PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Sedation strategy and ICU Delirium: A multi-center, population-based propensity score-matched cohort study
AUTHORS	Casault, Colin; Soo, Andrea; Lee, Chel; Couillard, Philippe; Niven, Daniel; Stelfox, Tom; Fiest, Kirsten

VERSION 1 – REVIEW

REVIEWER	Fuller, Brian Washington University in Saint Louis, Anesthesiology and Emergency Medicine
REVIEW RETURNED	19-Nov-2020

GENERAL COMMENTS	This is a very pertinent research question for both researchers and practicing clinicians. While some of the results are not novel and somewhat expected, the authors have taken a new look and approach to this topic. The fentanyl findings are especially thought provoking. It is an informative and interesting study that I enjoyed reviewing. Some specifics Methods Thank you for reporting adherence to STROBE, as it increases transparency and quality. Study setting and population Can you tell us how consecutively mechanically ventilated patients were identified, and have you used this methodology before. Just to provide us some assurance that you indeed are capturing consecutive patients of interest. Exposures and definitions It's probably fine to have only included fentanyl, propofol, and midazolam as the "sedation strategies" since they are the most common infusions used, but I'm wondering why dexmedetomidine was also not included. The definition of "dominant sedation strategy" is obviously left open for debate and should be addressed in the Limitations. If one strategy dominates for only 5 or 10 minutes, that is not very dominant, nor likely to induce a clinical signal. Did the authors consider a sensitivity-type analysis, whereby they tested different levels of "dominance' based on time differences? Statistical analysis

	Please provide some rationale/justification for why those particular variables were chosen for the propensity score. If the authors are adhering to STROBE, they haven't reported anything about power and sample size rationale.

REVIEWER	Reade, Michael The University of Queensland
REVIEW RETURNED	03-Feb-2021

GENERAL COMMENTS	The manuscript presents the results of a retrospective propensity matched cohort study in which 2837 mechanically ventilated patients were categorised into groups according to “dominant” sedation strategy, defined as “the sedative infusion, including midazolam, propofol or fentanyl, with the longest duration prior to first delirium assessment”. Only these three drugs were chosen, based on the results of a preliminary study that found few others were prescribed in these 4 ICUs. A small number of patients (n=57) were excluded, after other exclusions, because two or more sedatives were commenced simultaneously (i.e. “had more than 1 dominant infusion prior to the first ICDSC assessment). Propensity-matched cohorts were compared (propofol vs. midazolam, n=712; propofol vs. fentanyl, n=1,732). Matching was performed according to age, sex, reason for admission to ICU, Charlson comorbidity category (0, 1, 2+), admission APACHE II score, use of vasoactive medications and use of continuous renal replacement therapy. In comparison to propofol, midazolam was associated with more ICDSC-identified delirium (occurring at any point in the ICU stay) and a longer duration of mechanical ventilation. In comparison to propofol, fentanyl was associated with longer duration of mechanical ventilation, longer ICU and hospital LOS, and more delirium days. With appropriate circumspection for a retrospective observational design, the authors conclude their study “highlights the need for additional research to further evaluate potentially negative effects of fentanyl and midazolam based sedation strategies.” The manuscript is clearly written and the ultimate conclusion justified by the results. However, I have several questions and comments that, if addressed, I believe would improve the utility of the manuscript: Major points: 1. The definition of “dominant” sedative infusion is novel, and so would benefit from the provision of additional detail. The definition used was “the longest continuous duration of infusion prior to the first ICDSC assessment”. This increased the size of the fentanyl group from 158 (who received fentanyl only) to 1069, midazolam from 124 to 356, and propofol from 887 to 1412. What time window was permitted in the definition of “longest”? i.e. if fentanyl had been infused for 11hr50 min and propofol for 11hr 40 min, a difference of only 10 minutes, was this patient classified in the fentanyl group? Clinically this would make little intuitive sense. If this is the way the analysis has been performed, at the very least this should be explained in detail in the discussion amongst the limitations of the study.

	2. “Sedation” with fentanyl alone is an unusual strategy, notwithstanding the results of the 2010 Strom Lancet trial of no sedation. Indeed, as the investigators observed, only 158 patients were treated with fentanyl alone prior to their first ICDSC assessment (table 2). Table S2 shows these 152 patients treated with fentanyl alone had a higher ICU and 30 day mortality. The possible biological reasons for this association are explored in the discussion, but the possibility of residual confounding not accounted for in the propensity score matching is not. This should be listed. Indeed, it would be interesting to explore just how unusual these patients were. Table 1 lists characteristics of the 1069 patients with a ‘fentanyl dominant’ strategy, but the paper does not provide similar data for the 158 who received fentanyl only. Indeed the characteristics of these 7 ‘sensitivity analyses’ groups could be listed as another supplementary analysis table, with appropriate commentary on the likelihood of residual confounding. 3. Did all included patients undergo an ICDSC and RASS assessment every 12 hours after admission (page 5)? Were there no missing data? Or, perhaps, was there a variable time between admission and the first ICDSC assessment (and hence variable sedative exposure between groups), and a variable frequency of ICDSC assessments (and hence likelihood of identifying the primary outcome?). If there are no missing data, or if this data is seemingly missing at random between the groups, this would be reassuring – and worth stating in the manuscript – but if these factors are not true, this should also be noted as a source of potential bias. 4. Methods: death data is available for “at least 18 months of follow-up from the ICU admission date”. Why is only 12 month mortality presented? 5. Results: “There were 2,837 patients in the study cohort (Figure 1). For those receiving propofol dominant sedation, it was common to receive a single agent (62.8%). While with fentanyl (14.8%) and midazolam dominant sedation strategies (34.8%) single agent use was less common.” And later “propensity scores were calculated for 1,409 patients receiving propofol dominant strategies, 1,067 patients receiving fentanyl dominant sedation strategies and 356 patients receiving midazolam dominant sedation strategies”. 14.8% of 2837 is 420 patients with a fentanyl-dominant strategy, not 1067, and 34.6% of 2837 is 982, not 356. Perhaps there is something in the manner of expression in these sentences that has led to my misunderstanding? 6. ICU mortality was 6.7-11%. How was the competing risk of death and delirium (the primary outcome) addressed? Minor points: 1. Introduction: “awareness” is not part of the DSM-V definition of delirium. Disturbance of awareness is part of the definition. 2. “Prospective cohort and randomized controlled trials support shorter durations of mechanical ventilation, more rapid awakening, and equally effective sedation with propofol or dexmedetomidine compared to midazolam”. This seems to be stating that these studies support equal effectiveness of propofol, dexmedetomidine and midazolam, which I think is not the authors’ intent (and which is not true). 3. The paper’s ultimate conclusion appropriately avoids attribution of causation to this observational study. However, several statements in the discussion do not, and need to be revised. Specifically:

	- “Regardless, these results should advise clinicians to be cautious when selecting their sedation strategy” - “Sedation using multiple agents also increased delirium risk, duration of mechanical ventilation, ICU LOS and hospital LOS” - “The mechanism that benzodiazepines would increase mortality is unclear”

VERSION 1 – AUTHOR RESPONSE

Response to Reviewers:

Reviewer: 1

1. Question #1: Study setting and population

Reviewer: Can you tell us how consecutively mechanically ventilated patients were identified, and have you used this methodology before. Just to provide us some assurance that you indeed are capturing consecutive patients of interest.

The study group received the data for all consecutive ICU patients receiving mechanical ventilation admitted during the study period from an established database using an electronic abstraction method. This methodology was published in Brundin-Mather *et al.* which compared electronic vs manual abstraction of mechanical ventilation as a yes-no variable and duration with a kappa of 1.0 and intraclass correlation coefficient of 1.0 for duration.

Ref: Brundin-Mather et al. (2018) Secondary EMR data for quality improvement and research: A comparison of manual and electronic data collection from an integrated critical care electronic medical record system. J Crit Care. DOI: 10.1016/j.jcrc.2018.07.021.

2. Question #2: Exposures and definitions:

Reviewer: It's probably fine to have only included fentanyl, propofol, and midazolam as the "sedation strategies" since they are the most common infusions used, but I'm wondering why dexmedetomidine was also not included.

During initial evaluation phase, the study team considered additional sedation strategies. A screening survey was conducted which demonstrated small populations utilizing alternative sedation strategies including dexmedetomidine, ketamine, lorazepam or alternative narcotics such as morphine or hydromorphone. The decision to limit the cohort to the above mentioned strategies was made after propensity matching, the remaining cohorts for those sedation strategies were too small for statistical analysis.

3. Question#3: Dominant sedation strategy definition:

Reviewer: The definition of "dominant sedation strategy" is obviously left open for debate and should be addressed in the Limitations. If one strategy dominates for only 5 or 10 minutes, that is not very dominant, nor likely to induce a clinical signal. Did the authors consider a sensitivity-type analysis, whereby they tested different levels of "dominance' based on time differences?

Due to the way dominant sedation strategy was defined, the difference could have hypothetically been that small. We present below the median difference in duration between each "dominant" strategy and the duration of the other 2 strategies. The results below are restricted to those who received both relevant agents.

For example, among those with a fentanyl dominant strategy who also received propofol, the median difference between the duration of fentanyl and the duration of propofol was 6.0 hours (IQR 2.3-15.6).

Dominant strategy	Difference between dominant strategy duration and propofol	Difference between dominant strategy duration and	Difference between dominant strategy duration and fentanyl
	duration (hours), median (IQR)	midazolam duration (hours), median (IQR)	duration (hours), median (IQR)
Propofol	-	12.5 (5.1-33.0)	4.7 (1.3-14.2)
Midazolam	9.9 (2.7-28.3)	-	6.2 (2.2-15.6)
Fentanyl	6.0 (2.3-15.6)	11.0 (3.8-25.7)	-

To best answer this question, the study team conducted a sensitivity analysis based on those patients receiving a dominant sedation strategy for ≥6 hours over the other 2 strategies and their patient centered outcomes. Using this restriction, fentanyl dominant strategies were associated with a composite of delirium or ICU death, ICU mortality, hospital mortality, 30 day & 1 year admission mortality, ICU & Hospital LOS and duration of mechanical ventilation. We have included this sensitivity analysis in the manuscript under Table 4. For the convenience of the reviewer, we have also included it below. Finally, we included a statement about describing the potential for bias created using this definition in the limitations segment.

Table 4: Sensitivity analyses based on those on a single sedation strategy or those whose sedation strategy was dominant for ≥6 hours over the other 2 strategies.

Outcome	**Dominant Sedation Strategy**	**Propensity score-matched odds ratio, mean ratio or rateratio (95% CI)[1]**
Delirium ever	Propofol	1.00 (reference group)
	Fentanyl	1.29 (0.99-1.69)
	Midazolam	**1.64 (1.12-2.41)**
Delirium or ICU death	Propofol	1.00 (reference group)
	Fentanyl	**1.38 (1.05-1.81)**
	Midazolam	**1.75 (1.18-2.60)**
ICU Mortality	Propofol	1.00 (reference group)
	Fentanyl	**1.82 (1.18-2.84)**
	Midazolam	1.31 (0.73-2.39)
Hospital Mortality	Propofol	1.00 (reference group)

	Fentanyl	**1.69 (1.19-2.42)**
	Midazolam	1.50 (0.92-2.49)
Died within 30 days of ICU admission	Propofol	1.00 (reference group)
	Fentanyl	**1.84 (1.27-2.68)**
	Midazolam	1.14 (0.69-1.89)
Died within 1 year of ICU admission	Propofol	1.00 (reference group)
	Fentanyl	**1.38 (1.02-1.86)**
	Midazolam	1.16 (0.77-1.76)
Died within 1.5 years of ICU admission	Propofol	1.00 (reference group)
	Fentanyl	1.25 (0.94-1.66)
	Midazolam	1.25 (0.84-1.85)
ICU length of stay, mean ratio (95% CI)	Propofol	1.00 (reference group)
	Fentanyl	**1.23 (1.09-1.39)**
	Midazolam	1.01 (0.86-1.20)
Hospital length of stay, mean ratio (95%CI)	Propofol	1.00 (reference group)
	Fentanyl	**1.31 (1.13-1.51)**
	Midazolam	1.01 (0.83-1.22)
Duration of invasive ventilation, meanratio (95% CI)	Propofol	1.00 (reference group)
	Fentanyl	**1.35 (1.14-1.59)**
	Midazolam	1.17 (0.94-1.46)
Number of delirium days, rate ratio(95% CI)	Propofol	1.00 (reference group)
	Fentanyl	1.19 (0.99-1.43)
	Midazolam	1.11 (0.85-1.44)

[1]Data presented as odds ratios unless otherwise indicated

4. Question #4: Statistical analysis

Reviewer: Please provide some rationale/justification for why those particular variables were chosen for the propensity score.

The study team considered the following prior to matching including, all measured baseline covariates, all

baseline covariates that are associated with treatment assignment (eg. sedation strategy), all baseline covariates that affect the outcome (ie. potential confounders) or all covariates that affect both treatment assignment and outcome (ie. true confounders).

Therefore, age, reason for admission, admission APACHE II, the charlson comorbidity index, use of vasoactives and renal replacement represented covariates which affected the outcome variables and therefore were selected to ensure patient severity and medical issues did not confound the outcome of the treatment assignment.

Ref: An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research, 46:399–424, 2011

5. Question #5:

Reviewer: If the authors are adhering to STROBE, they haven't reported anything about power and sample size rationale.

There was no sample size or power calculation, but this was a population-based study and we included all possible mechanically ventilated patients during the study timeframe.

Reviewer: 2

1. *Question #1:*

Reviewer: The definition of "dominant" sedative infusion is novel, and so would benefit from the provision of additional detail. The definition used was "the longest continuous duration of infusion prior to the first ICDSC assessment". This increased the size of the fentanyl group from 158 (who received fentanyl only) to 1069, midazolam from 124 to 356, and propofol from 887 to 1412. What time window was permitted in the definition of "longest"? i.e. if fentanyl had been infused for 11hr50 min and propofol for 11hr 40 min, a difference of only 10 minutes, was this patient classified in the fentanyl group? Clinically this would make little intuitive sense. If this is the way the analysis has been performed, at the very least this should be explained in detail in the discussion amongst the limitations of the study.

The study team greatly appreciates the constructive feedback provided by the review committee. As the comment is similar to the first reviewers question #3, we will recommend they review our response above, the associated sensitivity analyses and the changes made to the limitations section of the manuscript.

2. Question #2:

Reviewer: "Sedation" with fentanyl alone is an unusual strategy, notwithstanding the results of the 2010 Strom Lancet trial of no sedation. Indeed, as the investigators observed, only 158 patients were treated with fentanyl alone prior to their first ICDSC assessment (table 2). Table S2 shows these 152 patients treated with fentanyl alone had a higher ICU and 30 day mortality. The possible biological reasons for this association are explored in the discussion, but the possibility of residual confounding not accounted for in the propensity score matching is not. This should be listed. Indeed, it would be interesting to explore just how unusual these patients were. Table 1 lists characteristics of the 1069 patients with a

'fentanyl dominant' strategy, but the paper does not provide similar data for the 158 who received

fentanyl only. Indeed the characteristics of these 7 'sensitivity analyses' groups could be listed as another

supplementary analysis table, with appropriate commentary on the likelihood of residual confounding.

The study team has prepared the baseline characteristics of the 7 groups and included it in the manuscript as supplementary table 3. Patients who were admitted with only fentanyl were more critically ill with a moderately older age, higher surgical admission rate and use of vasopressors and CRRT compared to the propofol only patients. That being said, the midazolam only patient cohort were very similar in all those parameters but did not have a statistically higher mortality based on table S2 suggesting the association is not driven by those specific variables. One area of difference in the fentanyl and midazolam only cohorts was the admission reason. fentanyl only patients were more likely to be surgical admissions. Based on data provided by Ball *et al.,* elderly surgical admissions in the ICU tend to have better outcomes compared to medical patients, therefore this difference in the cohort may be less likely to be a confounder to the mortality outcome observed.

Ref: Ball et al. (2017) Outcomes of elderly critically ill medical and surgical patients: a multicentre cohort study. Can J Anaesth. DOI: 10.1007/s12630-016-0798-4

Question #3:

Reviewer: Did all included patients undergo an ICDSC and RASS assessment every 12 hours after admission (page 5)? Were there no missing data? Or, perhaps, was there a variable time between admission and the first ICDSC assessment (and hence variable sedative exposure between groups), and a variable frequency of ICDSC assessments (and hence likelihood of identifying the primary outcome?). If there are no missing data, or if this data is seemingly missing at random between the groups, this would be reassuring – and worth stating in the manuscript – but if these factors are not true, this should also be noted as a source of potential bias.

We appreciate the reviewer's diligence to data integrity. Understandably, there are limitations to database analysis secondary to human error. There was a variable time between admission and first ICDSC assessment and therefore, hypothetically, a different sedative exposure time, however the median time to first ICDSC in hours were similar between the propofol(median time = 17.1hrs (IQR = 8.5-34.7)), midazolam((median time =17.6 hrs (IQR = 8.8-41.2)) and fentanyl (median time =16.5 hrs (8.8-35.4) dominant strategies.

Similarly, the median number of ICDSC assessments per ICU day was similar for propofol(1.4 (IQR =1.0- 1.8) , fentanyl(1.4 (IQR 1.0-1.8), midazolam(1.3 (IQR 1.0-1.7). Therefore, any missing data was largely atrandom between the three groups and are less likely to introduce bias.

Question #4:
Reviewer: Methods: death data is available for "at least 18 months of follow-up from the ICU admission

date". Why is only 12 month mortality presented?

During the study design, 30 day and 12 month mortality was chosen a priori as commonly used mortality time frames. At the request of the reviewer, we have added the 18 month mortality to Table S2, which demonstrated no statistical difference from 12 months.

Question #5:
Reviewer: Results: "There were 2,837 patients in the study cohort (Figure 1). For those receiving propofol dominant sedation, it was common to receive a single agent (62.8%). While with fentanyl (14.8%) and midazolam dominant sedation strategies (34.8%) single agent use was less common." And later

"propensity scores were calculated for 1,409 patients receiving propofol dominant strategies, 1,067 patients receiving fentanyl dominant sedation strategies and 356 patients receiving midazolam dominant sedation strategies". 14.8% of 2837 is 420 patients with a fentanyl-dominant strategy, not 1067, and 34.6% of 2837 is 982, not 356. Perhaps there is something in the manner of expression in these sentences that has led to my misunderstanding?

To clarify, among those who received a fentanyl dominant strategy, 14.8% of those patients received only the single agent of fentanyl and among those who received a midazolam dominant sedation strategy, 34.8% of those patients received only the single agent of midazolam. We appreciate the opportunity to clarify and have added more detail and revised one of the earlier sentences in the manuscript for clarity. We have attached the change below for your convenience.

"There were 2,837 patients in the study cohort (Figure 1), which included 1412 patients (49.8%) receivinga propofol dominant strategy, 356 patients (12.5%) receiving a midazolam dominant strategy and 1069 patients (37.7%) receiving a fentanyl dominant strategy. For those receiving propofol dominant sedation,it was common to receive a single agent of only propofol (62.8%). While among those receiving fentanyl and midazolam dominant sedation strategies, single agent use was less common with 14.8% and 34.8% only receiving a single agent, respectively."

Question #6:
Reviewer: ICU mortality was 6.7-11%. How was the competing risk of death and delirium (the primaryoutcome) addressed?

To address the competing risk of death and delirium, we have provided the reviewer with a compositeoutcome of both death and delirium based on a propensity matched dominant sedation strategy. Thisdemonstrated very similar outcomes to the delirium for each dominant sedation strategy. We have subsequently included this in the results. The change to the manuscript can be found below:

*"*To control for the effects of death on delirium rates, a sensitivity analysis of the matched cohort was performed using a composite outcome of delirium or death. A statistically significant association between the composite outcome of delirium and death with midazolam dominant(OR 1.53 (95% CI 1.10-2.12; p=0.011) and fentanyl dominant(OR 1.27 (95% CI 1.04-1.55; p=0.020) strategies was observed, however not for propofol dominant strategies. *"*

Question #7:
A number of minor changes were recommended below:
1. *Introduction: "awareness" is not part of the DSM-V definition of delirium. Disturbance of awarenessis part of the definition.*
2. *"Prospective cohort and randomized controlled trials support shorter durations of mechanical ventilation, more rapid awakening, and equally effective sedation with propofol or dexmedetomidinecompared to midazolam". This seems to be stating that these studies support equal effectiveness of propofol, dexmedetomidine and midazolam, which I think is not the authors' intent (and which is nottrue).*
3. *The paper's ultimate conclusion appropriately avoids attribution of causation to this observationalstudy. However, several statements in the discussion do not, and need to be revised. Specifically:*

- *"Regardless, these results should advise clinicians to be cautious when selecting their sedationstrategy"*
- *"Sedation using multiple agents also increased delirium risk, duration of mechanical ventilation, ICULOS and hospital LOS"*
- *"The mechanism that benzodiazepines would increase mortality is unclear"*

The study team has adjusted the manuscript to ensure these minor corrections have been resolved.

VERSION 2 – REVIEW

REVIEWER	Reade, Michael The University of Queensland
REVIEW RETURNED	16-Apr-2021

GENERAL COMMENTS	The authors have provided responses to the questions in my original review. Commenting upon these in turn:

	1. Question #1: Reviewer: The definition of “dominant” sedative infusion is novel, and so would benefit from the provision of additional detail. The definition used was “the longest continuous duration of infusion prior to the first ICDSC assessment”. This increased the size of the fentanyl group from 158 (who received fentanyl only) to 1069, midazolam from 124 to 356, and propofol from 887 to 1412. What time window was permitted in the definition of “longest”? i.e. if fentanyl had been infused for 11hr50 min and propofol for 11hr 40 min, a difference of only 10 minutes, was this patient classified in the fentanyl group? Clinically this would make little intuitive sense. If this is the way the analysis has been performed, at the very least this should be explained in detail in the discussion amongst the limitations of the study. The study team greatly appreciates the constructive feedback provided by the review committee. As the comment is similar to the first reviewers question #3, we will recommend they review our response above, the associated sensitivity analyses and the changes made to the limitations section of the manuscript. The response that was provided is: Due to the way dominant sedation strategy was defined, the difference could have hypothetically been that small. We present below the median difference in duration between each “dominant” strategy and the duration of the other 2 strategies. The results below are restricted to those who received both relevant agents. For example, among those with a fentanyl dominant strategy who also received propofol, the median difference between the duration of fentanyl and the duration of propofol was 6.0 hours (IQR 2.3-15.6). (TABLE INSERTED) To best answer this question, the study team conducted a sensitivity analysis based on those patients receiving a dominant sedation strategy for ≥6 hours over the other 2 strategies and their patient centered outcomes. Using this restriction, fentanyl dominant strategies were associated with a composite of delirium or ICU death, ICU mortality, hospital mortality, 30 day & 1 year admission mortality, ICU & Hospital LOS and duration of mechanical ventilation. We have included this sensitivity analysis in the manuscript under Table 4. For the convenience of the reviewer, we have also included it below. Finally, we included a statement about describing the potential for bias created using this definition in the limitations segment. (TABLE INSERTED) Comment on second review: This does go some way to answering the question I asked. However: - In the sensitivity analysis presented above and added to the manuscript, what numbers of patients contribute to the data presented in table 4? - Why was 6 hours chosen as the threshold for table 4? Did other thresholds produce different results?

	2. Question #2: Reviewer: “Sedation” with fentanyl alone is an unusual strategy, notwithstanding the results of the 2010 Strom Lancet trial of no sedation. Indeed, as the investigators observed, only 158 patients were treated with fentanyl alone prior to their first ICDSC assessment (table 2). Table S2 shows these 152 patients treated with fentanyl alone had a higher ICU and 30 day mortality. The possible biological reasons for this association are explored in the discussion, but the possibility of residual confounding not accounted for in the propensity score matching is not. This should be listed. Indeed, it would be interesting to explore just how unusual these patients were. Table 1 lists characteristics of the 1069 patients with a ‘fentanyl dominant’ strategy, but the paper does not provide similar data for the 158 who received fentanyl only. Indeed the characteristics of these 7 ‘sensitivity analyses’ groups could be listed as another supplementary analysis table, with appropriate commentary on the likelihood of residual confounding. The study team has prepared the baseline characteristics of the 7 groups and included it in the manuscript as supplementary table 3. Patients who were admitted with only fentanyl were more critically ill with a moderately older age, higher surgical admission rate and use of vasopressors and CRRT compared to the propofol only patients. That being said, the midazolam only patient cohort were very similar in all those parameters but did not have a statistically higher mortality based on table S2 suggesting the association is not driven by those specific variables. One area of difference in the fentanyl and midazolam only cohorts was the admission reason. fentanyl only patients were more likely to be surgical admissions. Based on data provided by Ball et al., elderly surgical admissions in the ICU tend to have better outcomes compared to medical patients, therefore this difference in the cohort may be less likely to be a confounder to the mortality outcome observed. Ref: Ball et al. (2017) Outcomes of elderly critically ill medical and surgical patients: a multicentre cohort study. Can J Anaesth. DOI: 10.1007/s12630-016-0798-4 Comment on second review: This does address the question. I think it would be worth commenting in the manuscript that fentanyl alone tended to be chosen more for patients with a higher severity of illness (whether they are surgical or medical is less relevant). First, this would be a reassuring observation, as it would demonstrate that analysis of the database has identified a feature that clinicians would recognise as part of their practice. Second and more importantly, on page 10, it would provide another possible explanation for the observed association between a fentanyl-dominant strategy and the negative outcomes observed. i.e. that it is not a detrimental effect of fentanyl that has been identified (as is hypothesised e.g. in immunomodulatory effects), but rather that the causative association is reversed i.e. simply that sicker patients tend to be given fentanyl. Question #3: Reviewer: Did all included patients undergo an ICDSC and RASS assessment every 12 hours after admission (page 5)? Were there no missing data? Or, perhaps, was there a variable time between

	admission and the first ICDSC assessment (and hence variable sedative exposure between groups), and a variable frequency of ICDSC assessments (and hence likelihood of identifying the primary outcome?). If there are no missing data, or if this data is seemingly missing at random between the groups, this would be reassuring – and worth stating in the manuscript – but if these factors are not true, this should also be noted as a source of potential bias. We appreciate the reviewer's diligence to data integrity. Understandably, there are limitations to database analysis secondary to human error. There was a variable time between admission and first ICDSC assessment and therefore, hypothetically, a different sedative exposure time, however the median time to first ICDSC in hours were similar between the propofol(median time = 17.1hrs (IQR = 8.5-34.7)), midazolam((median time =17.6 hrs (IQR = 8.8-41.2)) and fentanyl (median time =16.5 hrs (8.8-35.4) dominant strategies. Similarly, the median number of ICDSC assessments per ICU day was similar for propofol(1.4 (IQR =1.0-1.8) , fentanyl(1.4 (IQR 1.0-1.8), midazolam(1.3 (IQR 1.0-1.7). Therefore, any missing data was largely at random between the three groups and are less likely to introduce bias. Comment on second review: I note this information has been added to the manuscript. This point has been satisfactorily addressed. Question #4: Reviewer: Methods: death data is available for "at least 18 months of follow-up from the ICU admission date". Why is only 12 month mortality presented? During the study design, 30 day and 12 month mortality was chosen a priori as commonly used mortality time frames. At the request of the reviewer, we have added the 18 month mortality to Table S2, which demonstrated no statistical difference from 12 months. Comment on second review: I agree that 12-month mortality is the more conventional endpoint. I would not have requested this information had the methods section not stated that follow-up had been undertaken for 18 months. It seemed illogical to stated that data had been collected and then not used. I am pleased at the intent to add this data to table S2, but in the version of this table I have been provided to review, only "Died within 1 year of ICU admission" is presented, not 18 month mortality. Either the methods section (stating that 18 month follow up has been undertaken) or table S2 should be modified to include this data. Question #5: Reviewer: Results: "There were 2,837 patients in the study cohort (Figure 1). For those receiving propofol dominant sedation, it was common to receive a single agent (62.8%). While with fentanyl (14.8%) and midazolam dominant sedation strategies (34.8%) single agent use was less common." And later "propensity scores were calculated for 1,409 patients receiving propofol dominant strategies, 1,067 patients receiving fentanyl dominant sedation strategies and 356 patients receiving midazolam dominant sedation strategies". 14.8% of 2837 is 420 patients with a fentanyl-

	dominant strategy, not 1067, and 34.6% of 2837 is 982, not 356. Perhaps there is something in the manner of expression in these sentences that has led to my misunderstanding? To clarify, among those who received a fentanyl dominant strategy, 14.8% of those patients received only the single agent of fentanyl and among those who received a midazolam dominant sedation strategy, 34.8% of those patients received only the single agent of midazolam. We appreciate the opportunity to clarify and have added more detail and revised one of the earlier sentences in the manuscript for clarity. We have attached the change below for your convenience. "There were 2,837 patients in the study cohort (Figure 1), which included 1412 patients (49.8%) receiving a propofol dominant strategy, 356 patients (12.5%) receiving a midazolam dominant strategy and 1069 patients (37.7%) receiving a fentanyl dominant strategy. For those receiving propofol dominant sedation, it was common to receive a single agent of only propofol (62.8%). While among those receiving fentanyl and midazolam dominant sedation strategies, single agent use was less common with 14.8% and 34.8% only receiving a single agent, respectively." Comment on second review: The manuscript now reads clearly. This point has been satisfactorily addressed. Question #6: Reviewer: ICU mortality was 6.7-11%. How was the competing risk of death and delirium (the primary outcome) addressed? To address the competing risk of death and delirium, we have provided the reviewer with a composite outcome of both death and delirium based on a propensity matched dominant sedation strategy. This demonstrated very similar outcomes to the delirium for each dominant sedation strategy. We have subsequently included this in the results. The change to the manuscript can be found below: "To control for the effects of death on delirium rates, a sensitivity analysis of the matched cohort was performed using a composite outcome of delirium or death. A statistically significant association between the composite outcome of delirium and death with midazolam dominant(OR 1.53 (95% CI 1.10-2.12; p=0.011) and fentanyl dominant(OR 1.27 (95% CI 1.04-1.55; p=0.020) strategies was observed, however not for propofol dominant strategies. " Comment on second review: This point has been satisfactorily addressed. Question #7: A number of minor changes were recommended below: 1. Introduction: "awareness" is not part of the DSM-V definition of delirium. Disturbance of awareness is part of the definition. 2. "Prospective cohort and randomized controlled trials support shorter durations of mechanical ventilation, more rapid awakening, and equally effective sedation with propofol or dexmedetomidine compared to midazolam". This seems to be stating that these studies support equal effectiveness of propofol, dexmedetomidine and midazolam, which I think is not the authors' intent (and which is not true).

	3. The paper's ultimate conclusion appropriately avoids attribution of causation to this observational study. However, several statements in the discussion do not, and need to be revised. Specifically: - "Regardless, these results should advise clinicians to be cautious when selecting their sedation strategy" - "Sedation using multiple agents also increased delirium risk, duration of mechanical ventilation, ICU LOS and hospital LOS" - "The mechanism that benzodiazepines would increase mortality is unclear" The study team has adjusted the manuscript to ensure these minor corrections have been resolved. Comment on second review: These points have been satisfactorily addressed. Just one causation sentence has slipped through, in the abstract: "Midazolam dominant sedation strategies increased delirium risk and duration of mechanical ventilation".

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Prof. Michael Reade, The University of Queensland
Comments to the Author:
The authors have provided responses to the questions in my original review. Commenting upon these in turn:

Question #1: Comment on second review:

This does go some way to answering the question I asked. However:
- In the sensitivity analysis presented above and added to the manuscript, what numbers of patients contribute to the data presented in table 4?
- Why was 6 hours chosen as the threshold for table 4? Did other thresholds produce different results?

Author Response:

When conducting the sensitivity analysis, the study team reviewed the sample size of the cohort at time points greater than 2, 4, 6 and 8 hours. Below, the reviewer will find the raw cohort size and propensity match cohort sizes. At 6 hours, there was a significant drop in the cohort size which is further impacted by the process of propensity matching. Therefore, there was concern that changes in the statistical results would be driven by conducting analyses on small samples size. When restricting for cases where the sedation strategy was dominant for greater than 6 hours, the cohort was reduced by approximately 55% and 65% for fentanyl and midazolam cohorts respectively. This was further reduced by the propensity matching process as not all patients could be matched secondary to their baseline characteristics.

Table: Cohort size based on duration of Dominant Sedation strategy

Dominant sedation strategy	*Total number of patients*	*Dominant for ≥ 2 hours over other strategies*	*Dominant for ≥ 3 hours over other strategies*	*Dominant for ≥ 4 hours over other strategies*	*Dominant for ≥ 5 hours over other strategies*	*Dominant for ≥ 6 hours over other strategies*	*Dominant for ≥ 8 hours over other strategies**
Propofol	*1412*	*1260 (89.2%)*	*1229 (87.0%)*	*1191 (84.3%)*	*1173 (83.1%)*	*1152 (81.6%)*	*1101 (78.0%)*
Fentanyl	*1069*	*852 (79.7%)*	*769 (71.9%)*	*693 (64.8%)*	*632 (59.1%)*	*584 (54.7%)*	*516 (48.3%)*
Midazolam	*356*	*298 (83.7%)*	*274 (77.0%)*	*252 (70.8%)*	*243 (68.3%)*	*231 (64.9%)*	*211 (59.3%)*

Table: Number of patients based on duration of dominant sedation strategy after propensity score matching:

Propensity matched dominant sedation strategy cohorts for comparison	*Dominant for ≥ 2 hours over other strategies*	*Dominant for ≥ 4 hours over other strategies*	*Dominant for ≥ 6 hours over other strategies*	*Dominant for ≥ 8 hours over other strategies*
Propofol vs Midazolam	*298 per group*	*252 per group*	*231 per group*	*211 per group*
Propofol vs Fentanyl	*672 per group*	*547 per group*	*476 per group*	*421 per group*

At the 2,4 and 6 hour cut off time points, the results remained significant for the composite of delirium or death, delirium, ICU mortality, hospital mortality, ICU & Hospital LOS and duration of mechanical ventilation(See below). At dominant for 8 hours or more, the mortality association remained significant, however the delirium and composite delirium end point odds ratio crossed 1.0 and lost significance. This was felt to be likely impacted by the substantial reductions in sample size as the effect was consistent at 2, 4 and 6 hours. The authors have included the sample sizes in the final version of table 4 in the manuscript for the reader as well as a brief mention in the limitations.

Table 4: Sensitivity analyses based on those on a single sedation strategy or those whose sedation strategy was dominant for ≥8, ≥6, ≥4, and ≥2 hours over the other 2 strategies.

Outcome		*Propensity score-matched odds ratio, mean ratio or rate ratio (95% CI)[1]*

	Dominant Sedation Strategy	*Dominant for ≥8 hours over the other 2 strategies*	*Dominant for ≥6 hours over the other 2 strategies*	*Dominant for ≥4 hours over the other 2 strategies*	*Dominant for ≥2 hours over the other 2 strategies*
Delirium ever	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	*1.28 (0.96-1.71)*	*1.29 (0.99-1.69)*	***1.32 (1.03-1.69)***	*1.20 (0.96-1.50)*
	Midazolam	*1.34 (0.90-2.02)*	***1.64 (1.12-2.41)***	***1.48 (1.02-2.15)***	***1.56 (1.11-2.21)***
Delirium or ICU death	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	*1.31 (0.98-1.75)*	***1.38 (1.05-1.81)***	***1.38 (1.07-1.77)***	*1.23 (0.99-1.54)*
	Midazolam	*1.49 (0.99-2.26)*	***1.75 (1.18-2.60)***	***1.55 (1.06-2.29)***	***1.66 (1.16-2.37)***
ICU Mortality	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.71 (1.11-2.67)***	***1.82 (1.18-2.84)***	***1.61 (1.09-2.41)***	***1.50 (1.04-2.18)***
	Midazolam	*1.38 (0.73-2.65)*	*1.31 (0.73-2.39)*	*1.26 (0.73-2.21)*	*1.27 (0.76-2.15)*
Hospital Mortality	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.61 (1.12-2.32)***	***1.69 (1.19-2.42)***	***1.48 (1.07-2.06)***	***1.36 (1.00-1.84)***
	Midazolam	*1.55 (0.91-2.65)*	*1.50 (0.92-2.49)*	*1.39 (0.88-2.23)*	*1.50 (0.97-2.36)*
Died within 30 days of ICU admission	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.55 (1.07-2.26)***	***1.84 (1.27-2.68)***	***1.58 (1.13-2.21)***	***1.40 (1.03-1.91)***
	Midazolam	*1.00 (0.58-1.72)*	*1.14 (0.69-1.89)*	*1.06 (0.66-1.69)*	*1.05 (0.67-1.64)*

Died within 1 year of ICU admission	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	*1.35 (1.00-1.84)*	***1.38 (1.02-1.86)***	*1.24 (0.94-1.63)*	*1.20 (0.93-1.55)*
	Midazolam	*1.22 (0.79-1.89)*	*1.16 (0.77-1.76)*	*1.10 (0.75-1.63)*	*1.19 (0.83-1.73)*
Died within 1.5 years of ICU admission	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	*1.23 (0.91-1.65)*	*1.25 (0.94-1.66)*	*1.16 (0.89-1.51)*	*1.17 (0.92-1.50)*
	Midazolam	*1.31 (0.86-1.99)*	*1.25 (0.84-1.85)*	*1.01 (0.85-1.21)*	*1.23 (0.87-1.75)*
ICU length of stay, mean ratio (95% CI)	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.15 (1.01-1.31)***	***1.23 (1.09-1.39)***	***1.16 (1.03-1.29)***	***1.15 (1.04-1.27)***
	Midazolam	*1.05 (0.88-1.26)*	*1.01 (0.86-1.20)*	*1.03 (0.88-1.21)*	*1.00 (0.86-1.16)*
Hospital length of stay, mean ratio (95% CI)	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.26 (1.08-1.46)***	***1.31 (1.13-1.51)***	***1.27 (1.12-1.45)***	***1.25 (1.11-1.41)***
	Midazolam	*1.05 (0.85-1.28)*	*1.01 (0.83-1.22)*	*1.01 (0.85-1.21)*	*1.01 (0.84-1.20)*
Duration of invasive ventilation, mean ratio (95% CI)	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*
	Fentanyl	***1.23 (1.03-1.47)***	***1.35 (1.14-1.59)***	***1.22 (1.05-1.42)***	***1.22 (1.06-1.40)***
	Midazolam	***1.34 (1.06-1.70)***	*1.17 (0.94-1.46)*	***1.28 (1.04-1.58)***	*1.17 (0.96-1.42)*
Number of delirium days,	*Propofol*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*	*1.00 (reference group)*

rate ratio (95% CI)	*Fentanyl*	*1.10 (0.91-1.33)*	*1.19 (0.99-1.43)*	*1.03 (0.87-1.23)*	*1.15 (0.98-1.35)*
	Midazolam	*1.01 (0.77-1.32)*	*1.11 (0.85-1.44)*	*1.03 (0.81-1.32)*	*1.05 (0.84-1.30)*

[1]Data presented as odds ratios unless otherwise indicated

Question #2:

Comment on second review:
This does address the question. I think it would be worth commenting in the manuscript that fentanyl alone tended to be chosen more for patients with a higher severity of illness (whether they are surgical or medical is less relevant). First, this would be a reassuring observation, as it would demonstrate that analysis of the database has identified a feature that clinicians would recognise as part of their practice. Second and more importantly, on page 10, it would provide another possible explanation for the observed association between a fentanyl-dominant strategy and the negative outcomes observed. i.e. that it is not a detrimental effect of fentanyl that has been identified (as is hypothesised e.g. in immunomodulatory effects), but rather that the causative association is reversed i.e. simply that sicker patients tend to be given fentanyl.

Author Response:

In response to the reviewer, we have added a line in the discussion commenting on this feature in the data.

Question #3:

Comment on second review:
I agree that 12-month mortality is the more conventional endpoint. I would not have requested this information had the methods section not stated that follow-up had been undertaken for 18 months. It seemed illogical to stated that data had been collected and then not used. I am pleased at the intent to add this data to table S2, but in the version of this table I have been provided to review, only "Died within 1 year of ICU admission" is presented, not 18 month mortality. Either the methods section (stating that 18 month follow up has been undertaken) or table S2 should be modified to include this data.

Author Response:

We appreciate the reviewers assistance. We have provided the updated Table S2 containing the 18 month follow up data.

Question #4:
Comment on second review:
These points have been satisfactorily addressed. Just one causation sentence has slipped through, in

the abstract: “Midazolam dominant sedation strategies increased delirium risk and duration of mechanical ventilation”.

Author Response:

We have adjusted the abstract conclusion sentence to the following, “Midazolam dominant sedation strategies were also associated with increased delirium risk and duration of mechanical ventilation”.

VERSION 3 – REVIEW

REVIEWER	Reade, Michael The University of Queensland
REVIEW RETURNED	18-Jun-2021

GENERAL COMMENTS	The reviewer also provided a marked copy with additional comments. Please contact the publisher for full details.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 2
Prof. Michael Reade, The University of Queensland
Comments to the Author:
see attached file. My comments have all been satisfactorily addressed.

Response: No response is required as Dr. Reade has confirmed that his comments have been addressed. We appreciate his contributions to the manuscript.