

Peer Review File

Article Information: <https://dx.doi.org/10.21037/jss-22-50>

Reviewer A

Research regarding treatments that benefit the veteran population is very important. This article describes a retrospective database study to determine if veterans undergoing lumbar spine surgery are able to reduce opioid use beyond 90 days.

<There are several basic manuscript structuring items:>

Comment 1. Put the tables in the order in which they appear in the manuscript. Table 3 appears on line 170 before Table 2 on line 195.

Reply 1: Thank you for your comment. We have reordered the tables and renumbered them in the manuscript to have them in order of appearance.

Changes in text: Table 3 is renumbered to table 2, table 4 is renumbered to table 3, and table 2 is renumbered to table 4.

Comment 2. Similarly, put the citations in the order in which they appear, for example lines 207-211, Dunn et al. (#21) is mentioned before Warner et al. (#20). Check them all.

Reply 2: Thank you for your comment. We agree and have ensured that all citations are now in order of appearance.

Changes in text: Citations have been reordered as necessary. Due to the software used to insert citations, these changes were unable to be tracked in the document.

Comment 3. In the statistical analysis paragraph of the methods section (lines 127-141) it is easier for the reader to follow if you describe the statistical tests in the order that they appear in the results, for example start with how you compared groups for Table 1 and so on and so forth.

Reply 3: Thank you for your comment. We have reordered this paragraph to describe the statistical tests in order that they appear in the results.

Changes in text: The paragraph now reads “All analyses were conducted using R statistical software (R Foundation for Statistical Computing, Vienna, Austria) with a significance level of 0.05. To compare baseline demographics of the cohorts, chi square test was used for categorical variables and Anova was used for continuous variables. The Cochran-Armitage test was used to perform trend analysis of opioid usage following lumbar surgery. For each stratified group, the distribution of cumulative MME was assessed using the Shapiro-Wilk normality test, which confirmed non-normality. The paired Wilcoxon signed rank test was subsequently used to compare cumulative preoperative MME (between days -365 and 0) to cumulative postoperative MME (between days 91 and 455). Multivariable logistic regression controlling for which controlled the significant differences between the cohorts including age, gender, race, year, tobacco use, obesity, hypertension, diabetes mellitus, depression, substance use disorder, anxiety, fusion, and number of levels treated was used to determine associations between covariates and opioid usage onwards of postoperative day 90 which was chosen as the acute pain period (as defined by at least 1 fill between postoperative days 91 and 455). Mann Whitney U test was used to compare cumulative MME from years 2010-2015 to cumulative MME from years 2016-2020. Chi square test was used to compare 30-day complication rates between the opioid naïve, low preoperative opioid usage, and high preoperative opioid usage groups.” (Line 129-145)

Comment 4. It is important to provide definitions or be more precise in your description of items that could be interpreted in more than one way by the reader, for example:

- line 110, define obesity using BMI

- line 112, define claim (is this an insurance claim? what information do you get with this - filled prescriptions with name, quantity, and sig? does this cover both inpatient and outpatient opioids?)

- line 121, did you calculate the MME using the information obtained from the claim or some other method?

- line 118, "the most commonly prescribed" is ambiguous, please provide a list of codes and names of medications (can be put in the supplemental material section)

Reply 4: Thank you for your comment. We have defined these below and added to the text to make these points clearer where appropriate.

Obesity: Obesity is not defined in ICD codes using specific BMI numbers. It is rather defined categorically. For example, ICD-10-D-E66.0 defines obesity due to excess calories whereas ICD-10-D-E66.1 defines morbid (severe) obesity due to excess calories. So, this definition is not directly linked to BMI.

Claim: Claim gives us all the information about the prescription that was filled. This includes type of medication, strength, and quantity for both inpatient and outpatient records.

MME: MME was calculated using information from the claim. MME was calculated using the following formula: Strength per Unit * (Number of Units/ Days Supply) * MME conversion factor = MME/Day.

Commonly prescribed: Information regarding the specific drugs included in the study was added to the appendix. The list of opioids included: Oxymorphone, tapentadol, buprenorphine, oliceridine, morphine, methadone, meperidine, hydromorphone, hydrocodone, fentanyl, codeine, oxycodone, butorphanol, dihydrocodeine, levorphanol, opium, pentazocine, tramadol

Changes in text: Appendix now contains a list of included opioids. Sentences added below:

“A claim is each instance that a patient filled a prescription for an opioid and the claim includes information on type of opioid, strength, and quantity for both inpatient and outpatient records.” (Line 115-117)

“The list of opioids included are included in **Appendix Table B**” (Line 123-124)

Comment 5. The description of obtaining comorbidities and complications starting at line 122 would flow better with the clinical characteristics above ending at line 110.

Reply 5: Thank you for your comment. We have moved this sentence from its original position to your suggested position after the sentence ending at line 110.

Changes in text: The sentence starting at line 122 now starts after the sentence ending at 110. This transition now reads “...other comorbidities listed in Table 1. Comorbidities and complications were defined using...” (Line 110-113)

Comment 6. It is not clearly stated what the primary and secondary aims of this study are. This is a necessity for the methods section. In the background of the abstract, starting at line 23, "The purpose of this study was to analyze the impact of lumbar spine surgery on postoperative opioid usage in the US veteran population." That statement needs to be broken down into specifics in methods:

- was the primary aim to determine the percentage of lumbar spine surgery patients who reduced opioids beyond 90 days?
- was the primary aim to to determine the percentage of lumbar spine surgery patients who discontinued opioids beyond 90 days?
- was the primary aim to determine what the odds are of opioid naïve patients to become dependent on opioids after surgery?
- was the primary aim to determine which factors are associated with continued opioid use beyond 90 days?

- was a secondary aim to determine if preoperative opioid usage was associated with postoperative 30-day complications?

Reply 6: We agree with your comment that we must clearly state our hypothesis and the authors of this manuscript feel that our current hypothesis does state our intention. Breaking the hypothesis down into specifics for 90 days, beyond 90 days, etc we feel is paramount for the extensive outcomes we have found, but are too granular for a hypothesis statement as we evaluated many metrics under this hypothesis and each of these comments are addressed directly in the methods.

Changes in text: None

<Accurately describing the cohort(s) is very important for the reader to be able to determine if they treat a similar population and could potentially apply the findings. Also, there must be agreement with the population you intended to study and the population you actually studied:>

Comment 7. Throughout the text, the terminology "lumbar spine surgery" or "spine surgery" is used. I do not have access to Appendix A.1., but as described in the paragraph starting at line 98, it seems that you selected a variety of CPT codes for lumbar decompressions. It is not explained if the other CPT codes were assessed to choose only patients having a decompression type surgery, or if your cohort includes patients that had a decompression as part of their fusion surgery. There is also no information on the number of levels treated (which may or may not have been accessible in the database, but should be, using the similar codes for single levels and additional levels). These surgical differences are important because there is a known difference between opioid utilization after smaller spine surgeries (example: microdiscectomy x 1 level) and larger spine surgeries (example: T11-pelvis fusion with laminectomy). Both of which would have a decompression code.

So if you excluded all decompressive codes that were combined with fusion codes to only include decompression-exclusive spine surgeries, then you should define your cohort as "lumbar decompression" or "lumbar decompressive spine surgery" patients throughout the manuscript including the title...or vice versa, so that this clear. **Please reference the supplementary online content eTable 1 in your reference #20 Warner et al as a good example of what the reader is looking for.**

Reply 7: Thank you for your comment and we agree that this was too ambiguous. The cohort included patients undergoing decompression alone as well as decompression with fusion so the cohort includes both. We have added a sentence in the methods to make sure this point is clear to the reader. We have also added the number of levels treated to Table 1. All codes used to define study populations are included in the supplementary appendix which has been added to the submission.

Changes in text: Sentence added to the methods reads: “Patients included in the cohort included patients who underwent decompression alone as well as patients who underwent decompression with fusion and all data was for inpatient procedures.” (Line 100-102)

Comment 8. In the introduction section, you define the veteran population as being high-risk and vulnerable in part due to the higher burden of mental health comorbidities, including: PTSD, depression, anxiety, and substance abuse disorder - which can make pain management challenging after spine surgery. Was this the intended population?

If yes, then why don't we see those same diagnoses in table 1? Only depression is included. Is this because this cohort of >15,000 patients did not have any these other diagnoses (if yes then need to include these as 0% in table 1). Or, was the database search ineffective at given you this information?

If no, then the cohort you describe does not appear to be more vulnerable or high-risk than the general population. Please compare your Table 1 to one of your references (#20) by Warner et al. - these mental health disorders are well defined and the percentage comparison of depression/tobacco alone to your cohort suggests your veteran population is at lower risk in comparison to the general population. Please rectify.

Reply 8: Thank you for your comment. We have added PTSD, anxiety, and substance use disorder to Table 1. The percentage with substance use disorder in our cohorts is higher than the average population.(1) A systematic review showed that the estimated lifetime prevalence of PTSD in civilians varies significantly from 3.4% to 26.9% and in military population varies from 7.7% to 17.0%.(2) As such the percentage with PTSD in our cohort is also in line with prior studies and higher than the general population. A national survey from 2013-2016 demonstrated that 8.1% of adults aged 20 and over had depression over a 2-week period.(3) In our cohorts the rate of depression ranged from 8.8%-11.1%. Since our cohorts demonstrated higher rates of substance use disorder, PTSD, and depression we feel the study is representative of the higher risk veteran population as described in the introduction.

Changes in text: PTSD, anxiety, and substance use disorder have been added to Table 1. An additional sentence describing significant differences between cohorts was added to the results. This sentence reads: “The three cohorts were significantly different in age, percentage undergoing fusion, number of levels with instrumentation, tobacco use, and substance use disorder.” (Line 153-155)

References:

1. Jones CM, Noonan RK, Compton WM. Prevalence and correlates of ever having a substance use problem and substance use recovery status among adults in the United States, 2018. Drug Alcohol Depend. 2020 Sep 1;214:108169.

2. Schein J, Houle C, Urganus A, Cloutier M, Patterson-Lomba O, Wang Y, et al. Prevalence of post-traumatic stress disorder in the United States: a systematic literature review. *Curr Med Res Opin.* 2021 Dec;37(12):2151–61.
3. Brody DJ, Pratt LA, Hughes JP. Prevalence of Depression Among Adults Aged 20 and Over: United States, 2013-2016. *NCHS Data Brief.* 2018 Feb;(303):1–8.

Comment 9. Table 1 often includes CCI or ASA, #levels treated, inpatient or outpatient, all to help the reader understand the overall cohort which can help predict risk of complications, etc.

Reply 9: Thank you for your comment. We have added CCI and number of levels with instrumentation to Table 1. The number of levels treated with decompression alone is not available in the database. All data was for inpatient procedures.

Changes in text: Additional rows containing number of instrumented levels and CCI were added to Table 1. An additional sentence describing significant differences between cohorts was added to results. This sentence reads: “The three cohorts were significantly different in age, percentage undergoing fusion, number of levels with instrumentation, tobacco use, and substance use disorder.” (Line 153-155)

Sentence added to make it clear all procedures were inpatient: “Patients included in the cohort included patients who underwent decompression alone as well as patients who underwent decompression with fusion and all data was for inpatient procedures.” (Line 100-102)

Comment 10. Lines 101-103 describe excluding readmissions. In your limitations section you mention that this artificially reduces the number of postoperative complications observed in the study population, but the bigger picture is missed. By excluding anyone with a postoperative complication that required readmission, you've excluded the patients that typically need more opioids and for longer duration. Again, not knowing if your study population includes only decompressions or also decompressions with fusions, the reader will be skeptical as to the ability of this paper to accurately report on opioid outcomes which effects validity and generalizability.

Reply 10: Thank you for your comment. We believe that adding patients who were readmitted in the analysis would skew the results as the opioid use may be due to other pathology related to the readmission and not necessarily tied to the initial procedure. We have added a sentence to our limitations to address your point.

Changes in text: Sentence added to limitations: “This also likely excluded patients who were more likely to need more opioids and for longer duration.” (Line 280-281)

<Statistical Analysis:>

Comment 11. What comparison test did you use for Table 1? Please include.

Reply 11: Thank you for your comment. We used Anova for continuous variables and chi square for categorical variables and have added this to our statistical analysis section. We have also added all p-values for the data in Table 1.

Changes in text: Added all p-values for the data in Table 1.

Comment 12. Lines 135-136, "Fisher's" not Fisher. What covariates did you examine? As best as I can tell, your results of Fisher's exact test are seen in lines 171-174, which are unadjusted results. If you aim to describe what factors are associated with continued postoperative opioid usage, I would suggest using a logistic regression model where the outcome of interest (dependent variable) is continued postoperative usage (yes/no) and the independent variables in the model (which need to be described in the text if you use this method) should include all of those that are known to contribute to opioid usage so that you can attempt to control for known/measured confounders (adjusted results).

Reply 12: Thank you for your comment and we agree. We have changed this analysis to a multivariable logistic regression which controlled for the significant differences between the groups that were noted in table 1. The variables controlled for in the regression included age, gender, race, year, tobacco use, obesity, hypertension, diabetes mellitus, depression, substance use disorder, anxiety, and number of levels treated. The new odds ratios have been included

Changes in text: “Multivariable logistic regression controlling for which controlled the significant differences between the cohorts including age, gender, race, year, tobacco use, obesity, hypertension, diabetes mellitus, depression, substance use disorder, anxiety, fusion, and number of levels treated was used to determine associations between covariates and opioid usage onwards of postoperative day 90 which was chosen as the acute pain period (as defined by at least 1 fill between postoperative days 91 and 455).” (Line 137-142)

“Preoperative opioid usage of more than 3 claims was the covariate most significantly associated with continued postoperative opioid usage (Odds Ratio 11.84, 95% CI 10.87 – 12.89). A history of depression (Odds Ratio 1.19, 95% CI 1.04 – 1.35), substance use disorder (Odds Ratio 1.20, 95% CI 1.05 – 1.38), and any surgical complication (Odds Ratio 1.20, 95% CI 1.02 – 1.42) were also associated with continued postoperative opioid usage.” (Line 175-180)

Comment 13. Lines 161-162: is the significant difference reported as $p < 0.001$ describing each timepoint or the change over time?

Reply 13: Thank you for your comment. This is describing rates at each time point.

Changes in text: "Postoperative opioid claims rates were significantly different among the three groups **at each timepoint** ($p < 0.001$). **(Figure 2)**" **(Line 166-167)**

Comment 14. In table 3:

the "All" row shows the median postop MME as 0 but the IQR = 0-900 (same problem in table 4). The median should be a number in the middle of 0-900. Same for row "1-3 prior claims."

The "opioid naive" row shows that no MME were take postoperatively - error? - the p-value wouldn't make sense if there were no difference in these values.

Round large values such as MME to the nearest tenth rather than hundredth or thousandth.

Reply 14: Thank you for your comment. We have rounded to the nearest tenth in all large values. As for the median MME, the median is defined as the middle value in a data set when the dataset is ordered from least to greatest. The median in this case is 0 because over 50% of the values in the postoperative set are 0. The interquartile range is defined as the middle 50% of a dataset. The postoperative IQR is 0-0 because the opioid naïve group had less than 25% of its patients with a postoperative MME of greater than 0. However, this number was non-zero as there were 10.0% of patients still on opioids postoperatively. As such, because the statistical test takes into account the entire dataset and not just the interquartile range, the p-value is not an error and is significant in this case.

Changes in text: Data in table 2 rounded to nearest tenth.

<Discussion:>

Comment 15. Lines 205-207: This study does not include pain levels or disability ratings, so there is no comparison able to be made to reference #19. The discussion section should be more of a literature comparison to similar studies rather than a place to site results from dissimilar studies.

Reply 15: Thank you for your comment and we agree. As such we have removed this sentence and reference.

Changes in text: Sentence removed: “Skolasky et al. demonstrated that patient reported pain and disability scores significantly improve in patients after lumbar spine surgery with over 70% of patients having improvement in pain intensity.(19)”

Comment 16. In the discussion and in table 4 you show a reduction in preoperative MME, lines 215-229, and discuss how several measures are in place to reduce opioid addiction. Do you think that these measures are part of why the results show a reduction of opioids after surgery (a confounder)? In order to control for that, you can build a time factor into the logistic regression model mentioned above (2010-2015 and 2016-2020).

Reply 16: Thank you for your comment and we agree. We have added in time (2010-2015 and 2016-2020) as a variable in to our multivariable logistic regression.

Changes in text: “Multivariable logistic regression controlling for which controlled the significant differences between the cohorts including age, gender, race, year, tobacco use, obesity, hypertension, diabetes mellitus, depression, substance use disorder, anxiety, fusion, and number of levels treated was used to determine associations between covariates and opioid usage onwards of postoperative day 90 which was chosen as the acute pain period (as defined by at least 1 fill between postoperative days 91 and 455).” (Line 137-142)

Comment 17. You briefly mention and show in Figure 3 that there is a decrease in lumbar surgeries in veterans over time (again need to define what lumbar surgery is exactly), but you don't return to that concept and explain how this could effect the results (or not).

Reply 17: Thank you for your comment. We have included the supplementary appendix with all CPT codes so the reader may see how the cohorts were defined. There is no bearing on the trends of lumbar surgery would effect our results as we only examined outcomes of patients who underwent surgery. Suggesting that surgeons have changed their indications or that patients have somehow changed in their own need for surgery is beyond the scope of this paper.

Changes in text: none

<Conclusion:>

Comment 18. Once you have your aims defined, you should go back to your results, discussion, and conclusion and order the text so that the primary aim is always first followed by the secondary aims.

Reply 18: We agree with your comment. However, as we stated in comment 6 the authors of this manuscript feel that our hypothesis does state our intention. Breaking the hypothesis down into specifics for 90 days, beyond 90 days, etc we feel is paramount for the extensive outcomes we have found, but are too granular for a hypothesis statement as we evaluated many metrics under this hypothesis and each of these comments are addressed directly in the methods. As such we have not redefined our aims and are unable to edit the order of the text.

Changes in text: None.

Reviewer B

The authors use the VINCI database to evaluate veterans opioid claims habits prior to and following lumbar decompression. The authors found that lumbar decompression was effective at reducing postoperative opioid prescription claims in patients who have low preoperative opioid claims and high preoperative opioid claims. Patients with high preoperative opioid claims had a greater reduction in compared to those with lower claims.

This is an interesting study on a highly relevant topic given the current opioid epidemic. The data appears reasonable and lines up relatively well with the current literature. Overall, the article is well written. I do have a couple minor comments that I believe should be addressed to improve the manuscript.

Comment 1. Although the authors have a table dedicated to 30-day complications, and opioid naive patients have significantly less complications than preoperative opioid users, this is not discussed. Why do the authors believe the 30-day complication rate is higher in opioid users. UTIs are significantly more likely, while wound infections and pneumonia are non-significantly greater. Do you believe this is due to confounding from demographic variables or the immunomodulatory effect of opioids? Please elaborate in the discussion.

Reply 1: Thank you for your comment. We agree that this should be elaborated on in the discussion and we have added the following to the discussion in order to explore the higher complication and UTI rate.

Changes in text: “Higher preoperative opioid usage was also associated with an increased risk of any complication and UTI. This association with preoperative opioid usage and postoperative

complications has been demonstrated in several studies of different orthopaedic procedures including lumbar spine surgery.(32–34) This increase in complications and UTI noted in this study and prior studies is possibly due to the immunomodulatory effect of opioids.(35)” (Line 242-247)

Added references:

32. Ridenour R, Kowalski C, Yadavalli A, Ba D, Liu G, Leslie D, et al. Preoperative Opioid Use Is Associated With Persistent Use, Readmission and Postoperative Complications After Arthroscopic Knee Surgery. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2021 May;37(5):1567–72.
33. Jain N, Phillips FM, Weaver T, Khan SN. Preoperative Chronic Opioid Therapy: A Risk Factor for Complications, Readmission, Continued Opioid Use and Increased Costs After One- and Two-Level Posterior Lumbar Fusion. *Spine*. 2018 Oct 1;43(19):1331–8.
34. Sing DC, Barry JJ, Cheah JW, Vail TP, Hansen EN. Long-Acting Opioid Use Independently Predicts Perioperative Complication in Total Joint Arthroplasty. *J Arthroplasty*. 2016 Sep;31(9 Suppl):170-174.e1.
35. Ninković J, Roy S. Role of the mu-opioid receptor in opioid modulation of immune function. *Amino Acids*. 2013 Jul;45(1):9–24.

Comment 2. I think adding an ANOVA analysis for age and chi squared tests for the other demographics would strengthen the fact that it appears there are no significant differences in the cohorts evaluated in table 1. If there are differences, a regression analysis may account for these findings to confirm that 30 day complications are different due to preoperative MME use, not baseline demographics.

Reply 2: Thank you for your comment and we agree. We have added the recommended statistical analyses to table 1 to explore if there are significant differences in the demographics of the cohorts. We also generated a logistic regression model to account for these findings and described them in the statistical analysis section.

Changes in text: Additional column with p-values added to table 1. Sentence added to methods: “Multivariable logistic regression controlling for which controlled the significant differences between the cohorts including age, gender, race, year, tobacco use, obesity, hypertension, diabetes mellitus, depression, substance use disorder, anxiety, fusion, and number of levels treated was used to determine associations between covariates and opioid usage onwards of postoperative day 90 which was chosen as the acute pain period (as defined by at least 1 fill between postoperative days 91 and 455).” (Line 137-142)

Comment 3. On page 7, line 214 of the discussion, I believe it should read "This also reinforces prior findings that lumbar spine surgery can be utilized to successfully reduce PREOPERATIVE opioid use in patients who meet criteria for a lumbar decompression"

Reply 3: Thank you for your comment. The paragraph preceding this sentence describes several studies which demonstrate a reduction in postoperative opioid usage in patients undergoing lumbar spine surgery. This study found that a similar reduction in postoperative opioid usage among our study population and as such we believe this sentence is correct in its usage of "postoperative" instead of "preoperative". We have added your suggestion of "in patients who meet criteria for a lumbar decompression" to the sentence.

Changes in text: "This also reinforces prior findings that lumbar spine surgery can be utilized to successfully reduce postoperative opioid use **in patients who meet criteria for lumbar decompression.**" (Line 216-217)