation using regional cerebral oxygen saturation (rSO_2).[2] In this small study of 32 patients, cerebral oxygenation, as assessed by rSO_2, increased slightly, which suggests that the procedure did not induce cerebral ischaemia. Closhen *et al.* also reported similar findings in a study that evaluated

changes in cerebral oxygen saturation following prone positioning for orthopaedic surgery under general anaesthesia.[3] In contrast, Sørensen H *et al,* performed a systematic literature search of 901 patients from 24 publications.[4] Interestingly, they found that a decrease in NIRS (>15% relative to baseline) manifested with reverse Trendelenburg position and in almost one quarter of especially elderly patients undergoing open surgery, and demonstrated a correlation to hospital stay (LOS). However, if cerebral deoxygenation was reversed promptly, improved postoperative cognitive function (28 versus 26; mini-mental state examination) and reduced LOS (14 versus 23 days) were seen. The effects of cerebral oxygenation in patients undergoing shoulder surgery in beach chair position have also been well described. Aguirre *et al.* showed that patients in the regional anaesthesia group showed significantly less cerebral desaturation events, and drops in regional cerebral oxygen saturation values compared to patients under general anaesthesia.[5] Importantly, none of the patients in our study were positioned in the beach chair position.

Reviewer 4: Dr Vranken's queries

We thank expert Dr Vranken for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** Abstract: the sentence "The difference in %ΔrSO_2 on both hemispheres between the two groups diverged with time with TMH exhibiting smaller percentage decrease over time compared to TN" is difficult to read and easy to misinterpret, please rephrase.

   Authors' response: Thank you for the excellent recommendation. We apologise for the potentially confusing statement within the abstract. We have reviewed the abstract rigorously and in order to improve clarity, we now state:

   "The absolute difference between the two groups in percentage change in rSO_2 from the baseline to the completion of surgery was 19.0% higher in both hemispheres with TMH (P<0.001). On both sides, the percentage change in rSO_2 was greater in the TMH group than the TN group throughout the duration of surgery. The difference became more noticeable over time. Furthermore, postoperative delirium was higher in the TN group (risk difference 0.3, 95% CI [0.1 to 0.5], P=0.02). Length of stay was similar between groups (5 days vs. 5 days; P=0.99)."

2. **Expert reviewer:** Last sentence in the abstract: "of TMH" should be "on TMH''.

   Authors' response: Thank you for the excellent suggestion. We have now corrected the statement to *"Our findings provide the rationale for larger studies on TMH during surgery."*

3. **Expert reviewer:** Page 4, line 23: the last part of the sentence '' rSO_2 measurements rely on the assumption that rSO_2 is homogenous in the brain'' is not correctly formulated (grammar).

   Authors' response: Thank you for the suggestion. We now state "*Interpretation of rSO_2 depends on an assumption that rSO_2 is the same in different regions of the brain."*

4. **Expert reviewer:** Page 5, line 15: increase of cerebral blood flow is not necessarily part of cerebral protection. Cerebral hyperperfusion, for example, is rather harmful. The references to support this statement are lacking.

   Authors' response: Thank you for the excellent observation. We stated *"The recent emergence of near-infrared spectroscopy (NIRS) cerebral oximetry has provided a practical method to measure rSO_2 continuously and non-invasively. This technology has gained substantial supportive evidence in resuscitation, critical care, and surgical applications. Numerous studies*

have shown that NIRS can be applied clinically in the resuscitation and cardiac surgery settings where cerebral desaturation events can be both effectively monitored and managed." As our expert Reviewer correctly pointed out, increase in cerebral blood flow is not necessarily part of cerebral protection. However, our study was not designed to investigate cerebral blood flow nor its relationship with cerebral protection mechanisms. As a result, we have not made conclusive statements under this topic.

5. **Expert reviewer:** The last paragraph of the introduction lacks justification for the proposed link with postoperative delirium. Do the authors hypothesize that delirium is (generally) a complication caused by cerebral ischemia? And that cerebral blood flow, oxygen delivery, cerebral glucose utilisation and oxidative metabolism, and activation of ATP-sensitive potassium channels are all factors affecting the risk of postoperative delirium?

   Authors' response: Thank you for these important questions. In the introduction, we stated *"The neuroprotective mechanisms of mild hypercapnia, whilst not completely understood, have been postulated to be a result of increase in cerebral blood flow, enhancement of oxygen delivery, improvements in cerebral glucose utilisation and oxidative metabolism, and activation of ATP-sensitive potassium channels to maintain normal neuronal activity in the setting of ischemia."* We conducted a randomized controlled trial to investigate the relationship between mild hypercapnia and regional cerebral oxygen saturation. We recorded the incidence of postoperative delirium in both groups as the secondary outcome. However, we did not investigate the relationship of cerebral ischaemia and delirium. We have not investigated any risk factors for postoperative delirium either.

6. **Expert reviewer:** Page 7, lines 35-36; '' trend oximetry data is defined as the change in regional oxygen saturation value measured by the oximetry probes''. The difference between which time points?

   Authors' response: Thank you for the important question. Trend oximetry value is defined as the change in rSO_2 from a user-specified value (usually the baseline rSO_2). We now state *"Absolute oximetry value is defined as rSO_2 value measured by the oximetry probe calibrated by a fixed ratio of arterial to venous blood, whereas trend oximetry value is defined as the change in rSO_2 from a user-specified value (usually the baseline rSO_2)."*

7. **Expert reviewer:** Page 7, line 38: ''tested against'' is not correct English.

   Authors' response: Thank you for pointing this out. We have now corrected the sentence to *"The measurement errors for absolute and trend data are reported to be approximately 4% and 3% respectively when checked against reference blood samples taken from the radial artery and internal jugular bulb vein."*

8. **Expert reviewer:** Page 7, line 42-45: the sentence suggests that in total four sensors were placed on the patients' forehead.

   Authors' response: Thank you for the great comment. We agree that the statement is confusing, and therefore we have rephrased the statement to the following, *"Regional cerebral oxygen saturation was measured in the two hemispheres separately with a NIRS sensor attached to each side of patient's forehead."*

9. **Expert reviewer:** What was the reason to only subtract absolute values? are trend values only meaningful during surgery in the context of patient monitoring? Please specify.

Authors' response: Thank you for the excellent question. The rationale behind subtracting the absolute values is that the percentage change of rSO_2 was calculated based on absolute rSO_2 measurements. Baseline rSO_2 was recorded as an absolute rSO_2 value before induction of anaesthesia. As a consequence, the difference between rSO_2 at any given time point and the baseline rSO_2 could only be computed if absolute values were used.

To improve the clarity of the paragraph, we now state *"Regional cerebral oxygen saturation was measured in the two hemispheres separately with a NIRS sensor attached to each side of patient's forehead. Only the absolute oximetry data were extracted and analysed. The baseline rSO_2 was recorded before commencing any premedication and before induction of anaesthesia. Subsequent rSO_2 measurements were recorded every two seconds until the last surgical suture was sited."*

10. **Expert reviewer:** Page 7 line 48: what do the authors mean by '' Regional cerebral oxygen saturation was collected''? ''collecting'' rSO2 is not correct English.

    Authors' response: We have now replaced the word collected with recorded. Thank you for this excellent recommendation.

11. **Expert reviewer:** Page 8, line 12: use the CAM abbreviation.

    Authors' response: We have now used the CAM abbreviation. Thank you for the excellent suggestion.

12. **Expert reviewer:** Page 8, line 17: ''carers'' is not the correct term, ''caretakers'' is.

    Authors' response: Thank you for the excellent suggestion, which has been discussed (and debated) amongst our authors. According to the Cambridge English Dictionary, "carer" is defined as someone who takes care of a person who is young, old, or sick. The term "carer" appropriately conveys the intended meaning in this context, and we respectfully have left this term unchanged.

13. **Expert reviewer:** Page 8, line 38: an arterial blood gas samples at the end of surgery does not prove that a certain arterial pCO_2 was maintained throughout surgery.

    Authors' response: Thank you for this important comment. An arterial blood gas at the end of surgery certainly does not prove that $PaCO_2$ was maintained throughout surgery. In our study, patients had gone under general anaesthesia and were mechanically ventilated through an endotracheal tube. Ventilation was adjusted according to the initial arterial blood gas result to achieve the desired $PaCO_2$ target ranges. The $PaCO_2$-$EtCO_2$ gradient was then maintained throughout surgery. The arterial blood gas result at the end of surgery enabled us to confirm $PaCO_2$ group adherence at the time of which the blood sample was taken. It is an assumption that the $PaCO_2$-$EtCO_2$ gradient remained constant between the $PaCO_2$ readings. In the Methods section, we stated *"an arterial blood gas (ABG) was obtained to check $PaCO_2$, and ventilation was further adjusted accordingly to achieve the desired $PaCO_2$ target ranges. The $PaCO_2$-$EtCO_2$ gradient was then maintained throughout the surgery, with the assumption that the $PaCO_2$ would remain constant. Additional ABGs were sampled at the discretion of the anaesthetist if the gradient required re-evaluation, for example, requirements for an adjustment of the ventilation setting. Finally, at completion of surgery, an ABG was sampled to accurately document the $PaCO_2$ value, and to assess whether $PaCO_2$ was being maintained within target values."*

14. **Expert reviewer:** Page 8, line 50-51: use the beforementioned abbreviations for pCO_2, pO_2, etc.

Authors' response: Thank you for this comment. We have made the correction and used the before mentioned abbreviations. Thank you for these excellent suggestions.

15. **Expert reviewer:** Page 9, line 19: use the beforementioned abbreviation for Hb.

    Authors' response: We have made the correction and used the before mentioned abbreviation. Thank you for the excellent suggestion.

16. **Expert reviewer:** Page 9, line 42: the abbrevation ECG is used without previous clarification.

    Authors' response: Please accept our apologies for the oversight. We now state *"Routine monitoring for all participants included continuous electrocardiogram (ECG), pulse oximetry, temperature, bispectral index (BIS) monitoring, and neuromuscular monitoring."*

17. **Expert reviewer:** Page 9, line 45: '' BIS reading'' should be ''BIS readings''.

    Authors' response: Please accept our apologies for the oversight. We now state "*Adequate depth of anaesthesia was ensured by targeting BIS readings between 40 and 60."*

18. **Expert reviewer:** Page 10, line 12: first, the authors discuss a relative change of 15% to be relevant, which translates into 12% absolute change in case of a baseline reading of 80%. What do the authors mean by ''common'' standard deviation of 10%?

    Authors' response: We agree with the Reviewer that the word "common" is ambiguous. For clarification in the revised manuscript we have replaced the words "common standard deviation" with "the standard deviation in both groups". Thank you for pointing this out.

19. **Expert reviewer:** Page 10. The sentence '' For the primary outcome we compared the absolute difference between the TMH and TN groups in percentage change in rSO2 from baseline to completion of surgery using an unpaired, two-tailed t-test.'' is unnecessary since it is a repetition of previous statements on the statistical analyses.

    Authors' response: We agree with the Reviewer that deleting the sentence does not compromise the content and yet, it improves readability. We have therefore made the necessary change in the manuscript. Thank you.

20. **Expert reviewer:** Page 11. The sentence '' The study was designed to investigation...'' is not correct English.

    Authors' response: Please accept our apologies for the oversight. In the discussion section, we now state *"our study was not powered to investigate postoperative delirium".*

21. **Expert reviewer:** Page 11. The paragraph '' Patient and Public Involvement'' is unnecessary and does not add valuable information to the manuscript.

Authors' response: Respectfully, we have included the paragraph "Patient and Public Involvement" as part of the submission requirement under the innovative Patient and Public partnership adopted since 2014 by *The BMJ*. The submission guidelines can be found on the link below (https://bmjopen.bmj.com/pages/authors/#reporting_patient_and_public_involvement_in_research

VERSION 3 – REVIEW

REVIEWER	Nousjka Vranken Maastricht University Medical Center, Maastricht, the Netherlands
REVIEW RETURNED	24-Nov-2019

GENERAL COMMENTS	The authors conducted a very interesting study with regards to the relationship between intraoperative CO2 management and cerebral oximetry readings. In general, the manuscript is well written. I do have some comments and concerns, with regards to the contents as well as the form/structure of the manuscript. The clinical relevance is missing; why do the authors wish to investigate the effect of partial CO2 on rSO2? On another note, why would the authors strive to increase the measured value captured by a particular monitoring tool? What would be the potential benefit of deliberately increasing regional cerebral tissue oxygen saturation while its effect on clinical outcome remains unknown? I feel that some background information and justification with regards to the hypothetical relationship between postoperative delirium and intraoperative mild hypercapnia is lacking in the manuscript. Was rSO2 monitoring part of routine intraoperative monitoring in the study centre? It appears that no blinding of the rSO2 measurement was applied. Staff including the anesthesiologist is able to intervene in order to strive for a particular saturation threshold or range of saturation values. This is logically more plausible in case of cerebral oximetry being part of routine monitoring. Explanation of the difference between absolute and trend monitoring values is unnecessary. In the current study, however, the authors note that only the absolute measurement values were used. Thus, a description of absolute and relative measurement of rSO2 is unnecessary. Page 7 lines 42-43: although a reference is used, the way the authors formulated the sentence makes it look as if they conducted the experiments with regards to assessing the measurement error of rSO2. Please reformulate the sentence. Page 10 sample size calculation: did the authors base their decision for a 15% change to be clinically relevant on previous research? Please provide supporting literature. Page 10 lines 39-40: the authors state that the skewness and kurtosis was inspected. Do the authors mean to say ‘’visual inspection of the data distribution using histograms?’’ Page 11: the paragraph ‘’Patient and public involvement’’ does not add any valuable information to the manuscript. Results section: what appears to be missing information regarding the occurrence of postoperative complications (besides delirium). Page 19. The following sentence should be reformulated to increase the readability: ‘’ Eastwood et al. found that mild hypercapnia resulted in higher rSO2 values in post-cardiac arrest patients when rSO2 values at the end of the normocapnic period and the end of the hypercapnic period were compared.31’’

	Page 20, line 8: the sentence ‘’Normocapnia was also found to be superior in preserving cerebral autoregulation.34’’ looks like it is wrongly placed in the text, please relocate this sentence and incorporate it into the tekst accordingly. Page 20: the paragraph following from the sentence ‘’Whilst theoretical absolute and relative saturation thresholds requiring prompt interventions have not...’’ is a repitition of information. Page 20: the sentence ‘’ There has been conflicting evidence in the literature regarding the relationship between rSO2 and LOS or postoperative cognitive performance.’’ requires support with a literature reference.

REVIEWER	Lynda Cochrane Clinical Statistics Consultants UK
REVIEW RETURNED	17-Nov-2019

GENERAL COMMENTS	I remain concerned that between-group differences in some baseline characteristics (e.g. comorbidities) have not been included in robustness analyses.

VERSION 3 – AUTHOR RESPONSE

Reviewer 4: Dr Vranken’s queries

We thank expert Dr Vranken for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** The clinical relevance is missing; why do the authors wish to investigate the effect of partial CO_2 on rSO_2? On another note, why would the authors strive to increase the measured value captured by a particular monitoring tool? What would be the potential benefit of deliberately increasing regional cerebral tissue oxygen saturation while its effect on clinical outcome remains unknown?

Authors’ response: Thank you for raising this important question. We have stated in the beginning of the *Introduction* section that the effects of mild hypercapnia on rSO_2, particularly its potential beneficial or harmful effects as a therapeutic ventilation strategy, have not been fully examined. Therefore, to address this question, a prospective clinical trial is clinically justified. One of the most common limitations in cerebral oximetry monitoring has been the absence of an intervention protocol to treat a decrease in regional brain oxygenation. Therefore, the use of therapeutic hypercapnia is physiologically plausible. As stated in our *Introduction* section, in animal models, CO_2 is a well-known vasodilator improving cerebral blood flow and the neuroprotective mechanisms of mild hypercapnia have been postulated to be a result of increase in cerebral blood flow, enhancement of oxygen delivery, improvements in cerebral glucose utilisation and oxidative metabolism, and activation of ATP-sensitive potassium channels to maintain normal neuronal activity in the setting of ischemia. Therefore, the effects of mild hypercapnia on impacting rSO_2 are both physiologically and biologically plausible.

Further, the literature on this subject is conflicting. Recently, our group performed a pilot randomized controlled study of mild hypercapnia during cardiac surgery with cardiopulmonary bypass.[1] Our findings showed that compared to targeted normocapnia, a target arterial carbon dioxide partial pressure between 50 and 55 mmHg did not increase rSO_2 appreciably during CPB but increased pulmonary artery pressures before and after CPB. To our knowledge there are no studies addressing this question in patients undergoing non-cardiac surgery. Rogers et al. found no clinical benefit in optimizing NIRS directed management of rSO_2 in patients undergoing cardiopulmonary bypass.[2] On the other hand, as previously discussed, there is clear evidence that low rSO_2 is linked to cerebral ischemia and neurological complications.[3-5]

As the reviewer astutely pointed out, the effects of rSO_2 on clinical outcome have not been fully examined, and as a result, there were no deliberate actions to adjusting the regional cerebral tissue oxygen saturation. As stated under *Randomisation and blinding,* the attending anaesthetists had no rSO_2 target to titrate to.

2. **Expert Reviewer:** I feel that some background information and justification with regards to the hypothetical relationship between postoperative delirium and intraoperative mild hypercapnia is lacking in the manuscript.

Authors' response: As discussed in detail in our manuscript, we have clearly stated that delirium is a secondary outcome and our findings of a greater incidence of early postoperative delirium in the targeted normocapnia group need to be interpreted with caution as confounders of postoperative delirium were not controlled, our study was not powered to investigation postoperative delirium, and mental state was only assessed by CAM, once pre-operatively and once postoperatively. Accordingly, our findings for delirium should be viewed as hypothesis generating.

We chose delirium as a secondary end point as it is a well-defined complication of surgery, and clearly outlined in the Diagnostic and Statistical Manual of Mental Disorders fourth edition (DSM-IV–TR; www.dsmivtr.org/). There are key characteristics of delirium which include classic symptoms that patient may express in hypoactive, hyperactive, or mixed psychomotor behaviours and importantly, tests have been developed and validated for use in diagnosis and grading of delirium, and as outlined in our manuscript where we used the Confusion Assessment Method (CAM), a validated tool to measure delirium.

Previously, other Reviewers have stated that postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation and certainly may be temporally associated with surgery. As previously discussed, postoperative cognitive dysfunction (POCD) is difficult to define. While the diagnosis of delirium requires a detection of symptoms, the diagnosis of POCD requires pre-operative neuropsychological testing (baseline) and a determination that defines how much of a decline is called cognitive dysfunction. The spectrum of abilities referred to as cognition is diverse, including learning and memory, verbal abilities, perception, attention, executive functions, and abstract thinking. It is possible to have a decrement in one area without a deficit in another. Self-reporting of cognitive symptoms has been shown to correlate poorly with objective testing, so valid pre- and postoperative testing is essential to the diagnosis of POCD.[6] Unfortunately, we did not have the resources to accurately undertake the accurate assessment of POCD in our study.

Further, there has not been a standard methodology used in the multiple studies within the POCD literature.[7] Selection of neuropsychological test instruments, the amount of change considered to be significant, timing of testing, and inclusion and exclusion criteria have all varied.[8] Therefore, for the present study, whilst we acknowledge that postoperative neurocognitive complications could potentially occur following episodes of profound cerebral desaturation, given the logistical barriers to accurately measure POCD, we focussed on delirium as a more relevant secondary end point.

3. **Expert reviewer:** Was rSO2 monitoring part of routine intraoperative monitoring in the study centre? It appears that no blinding of the rSO2 measurement was applied. Staff including the anesthesiologist is able to intervene in order to strive for a particular saturation threshold or range of saturation values. This is logically more plausible in case of cerebral oximetry being part of routine monitoring. Explanation of the difference between absolute and trend monitoring values is unnecessary. In the current study, however, the authors note that only the absolute measurement values were used. Thus, a description of absolute and relative measurement of rSO2 is unnecessary.

Authors' response: Thank you for these excellent comments.

Whilst cerebral oximetry is readily available in our institution, rSO_2 monitoring is not included as routine intra-operative monitoring. During anaesthesia, cerebral oximetry values must not be interpreted in isolation; alterations in cerebral oximetry measurements must take into consideration all available clinical information and physiological state of the patient. It is crucial to take notice that in our study, the attending anaesthetists had no rSO_2 value to target. We acknowledge that blinding was not applied to the attending anaesthetists, and this is highlighted as one of the limitations of the study under the *Discussion* section.

One of the most common limitations in cerebral oximetry monitoring has been the absence of an intervention protocol to treat a decrease in regional brain oxygenation. Hence, the clinical justification of our trial. This was one of the motivating factors to design and execute this RCT where we specifically tested the hypothesis that targeted mild hypercapnia (defined as partial pressure of carbon dioxide in arterial blood between 45 and 55 mmHg), during elective major surgery would increase cerebral oxygen saturation compared to targeted normocapnia (defined as partial pressure of carbon dioxide between 35 and 40 mmHg).

In response to the description of absolute and tread measurement of rSO_2, we agree entirely with the reviewer that the description of trend measurement may not be directly related to the context of the manuscript. As a consequence, we have now modified the paragraph under *Measurement of rSO_2* to the following.

"The absolute oximetry value is defined as the rSO_2 value measured by the oximetry probe calibrated by a fixed ratio of arterial to venous blood. In our study, only the absolute oximetry data were extracted and analysed. The accuracy of the Masimo O_3^{TM} regional oximetry was investigated by Redford et al. previously, and the measurement error was reported to be approximately 4% when checked against reference blood samples taken from the radial artery and internal jugular bulb vein."

4. **Expert reviewer:** Page 7 lines 42-43: although a reference is used, the way the authors formulated the sentence makes it look as if they conducted the experiments with regards to assessing the measurement error of rSO2. Please reformulate the sentence.

Authors' response: Thank you for the excellent suggestion. We agree with the Reviewer that the sentence can be potentially misleading. The sentence has been reformatted to *"The accuracy of the Masimo O_3^{TM} regional oximetry was investigated by Redford et al. previously, and the measurement error was reported to be approximately 4% when checked against reference blood samples taken from the radial artery and internal jugular bulb vein."*

5. **Expert reviewer:** Page 10 sample size calculation: did the authors base their decision for a 15% change to be clinically relevant on previous research? Please provide supporting literature.

Authors' response: Thank you for this important question. There is unfortunately no universally accepted threshold to identify pathological cerebral saturation. The threshold for identifying

cerebral ischemia may be influenced by a number of patient-specific or technology-dependent variables.

Numerous studies have utilised threshold of approximately 20% decrease in rSO_2 as the threshold for abnormal rSO_2. Murphy et al. defined cerebral desaturation events as a greater than or equal to 20% decrease in rSO_2 values from baseline measures or an rSO_2 of less than or equal to 55%.[3] Denault et al. proposed a clinical algorithm with the use of NIRS intra-operatively to monitor for cerebral saturation level and defined abnormal rSO_2 as a 20% reduction from baseline or an absolute decrease below 50%.[9] Levy et al. compared changes in oxygen saturation with electroencephalographic evidence of cerebral ischaemia and estimated an ischaemic threshold of 47% in cerebral saturation.[10]

If we consider 65% as the baseline cerebral oxygen saturation, a 15% increase from baseline would approximate 75% and a 15% decrease from baseline would approximate 55%. If rSO_2 of 47% is the threshold for cerebral ischaemia, rSO_2 of 55% can be considered as a conservatively approach to define cerebral desaturation.

6. **Expert reviewer:** Page 10 lines 39-40: the authors state that the skewness and kurtosis was inspected. Do the authors mean to say ''visual inspection of the data distribution using histograms?''

Authors' response: Thank you for raising this important point. We used the Shapiro-Wilk test to determine normality. We have also inspected the numerical values of skewness and kurtosis of a distribution. Although there has not been a set of agreed cut off values for skewness and kurtosis to determine normality, inspection of these matrices is generally a more objective method than visual inspection of the distributions.[11,12]

7. **Expert reviewer:** Page 11: the paragraph ''Patient and public involvement'' does not add any valuable information to the manuscript.

Authors' response: Respectfully, we have included the paragraph "Patient and Public Involvement" as part of the submission requirement under the innovative Patient and Public partnership adopted since 2014 by *The BMJ*. The submission guidelines can be found on the link below (https://bmjopen.bmj.com/pages/authors/#reporting_patient_and_public_involvement_in_research). We agree with the BMJ Open Editorial policy that public and patient involvement in randomised controlled trial is paramount to improving the quality and safe of clinical research. Active patient and public involvement are also considered as best practice by ethical review boards.

As per the Editor's request, we have now reviewed this section and made the paragraph more concise. We now state *"Patients were involved in the study from the initial pre-admission consultation appointment where the rationale of the study, potential applications of the study outcomes, data privacy and management, and potential harmful effects were explained in detail. Patients were not directly involved in the development of the research question and outcome measures, and they were not involved in the design and conduct of the study. Potential burden of the intervention was not rated by the patients themselves; rather, potential harmful effects were monitored by the attending anaesthetist as part of routine clinical care. Study results and outcomes, once finalised, will be mailed out to study participants."*

8. **Expert reviewer:** Results section: what appears to be missing information regarding the occurrence of postoperative complications (besides delirium).

Authors' response: Thank you for raising this important point. There were no intra-operative complications in either group. Four patients in the normocapnic group and three patients in the hypercapnic group developed a postoperative complication. In the normocapnic group, complications included atrial fibrillation (n=1), rectal bleeding resolving without intervention (n=1), blood transfusion (n=1), blood transfusion with a transfusion reaction (n=1). Postoperative complications in the hypercapnic group included death due to complications of surgery (n=1), stoma cellulitis requiring antibiotic therapy (n=1), and wound infection requiring antibiotic therapy (n=1). Since postoperative complications are not part of primary or secondary outcome measures of the study, and they were unrelated to the intervention delivered intra-operatively, we have respectfully reported them as above.

9. **Expert reviewer:** Page 19. The following sentence should be reformulated to increase the readability: '' Eastwood et al. found that mild hypercapnia resulted in higher rSO_2 values in post-cardiac arrest patients when rSO_2 values at the end of the normocapnic period and the end of the hypercapnic period were compared.''

Authors' response: Thank you for this excellent suggestion. We now state *"Eastwood et al. compared rSO_2 values at the end of alternating hypercapnic and normocapnic periods in post-cardiac arrest patients in a double cross-over study and discovered that mild hypercapnia resulted in higher rSO_2."*

10. **Expert reviewer:** Page 20, line 8: the sentence ''Normocapnia was also found to be superior in preserving cerebral autoregulation.'' looks like it is wrongly placed in the text, please relocate this sentence and incorporate it into the text accordingly.

Authors' response: Thank you for pointing this out. We completely agree with Dr Vranken that the sentence was misplaced and subsequently it did not convey the intended meaning of the paragraph. The objective of the paragraph was to compare our study findings with the current literature. Our study focused on measuring rSO_2 rather than cerebral blood flow, and therefore, we have not collected important metrics to inform the effects of mild hypercapnia on cerebral blood flow or cerebral autoregulation. The relationship between hypercapnia or normocapnia on cerebral autoregulation in the setting of hemodilution (hematocrit lower than 28%) was extensively investigated by Dr Vranken's study group.[13] This important finding was articulated by the following statement in the same paragraph. "Numerous factors, for instance, cardiac output, haemoglobin affinity for oxygen, cerebral autoregulation, and the ratio of cerebral arterial to venous blood volume, affect rSO_2 in the setting of hypercapnia, but changes in $PaCO_2$ and CBF, in turn, have a direct influence on these factors." We have, therefore, incorporated this important reference into the above statement. Thank you for this excellent suggestion once again.

11. **Expert reviewer:** Page 20: the paragraph following from the sentence ''Whilst theoretical absolute and relative saturation thresholds requiring prompt interventions have not...'' is a repetition of information.

Authors' response: We agree with the reviewer that the sentence can be improved in order to convey the intended meaning and provide insightful information to the paragraph. In particular, we have now stated that in our study, no interventions were performed intra-operatively in response to changes in rSO_2. We have added this key piece of information to the sentence.

"In our study, the reduction in rSO_2 from the baseline was small in the majority of patients in the TN group, and the attending anaesthetists had no rSO_2 target to titrate to. As a result, no interventions were performed intra-operatively in response to changes in rSO_2."

12. **Expert reviewer:** Page 20: the sentence '' There has been conflicting evidence in the literature regarding the relationship between rSO_2 and LOS or postoperative cognitive performance'' requires support with a literature reference.

Authors' response: Thank you for the excellent suggestion. The topic statement "There has been conflicting evidence in the literature regarding the relationship between rSO_2 and LOS on postoperative cognitive performance" was intended to signpost readers to the remaining paragraph, which contained a summary of current evidence on the topic with literature reference to each cited study. We have, therefore, respectfully kept the statement unchanged but we welcome the idea of providing the citations immediately following the topic sentence rather than providing the citations after quoting study results from each study group.

Reviewer 5: Dr Cochrane's queries

We thank expert Dr Cochrane for her valuable time in reviewing our manuscript. We genuinely appreciate the expert comments provided.

1. **Expert reviewer:** I remain concerned that between-group differences in some baseline characteristics (e.g. comorbidities) have not been included in robustness analyses.

Authors' response: We assure Dr Cochrane that between group differences baseline characteristics do not influence the outcomes we have reported. As we have stated in our previous response, for the primary outcome, we compared the absolute difference between the targeted normocapnia and targeted mild hypercapnia groups in percentage change in rSO_2 from baseline to completion of surgery using an unpaired, two-tailed t-test. Furthermore, a more detailed longitudinal analysis of time-by-treatment interaction was also conducted using a random effect generalized least squares regression model. The nature of the effect of time on the outcome measure is not relevant for any outcome that is measured at a prespecified time point, as both groups are compared at the same timepoint on either a value of an outcome of interest or the change from the baseline.

We have already stated that there are no clinically meaningful differences in the baseline patient characteristics between the two groups, so there is little evidence to suspect that the randomization procedure did not work. This is why we provided the results of additional analysis in our previous response to the Reviewer but, as stated previously, we would not regard including this post hoc additional analysis in the manuscript as appropriate.

We present out argument logically:

1) First, we observed no major clinical differences between groups based on Table 1 – i.e. randomization worked,
2) No a priori known prognostic values – hence the original test is the t-test as it relies on randomized nature of the study and there is no a priori need to adjust for known prognostic covariates,
3) We undertook extra exploratory analyses for the primary outcome and reported in greater detail our longitudinal analyses,
4) On the exploratory analyses of the primary outcome, we used an ANCOVA model with the change as the dependent variable, group as a factor, and baseline value of cerebral saturations as covariate. The results confirm the original findings, i.e. the adjusted difference between the intervention and control groups is 23% (95%CI: 10%-36%; p=0.001) for the left and 20% (95%CI: 8%-31%; p=0.001) for the right hemisphere. Additionally, when age or preoperative hemoglobin were included as extra covariates in the respective models, the effects remain highly significant for both hemispheres with the magnitudes ranging from 15% to 23% for different models,
5) In addition, our findings from the longitudinal random-effect analysis are consistent with both the primary analysis presented in the earlier version of the manuscript and extra analysis outlines above in our response. The estimates of the mean difference between the groups adjusted for time and baseline cerebral saturations are: 19% (95%CI: 9%-29%, p<0.001) for left hemisphere and 19% (95%CI: 11%-27%; p<0.001) for the right hemisphere. We chose not to report these in

the manuscript in the context of the longitudinal analysis as in the presence of significant time-by-group interactions, these averaged values, although consistent with the primary outcome, could mask the fact that the extent of change over time differs by treatment group.

Thank you for taking the time to review and consider the above manuscript for publication in BMJ Open.

A/Prof Laurence Weinberg
(BSc, MBBCh,MRCP,DPCritCareEcho,FANZCA,MD)

Director, Department of Anesthesia, Austin Hospital
Associate Professor, Department of Surgery, University of Melbourne
Associate Professor, Perioperative Pain and Medicine Unit, Department of Surgery, University of Melbourne

References

1. Chan MJ, Lucchetta L, Cutuli S, et al. A Pilot Randomized Controlled Study of Mild Hypercapnia During Cardiac Surgery With Cardiopulmonary Bypass. *Journal of cardiothoracic and vascular anesthesia.* 2019;33(11):2968-2978.
2. Rogers CA, Stoica S, Ellis L, et al. Randomized trial of near-infrared spectroscopy for personalized optimization of cerebral tissue oxygenation during cardiac surgery. *British journal of anaesthesia.* 2017;119(3):384-393.
3. Murphy GS, Szokol JW, Marymont JH, et al. Cerebral oxygen desaturation events assessed by near-infrared spectroscopy during shoulder arthroscopy in the beach chair and lateral decubitus positions. *Anesthesia and analgesia.* 2010;111(2):496-505.
4. Slater JP, Guarino T, Stack J, et al. Cerebral oxygen desaturation predicts cognitive decline and longer hospital stay after cardiac surgery. *The Annals of thoracic surgery.* 2009;87(1):36-44; discussion 44-35.
5. Gale SD, Hopkins RO. Effects of hypoxia on the brain: neuroimaging and neuropsychological findings following carbon monoxide poisoning and obstructive sleep apnea. *Journal of the International Neuropsychological Society : JINS.* 2004;10(1):60-71.
6. Deiner S, Silverstein JH. Postoperative delirium and cognitive dysfunction. *BJA: British Journal of Anaesthesia.* 2009;103(suppl_1):i41-i46.
7. Newman S, Stygall J, Hirani S, Shaefi S, Maze M. Postoperative cognitive dysfunction after noncardiac surgery: a systematic review. *Anesthesiology.* 2007;106(3):572-590.
8. Rasmussen LS, Larsen K, Houx P, Skovgaard LT, Hanning CD, Moller JT. The assessment of postoperative cognitive function. *Acta anaesthesiologica Scandinavica.* 2001;45(3):275-289.
9. Denault A, Deschamps A, Murkin JM. A proposed algorithm for the intraoperative use of cerebral near-infrared spectroscopy. *Seminars in cardiothoracic and vascular anesthesia.* 2007;11(4):274-281.
10. Levy WJ, MD, Levin S, MD, Chance B, PhD. Near-infrared Measurement of Cerebral Oxygenation: Correlation with Electroencephalographic Ischemia during Ventricular Fibrillation. *Anesthesiology: The Journal of the American Society of Anesthesiologists.* 1995;83(4):738-746.
11. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod.* 2013;38(1):52-54.
12. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab.* 2012;10(2):486-489.
13. Severdija EE, Vranken NP, Simons AP, et al. Hemodilution Combined With Hypercapnia Impairs Cerebral Autoregulation During Normothermic Cardiopulmonary Bypass. *Journal of cardiothoracic and vascular anesthesia.* 2015;29(5):1194-1199.