

**Reviewer: 1**

## Comments:

This is a timely study providing a new level of knowledge in an area that is lacking good quality data. The data provided in this article helps to clear out a lot of misconceptions in the area. It has real power to influence practice not only in the surgical field, but opioid prescribing in general. I congratulate the authors.

Overall the manuscript is well written. I have few specific comments.

## General comments

In general, what do you mean by the term “misuse” and “abuse”? These terms have specific definitions. These terms are interchangeably used throughout the manuscript. I am guessing that these terms indicate a diagnosis of dependence, abuse or overdose. I would recommend changing the terminology and being consistent in using only one term, as it is a bit misleading and confusing in the current form. Something like “Problematic use” or “or opioid related problems” may be better. As authors pointed out, most overdoses are among short term users, not those who abuse, and “opioid abuse” is also just a subjective diagnostic code.

This is an excellent point that is consistent with our findings. We have taken the opportunity to clarify our use of “misuse.” We define what we mean by misuse in the methodology as the composite of all identified dependence, abuse, and overdose. We have now replaced several uses of the word “abuse” with “misuse.”

## Introduction:

Para 2: “Surgeons play a large role in this epidemic by serving as a gateway to overprescribing medications that opioid diversion and epidemic”

This is a bit of a stretch. Surgeons prescribe a lot which can putatively increase opioid availability for diversion. But I don’t think there is evidence that this contributed to opioid epidemic. The reasons for the opioid epidemic is much more complex than the often-reiterated statement that “doctors overprescribing caused the current opioid epidemic.”

We changed this sentence and included a reference to say: “Overprescribing enables opioid diversion and increases the potential for addiction.”

Methods: Solid, sensitivity analyses excellent.

Results:

Page 10: Several times the terminology “CP100K” is used. Provide the full expansion of the term “cases per 100,000” at least once in the results section for ease of reading. The authors have provided the full term in methods. Or they can use “case/100,000”, CP100K is not intuitive.

Thank you for this suggestion. We now use “cases/100,000” throughout.

Page 12, line 25: “For non-chronic opioid users, higher doses of opioids had smaller effects on the rate of misuse than additional weeks of exposure.”

Please explain this. What do you mean by non-chronic or chronic user? Is this 90 days? I am guessing you mean 13 weeks.

Thank you for this comment, which highlights a critical point. We have clarified this sentence to read: *“For short-term opioid use <90 days, higher doses of opioids had smaller effects on the rate of misuse than additional weeks of exposure.”*

eFigure 2 A and/or B are too important to be left in the appendix. Combined with figure 1 B, it demonstrates the persistence/worsening of “opioid related problems” despite the pendulum swing of number of prescriptions from 2008-2014 and stable dosage.

We agree that eFigure 2 A and B help to tell an important story about the rate of post-surgical opioid fills, and have incorporated the addition of 2A into the main manuscript as figure 1C. They are further nuance to the main point (as mentioned above) that is shown in Table 1, where the fill rates by year are depicted alongside rates of misuse. We believe that eFigure 2 helps to show the further nuance that this trajectory remains even when you stratify by various categories.

Page 11, line 51: “For both refills and duration, risk of misuse increased sharply at shorter periods and began to taper at higher levels of exposure (11 weeks of duration).”  
Is there data point connected to this statement? If so, please cite.

We changed this sentence to read: “For both refills and duration, Figure 2A shows that the initial risk of misuse follows the trend-line and begins to taper at higher levels of exposures 11 weeks of duration.”

In additional risk factors of misuse, was a history of substance use other than tobacco available? This is an important factor. If not available or insignificant on analyses, please state that in results or discussion.

Other diagnosis codes for abuse were either among our exclusion criteria or were not prevalent enough in our cohort for inclusion in analysis. It is important to note, however, that our sensitivity analysis specifically targets the problem of missing confounders. In the sensitivity analysis, even

when we generated an unrealistically large confounder which was much more prevalent than any of the excluded codes, it was not sufficient to explain away the association that we observed. To make this point clearly, we added a sentence in the results: “An example of such a confounder could be an undiagnosed risk factor for post-surgical misuse, such as alcohol dependence.”

#### Discussion:

It would be comforting for surgeons and their patients to know that the rate of opioid related problems were small (0.2% in one year), but still significant because of the large numbers of surgeries done every year. So, I would encourage more forceful statement of that finding.

Thank you for this important point. We added a sentence in the first paragraph of the discussion to highlight the overall importance of these findings.

Page 14, second para: “Our data is consistent with several studies in surgical patients that have shown that early opioid administration after surgery is associated with subsequent long-term usage,(30, 31) a proxy for abuse.”

A proxy for abuse here is a bit harsh given that only 0.2% develop abuse. Besides administering opioids just after surgery is necessary in many, but not without risks.

We removed the phrase and re-organized the paragraph to emphasize the importance of the idea that refilling of opioids is strongly associated with ultimate misuse—as supported by several previous circumstantial findings in the literature.

Please discuss the persistence/worsening of “opioid related problems” despite the pendulum swing of the number of prescriptions from 2008-2014 and stable dosage.

We agree that this is an important point to clarify. The opioid epidemic is multi-factorial. Surgical opioid prescribing is only one cause of the overall increase in opioid misuse and abuse among all populations. We think it would have been more surprising, amid a worsening opioid crisis, to have observed stable rates of misuse. The overall rate of opioid misuse surely has many causes, so it would be surprising if it mirrored the potential causes that we measure in detail here.

In our discussion, we added the following sentences: Furthermore, our adjusted models suggest that the effect of duration is not explained by temporal changes in physician behavior or patient population. Overall rates of opioid misuse have continued to increase (Figure 1B); this epidemic is multi-factorial and is not only driven by duration of exposure on an aggregate level. During the later parts of the study period, surgeons appear to have reduced the number of patients receiving opioids and the number of patients with short prescriptions (<7 days). They increased rates of longer prescriptions for a subset of patients (see figure 2D). Despite changing exposure groups and behaviors over time, we specifically found that the effect of duration was

stable across the study years and was unchanged by the changing prescribing and misuse rates in the population. The stable relationship shown in Figure 3 is suggestive of an independent effect.

## Reviewer: 2

### Comments:

This is an interesting manuscript based on an Aetna database evaluating associations between post-operative opioid prescriptions and opioid misuse. They find a correlation between numbers of refills and duration of prescriptions in opioid naive post-surgical patients and the risk of opioid misuse. There are a number of minor issues in the manuscript including terribly difficult to read Tables and Figures (these should be trimmed in number if necessary for journal page length requirements but expanded in size those that remain). Methodology is also difficult to read at times (What is cabinet methodology?). But in general the conclusions are largely impartial, nonopiophobic and data driven (with the exception of the phrase at the end of page 14 stating the "long-term usage is a proxy for abuse" - is this true for other medications they may be on long term after surgery?). However, there are three primary concerns I have with this manuscript which may or may not be due to a need for clarification only:

We have attempted to clarify the term "cabinet methodology." We changed the sentence in the methods section to read: "This identified the "cabinet supply" of opioids acquired by a patient as outlined by Mosher."

1) Do these data sets include discharge meds? or are all of the data outpatient meds which follow whatever prescription was given on discharge. If these discharge prescriptions were based on usual requirements it certainly puts the risks of one refill in a different light.

All prescriptions on or after the date of discharge were included. This allowed us to identify index discharge prescribing habits as well as refill effects. We added the following sentence at in the results to clarify: "90% of the included patients filled an opioid prescription within 3 days of discharge."

2) The opioid misuse characteristics includes quite a grab bag of diagnoses and one wonders whether who labelled the patients as having opioid dependence disorder. A physician frustrated at a patient's failure to get better despite multiple treatments including surgery? What happens when a smaller subset of opioid misuse outcomes (overdose?) are evaluated? At the very least this table of misuse diagnoses needs to be a NON-supplemental Table to make evaluation of the data easier for the reader.

Numerous studies have been forced to use similarly coarse outcome measures, which is an effect of using claims data. We acknowledge this limitation in our discussion. We acknowledge that the quantification of opioid misuse is difficult with this relatively rare outcome. We performed a further sensitivity analysis to address this point as well as that made by another

reviewer. Specifically, we analyzed the effect of removing the most common outcome ICD-9 code, 304.00 (and its corresponding ICD-10 code, F11.20), a possible marker of chronic opioid use without misuse. Even after excluding this code, results are virtually unchanged. For example, each additional refill is associated with an increased risk of 70.7% using the original outcome. Using the modified outcome, the associated increase is 70.9%. We added the following sentences to the results: “To ensure that our outcome analysis was not biased by a specific ICD9 code, we removed 304.00 opioid dependence, the most common code, leaving only specific abuse and overdose codes. The findings of the model were virtually unchanged with this smaller subset: each additional refill was associated with an increased risk of 70.9% vs. 70.7% in the comprehensive model.”

3) Is it surprising to the authors given their conclusions that misuse numbers continued to rise despite the mean duration of exposure remaining stable (as described in the middle of page 12)? It is to me.

We agree that this is an important point to clarify. We think it would have been more surprising, amid a worsening opioid crisis, to have observed stable rates of misuse. The overall rate of opioid misuse surely has many causes, so it would be surprising if it mirrored just this one potential cause (duration) that we are measuring in detail here. The explanation of the rising rates of misuse is multi-factorial and are not only driven by duration of exposure on an aggregate level. However, we specifically found that the effect of duration was stable across the study years and was unchanged by the changing misuse rates in the population. We removed the confusing point from page 12 and added the following element to our sensitivity analysis to focus on the importance of the persistence of this effect despite changes in prescribing habits among clinicians over time. This was the important point we were hoping to make.

“Figure 3A shows that surgeons reduced the mean dosage within their specialty during the study period. Typical reductions ranged from 3 to 18 MME/day (4 to 24%) over the duration of the study. While dosage fell, mean duration of exposure during the years of the study remained relatively stable (Figure 3B).

Despite changing clinician behavior over time, the relationship between duration of exposure and misuse was persistent (see Figure 3C). Such stability is further evidence of the robustness of this effect.”

In short this is an interesting dive into a fertile dataset and if these confounds can be explained away or stated there are a number of interesting findings including the role of other pre-operative addictive disorders (tobacco and obesity) and pre-operative mood influencing comorbidities (chronic pain and depression) and the absence of risk for higher doses UNTIL several weeks into the post-operative period (where many epidemiologists might class them as chronic pain patients). I agree with the authors that for this reason this latter finding is certainly NOT in contradistinction to the chronic pain risk factor literature but an interesting supplement to it.

**Reviewer: 3**

## Comments:

Thank you for submitting this interesting article. I completely agree that there is a need to address acute care prescribing of opioids, and there are still a very limited number of studies addressing overdose and misuse after surgery. There are very limited data of misuse, abuse and overdose after acute care prescribing. The article is very well written, and the data collection and analyses (including multiple sensitivity analyses) were comprehensive.

Despite my enthusiasm for this work, I have some concerns and critiques that temper my enthusiasm.

## Major critiques/overall comments:

1. Definition of misuse- My major concern is about the number of patients meeting the "misuse" criteria based on a diagnosis of "dependence." I worry that many clinicians will use this diagnosis to represent uncomplicated continued opioid use, which is therefore much different than an opioid use disorder or overdose. "Chronic, uncomplicated opioid use" has an ICD9 code of 304.00, which is the same as "opioid dependence." These overlap with opioid use disorder ICD9 codes, and it is an inherent challenge of administrative data. The other diagnoses included are clearly misuse/abuse; however, the majority of the cohort (n=2203) have this more nebulous diagnostic code. It would not be surprising that those patients that continue to get refills are more likely to get this diagnosis, even if they are not misusing or abusing the opioid. In addition, I sometimes found the narrative unclear regarding misuse. For this to be considered a study of "misuse" the authors need to make a solid argument that these patients in fact are misusing or abusing.

Thank you for raising this issue. We analyzed a modified outcome after excluding 304.00 (and its corresponding ICD-10 code, F11.20). Results are virtually unchanged when using the modified outcome. For example, each additional refill is associated with an increased risk of 70.7% using the original outcome. Using the modified outcome, the associated increase is 70.9%.

We now describe this specifically in our sensitivity analysis by adding the following statements: "To ensure that our outcome analysis was not biased by a specific ICD9 code, we removed 304.00 opioid dependence, the most common code, leaving only specific abuse and overdose codes. The findings of the model were virtually unchanged with this smaller subset: each additional refill was associated with an increased risk of 70.9% vs. 70.7% in the comprehensive model."

2. Overdose- important to note that the ICD9 code for overdose may not be very sensitive, as this will not detect those who die, nor will it detect those who are resuscitated in the field

(EMS) and do not seek care to follow. There are groups working to create these types of data, but it is not easy. This is a limitation you cannot overcome.

Agreed. We mention this limitation specifically in the limitations section of our discussion.

3. Analysis- for the association between opioid use and the rate of misuse, does the relationship hold true if you simply analyze the initial prescription? Given the outcomes definitions, it seems obvious that refills would increase the rate of “opioid dependence,” if that includes chronic uncomplicated opioid use. Confused, as Fig2A calls it “initial exposure” but the prose states that you analyzed duration of exposure by week including refills.

Please see the response to question 1. We additionally analyzed just the first filled prescription within 3 days after discharge. Using this exposure, we find that each week of additional supply is associated with a 23.4% increase in risk of misuse. This finding is entirely consistent with our main analysis, as this analysis uses a cruder measurement of post-surgical opioid exposure and has a correspondingly smaller (but still sizable) effect estimate.

4. Days supplied vs dose- I am surprised that you do not see a more clear association with dose. While days supplied is a measure in claims data, it is not really how clinicians think when they prescribe. There may be some sense of an approximate time, but in the end, physicians prescribe a number of pills with less thought on the days this is intended to cover. We also found a weak association with dose when considering new chronic use of opioids.

While it may be that some surgeons think of the total number of pills and then define the number of days of prescription, there is an equal number of clinicians who define their prescription by a dose and a length and then tally the total number of pills. Additionally, this exposure is a strong nudge for both the patient and the clinicians caring for the patient.

Additional comments:

1. Abstract- Concept of “duration” of opioid prescription is confusing, as many will assume this means the first prescription; however, this is about refills

We have changed this to say “total duration” in order to ensure that readers are aware that we are measuring total exposure.

2. Abstract- should clearly describe that this is a composite measure for the primary outcome

Thank you for this comment. We have changed the sentence in the abstract to read the following: “The primary outcome was a composite outcome of *misuse* identified by a diagnostic code of opioid dependence, abuse, or overdose.”

3. P5- why not match the medical and pharmacy coverage? Seems that you would want a longer assessment of opioid use before and after surgery to match your medical outcomes.

We were interested in post-operative effects so we focused on maximizing the number of patients to be included in the sample. We followed a standard methodology used by several previous authors (for example, Jena et al. 2016) to identify opioid use and those with an abuse history.

4. P5 bottom- members were followed until they experienced an outcome or last month of coverage---why vary the coverage?

We model the time to event, so it is easy to accommodate different lengths of followup. However, we also looked at events within one year (which corresponds to only following everyone for a year) as part of the sensitivity analysis and did not find significant differences in our results. Otherwise, we tried to maximize the number of patients followed for the longest amount of time.

5. Methods- would be good to include a supplement of the CPT codes included to derive the cohort, as well as those used as covariates.

We reference the proprietary source of the CPT codes for inclusion in the methodology as specifically derived from the NSQIP institutional list. Given the proprietary nature of the CPT code list as well as its length, we are not able to list all codes used beyond providing a reference to the source document. To better explain the meta-categorization of surgeries by organ system, we added the following sentence to the methods: "Organ-based categories were derived from top-level CPT headers (e.g.10030-19499 for surgeries of the integumentary system)."

6. P6- 30-day postop for inclusion for postop opioid use is liberal. I would suggest a 3-7 day window to ensure that you can relate the script to surgical care—see Barnett NEJM 2017 Emergency Medicine

We additionally analyzed just the first filled prescription within 3 days after discharge. Using this exposure, we find that each week of additional supply is associated with a 23.4% increase in risk of misuse. This finding is entirely consistent with our main analysis, as this analysis uses a cruder measurement of post-surgical opioid exposure and has a correspondingly smaller (but still sizable) effect estimate.

7. Did you do a subanalysis of those using opioids for <7days before surgery—these people may not really be “naïve”---at minimum, my sense is that they may behave differently

It is difficult know when someone is truly naïve. We used a standard practice among previous authors to identify patients as naïve who had less than 7 days of documented pre-surgical use. Further, as Brummett 2017 pointed out, there is the possibility of pre-surgical prescribing for

post-surgical use in a small number of patients. Analysis of these effects are the subject of a future publication.

8. P7 given the primary outcomes, why truncate the refills to 5? Did you only exclude those with 350mg OME in the first prescription

We truncated the number of refills because there can be instability in the estimate of the rate of misuse when fewer patients receive that number of refills. We can get a much better estimate of the rate of misuse among those who receive more than 5 refills than we can among those who receive exactly 11 refills, for example. Combining in this way allows us to better estimate the risk in these groups and prevents any one group from exercising undue leverage over the analysis. In total, only 1% of patients received more than 5 refills in the time frame.

No patients were excluded from the study based on dosage. Patients with more than 350 MME/day over their entire post-surgical exposure were truncated to 350 for the purposes of analysis.

9. P7- why require that there not be a 30 day gap in prescribing? Getting a refill after a month could still be relevant to your outcomes---you are not trying to define chronicity but instead the effect of prescribing on dependence, abuse, overdose---so, all scripts seem relevant to me

We were interested in post-operative effects so we decided to focus on the period just after surgery. Of course a different analysis could look at all opioid fills, but that would change the scientific question.

10. Results- our group recently published an article that found no association between the initial prescription size and refill rate---seems relevant to this work and should be considered. Thank you for this important citation. We have added it to our introduction.

11. Table 1- why not consider age at time of surgery instead of birth year?

Thank you for this point. We changed our table to specifically look at age at time of surgery.

12. Figure 1- I would again prefer age at the time of surgery if you decide to keep Figure 2B in this manuscript. I wonder, however, if it is worth it to include Figure2B. There are some distinct findings, but the concept is the same.

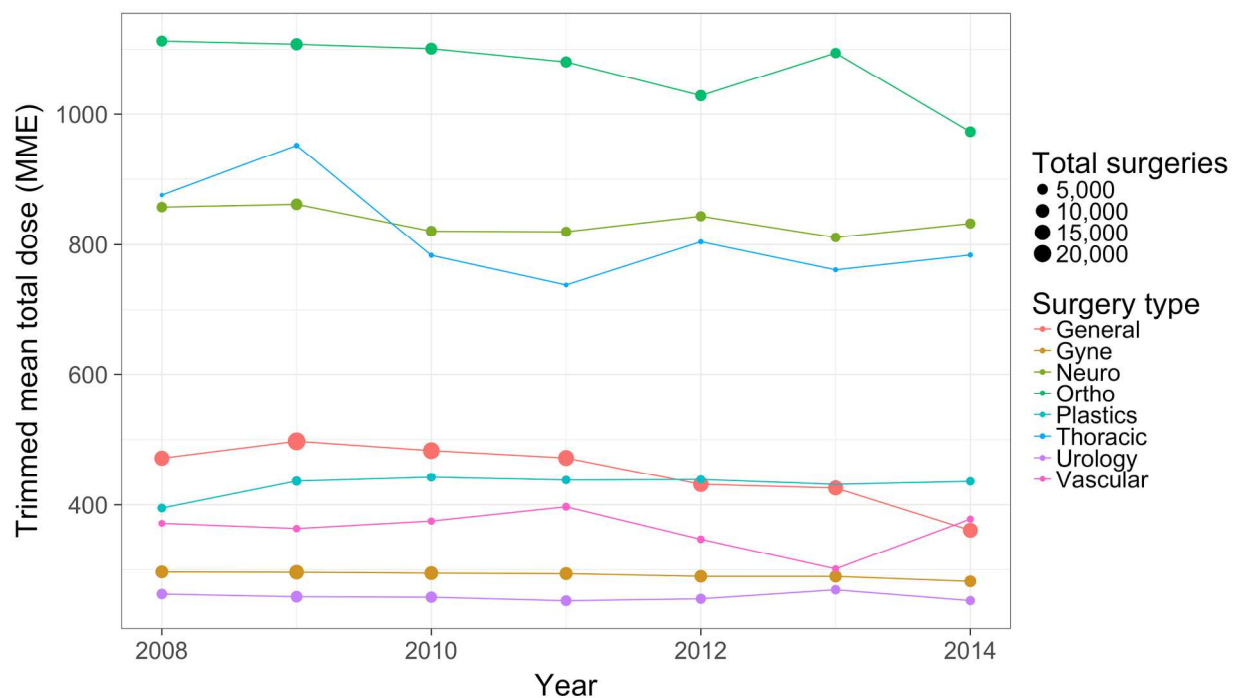
Again, thank you for this important point. We changed figure 1B to misuse vs year for the different age groups, as recommended. In compromise between the first reviewer's recommendation that we emphasize 1B and this recommendation that we consider removing the graph, we added prescribing over time by age cohorts.

13. P12, line25-27- I do not follow this comment. All were opioid naïve preop.

Line 25 refers to the mean post-operative dosage changes over time. We changed the sentence to emphasize “post-discharge prescription duration”

14. Figure 3a- why MME/day instead of MME? Total MME accounts for dose and duration and may be a better measure.

This is a very interesting point. It is possible to collapse both length of exposure and dose to total MME. However, we are interested in looking at both characteristics of the prescriptions. When we collapsed to create a total MME graph, it reflects our previous findings: dosage fell while duration was stable or increased. As a result, total MMEs prescribed did not change dramatically during the period of the study.



15. P14- second sentence of discussion- again, you describe “abuse” and “overdose” but some of this may have been what a physician deemed to be uncomplicated chronic use or dependence.

We have agree and have changed the wording of this sentence to “misuse” to be more inclusive. Additionally, we have performed a sensitivity analysis, mentioned above and in the main text, that suggests exclusion of the unclear variable of “dependence” does not change the association identified.

16. Discussion- I like that you refer back to the figures in the discussion