PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Does Cannabis Legalization Change Healthcare Utilization? A Population-based Study Using the Healthcare Cost and Utilization Project in Colorado, USA
AUTHORS	Delling, Francesca; Vittinghoff, Eric; Dewland, Thomas; Pletcher, Mark; Olgin, J; Nah, Gregory; Aschbacher, Kirstin; Fang, Christina; Lee, Emily; Fan, Shannon; Kazi, Dhruv; Marcus, Gregory

VERSION 1 - REVIEW

REVIEWER	Richard A Grucza Washington University School of Medicine, USA
REVIEW RETURNED	19-Dec-2018

GENERAL COMMENTS	I am very enthusiastic about this topic and the approach. Using the NAS report to develop hypotheses about which conditions might be impacted by cannabis legalization is a strength. But there are a number of weaknesses that will require significant effort to address. First is the choice of control states. New York state seems to be chosen as the primary control with no justification given—my guess is because of its size. Oklahoma is chosen as a second control because it is a neighboring state. Why not use multiple states as the control condition – For example all surrounding states, or all other states, or all states available in the whichever data sets are being analyzed? But there is a second issue with the controls which is that Colorado had legalized medical cannabis many year before 2012 and, in terms of the putative beneficial effects of cannabis legalization, people with Tourette's, PTSD, etc. had access well before 2012. Choosing states that already had medicalized cannabis might also make sense as a control strategy – but it is important that the authors consider the complex differences among medical legalization regimes (See Williams & Olfson's 2016 article in Health Affairs, for example). But the bottom line is that the prior medical legalization of cannabis in Colorado needs to be considered and addressed. Using two different controls make the presentation of results very confusing. For example Figure 3 presents *two* risk ratios for each outcome. There are no error bars or p-values, and so the reader has to refer to the legend to determine whether the two RRs were significantly different from each other and whether they were significant estimates in their own right. The authors never describe which ARHQ data sets they are using. Are they national or state?

	A summary of the NAS review in the introduction would be helpful. At least a partial list of what conditions are going to be analyzed and the different levels of evidence would help orient the reader to the eventual presentation of results. The choice to apply a Bonferroni correction is questionable. Presentation of 95% confidence intervals would be preferable. The "validation analysis" showing in increase in cannabis abuse diagnoses is interesting, but should be presented with the caveat that detection bias is a possibility. That is, patients may be more likely to disclose cannabis use, clinicians may be more likely to ask for / test for it after legalization. The authors need to clarify when they are using individual demographic variables from the AHRQ data and when they are using geographic-level variables. Also where do the "lifestyle" variables (alcohol use and smoking) come from? In sum, the consideration of outcomes and statistical modeling are very well thought out. The control strategy needs rethinking and the presentation needs improvement.

REVIEWER	David A. Gorelick, MD, PhD University of Maryland School of Medicine, USA
REVIEW RETURNED	26-Dec-2018

GENERAL COMMENTS	The authors have performed an extensive secondary analysis of data from the US Healthcare Cost and Utilization Project (HCUP) with the goal of evaluating whether the December 2012 legalization of recreational cannabis use in Colorado (CO) influenced inpatient healthcare utilization rates, both overall and for specific conditions (ICD-9 diagnoses) that might be influenced by cannabis use (either positively or negatively). To accomplish this, they compared utilization rates before (2010-2012) and after (2013-2014) legalization in Colorado with rates in the same time periods in two comparison states: New York (NY) and Oklahoma (OK). The question addressed is certainly of great public health and public significance. This would be the first study of which I am aware to evaluate the influence of recreational cannabis legalization (as opposed to medical cannabis legalization). However, I have serious concerns about how the authors conducted their analysis, making me doubt the validity of their findings. Most concerning is their use of 2013 as the dividing date for recreational cannabis becoming legal in CO. It is true that the referendum passed Nov. 6, 2012 and was certified as official on Dec. 10, 2012 (as stated in the manuscript). However, the law was not implemented (i.e., retail sales begun) until Jan., 2014. Thus, half the period treated as during legalization was actually before effective legalization. This blurring may account for the weak and inconsistent findings. The authors should redo their analyses using more appropriate time periods, e.g., 2011-2013 as pre-legalization and 2014-2015 as post-legalization. Another major concern is the authors' failure to use all the data available in the HCUP data set, which increases the possibility of confounding by unaccounted for factors. Potential confounding by variables associated with cannabis use (other than the identified "independent" variable, in this case, legalization) is clearly an issue in this type of population-level (epidemiological) analysis. This argues for adjusting for as many variables as possible in the statistical analyses (in other terms, using as many covariates as possible).

	1) The authors adjust only for age, race, and sex in their analyses, ignoring other relevant variables for which they have data, e.g., income level, “tobacco use,” and “alcohol abuse” (Table 1). In addition, the HCUP database records other relevant variables that the authors do not present, e.g., urban/rural location of hospital and presence/absence of psychiatric or alcohol abuse comorbidity. The authors should include all these variables in their analyses. 2) The authors use only two states as comparators for CO. They picked NY as the most populous state with data available for the years under consideration, and OK as a contrasting largely rural state adjoining CO. The differences in pairwise comparisons between CO vs. NY and CO vs. OK suggest that the findings are sensitive to unaccounted for differences in state-level characteristics (e.g., urbanicity). They also tend to confirm the authors’ concern that “secular trends…might differ across different populations” (p. 8) [or states]. All the more reason to compare CO with all other states which have data available for the relevant years, not just with 2 states. The larger sample size would hopefully average out any state-level differences in potentially confounding variables. 3) For similar reasons, the authors should include more than one index legalization state in their analyses. Why not also use Washington State (effective legalization July, 2014, similar to CO), Oregon (2015), and Alaska (2015). The inconsistent pattern of findings, as shown in Fig. 3, also raises my concern about validity. If I interpret the figure correctly, the majority of significant differences between CO and NY/OK are in the opposite direction from those the authors predict based on the health effects of cannabis summarized in the National Academy of Sciences (NAS) 2017 review. For only 3 of more than 2 dozen diagnoses did the authors find significant benefit (in terms of lower inpatient hospitalization rates) from recreational cannabis legalization, consistent with the NAS review. For several other diagnoses, the authors find significant benefit, while the NAS review suggests harm. The authors should discuss these discrepancies in terms of possible sources and its implications for their findings. Given that the NAS review is comprehensive and highly regarded in the field, these discrepancies cast doubt on the authors’ analyses more than on the validity of the NAS report. Other changes would improve the manuscript: 4) More information should be provided about use of the HCUP database. a. Which specific HCUP database was the source of the data analyzed? Was it the National Inpatient Sample (NIS)? b. How were diagnoses extracted from the HCUP database? Was only the primary diagnosis for each admission used? If so, this disregards some information, as multiple diagnoses may be coded for each inpatient admission: up to 30 in CO, 25 in NY, and 16 in OK. c. Several of the substance use disorder diagnoses listed in eTable 2 are “in remission” diagnoses, raising the question of whether they were incidental to the diagnosis actually responsible for the hospitalization. Were these “in remission” diagnoses included in the statistical analyses? 5) The authors should provide specific hypotheses regarding the association between recreational cannabis legalization and hospitalization rates, based on the NAS review. They can then use their statistical findings to confirm or dispute the specific hypotheses. This is the expected structure of a typical scientific

	paper. I suggest that the authors limit their hypotheses to putative cannabis health effects (either beneficial or harmful) for which the NAS review found either “conclusive,” “substantial,” or “moderate” evidence of an association. This greatly reduces the number of diagnoses being evaluated for influence by recreational cannabis legalization and lessens the possibility that failure to observe a significant influence was a false negative finding, i.e., due to absence of a significant association with cannabis, rather than true absence of an influence by legalization. I further suggest that the authors structure their discussion by the strength of evidence for the cannabis health effect. 6) The authors should be more precise in some of their terminology and chains of reasoning. a. The term “abuse” is used to refer to both “abuse” and “dependence” diagnoses (legend to Table 1). This will be confusing for readers and is contrary to the terminology of both DSM-IV and DSM-5. I suggest the authors use the term “substance use disorder,” which encompasses both abuse and dependence and is the term used in DSM-5. b. The manuscript frequently mentions “Amendment 64” in isolation. Although this term is defined in the Introduction (but first mentioned earlier in “Strengths and limitations”), it will not be readily familiar to readers outside of CO. More informative would be to mention what it stands for, i.e., “2012 legalization of recreational cannabis.” c. The authors are rightly concerned that secular trends might differ across states (p. 11; see item #3 above). However, a single pair-wise comparison between CO and another state does not alleviate this concern, as the concern is generated by the possibility that multiple states vary in secular trends. The manuscript contains evidence that this is the case: pairwise comparisons between CO-NY, CO-OK, and NY-OK are inconsistent for several diagnoses. The solution is to include as many states as possible in the analyses. d. The authors state that the “increased frequency of cannabis abuse diagnoses in CO helps validate…that legalization would result in greater use” (Discussion, p. 12). In fact, they observed an increased frequency of hospitalizations for cannabis abuse. This does not necessarily imply greater prevalence of cannabis use or even of abuse. The association could reflect increased severity of cannabis abuse, resulting in a higher rate of hospitalization for treatment (rather than, e.g., outpatient treatment). e. The manuscript frequently mentions “cannabis legalization” without clarifying that it is referring to legalization of recreational cannabis. CO already had legalized medical cannabis throughout the period of this study. 7) The authors should put their findings regarding hospitalizations for cannabis and alcohol abuse in the context of large, multi-state, population-based analyses of medical cannabis legalization. For example, some studies find legalization of medical cannabis associated with increased cannabis use but not increased cannabis use disorder (e.g., Williams et al., Addiction, 2017). Could legalization of recreational cannabis have different effects than legalization of medical cannabis? What would explain any such differences? 8) Literature cited should be updated in some areas. a. Reference #10 (O’Neill et al., 2017), cited to support cannabis treatment for PTSD symptoms, actually concludes that “Evidence is insufficient to draw conclusions about the benefits and harms of plant-based cannabis preparations in patients with PTSD.”

	b. Reference #12 (Li et al., 2012), cited for the association of cannabis with motor vehicle accidents, should be replaced with a more recent review, e.g., Rogeberg, Accid Anal Prevention, 2018; Hostiuc et al., Front Pharmacol, 2018. 9) The Abstract should include a brief description of statistical methods. 10) There are currently 31 (not 29) US state with legal medical cannabis (Introduction, p. 5, l. 15). 11) Some headings in Table 1 should be clarified. a. Title mentions "baseline" characteristics of the 3 study states, while the legend states that "All variables pertain to the 2010-2014 period," i.e., the entire study period, not just 2010. b. The heading "NAS Diagnoses" actually refers to the hospital admission diagnoses from HCUP. The table legend should explain that the admission diagnoses listed were based on those found influenced by cannabis in the 2017 NAS review. c. The NAS review should be explicitly cited in the table legend.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

I am very enthusiastic about this topic and the approach. Using the NAS report to develop hypotheses about which conditions might be impacted by cannabis legalization is a strength. But there are a number of weaknesses that will require significant effort to address.

We appreciate the Reviewer's positive feedback and the constructive suggestions, which we have addressed as follows to strengthen the paper:

1. First is the choice of control states. New York state seems to be chosen as the primary control with no justification given—my guess is because of its size. Oklahoma is chosen as a second control because it is a neighboring state. Why not use multiple states as the control condition – For example all surrounding states, or all other states, or all states available in the whichever data sets are being analyzed?

We understand the Reviewer's concerns. These were 2 of few states with HCUP data up until and including 2014. In addition, each state varies significantly in cost, but some cost tens of thousands of dollars (per state). We therefore opted to utilize two deliberately different states, that were both geographically and demographically different, with available data up until 2014. We believe selection of these two very different states is a strength of our manuscript rather than a weakness. Specifically, it allowed us to demonstrate that the effects of recreational cannabis legalization on health care utilization were independent of living in an urban or rural setting. We expanded on these concepts in the Discussion section, Strengths and Limitations (Page 17, Lines 340-344).

"As NY and OK are very different geographically and demographically, we believe their inclusion as control states represents a strength of our manuscript. Specifically, their selection allowed us to demonstrate that the effects of recreational cannabis legalization on health care utilization were likely independent of living in an urban or rural setting."

We also explain why we did not use all US states as controls in response to both Reviewer #1 and Reviewer #2 in the Discussion section, Strengths and Limitations (Page 18, Lines 363-368).

"We did not select all US states as controls. Most states do not have HCUP data up until the end of 2014, each requires a separate HCUP application, and each carries significant costs. We acknowledge this represents a limitation as we could not fully account for secular trends that might

influence results. However, we have adjusted all analyses for urbanicity and other state-level differences in order to minimize the impact of secular trends."

Regarding the reasoning behind selecting specific states, we have clarified in the Methods section (Page 6, Lines 98-101):

"New York (NY), the most populous state with inpatient HCUP data available up to 2014 was selected, and, to counter this coastal and largely urban state, we also selected Oklahoma (OK), a predominately rural state directly adjacent to CO with HCUP hospitalization data up to 2014."

2. But there is a second issue with the controls which is that Colorado had legalized medical cannabis many years before 2012 and, in terms of the putative beneficial effects of cannabis legalization, people with Tourette's, PTSD, etc. had access well before 2012. Choosing states that already had medicalized cannabis might also make sense as a control strategy – but it is important that the authors consider the complex differences among medical legalization regimes (See Williams & Olfson's 2016 article in Health Affairs, for example). But the bottom line is that the prior medical legalization of cannabis in Colorado needs to be considered and addressed.

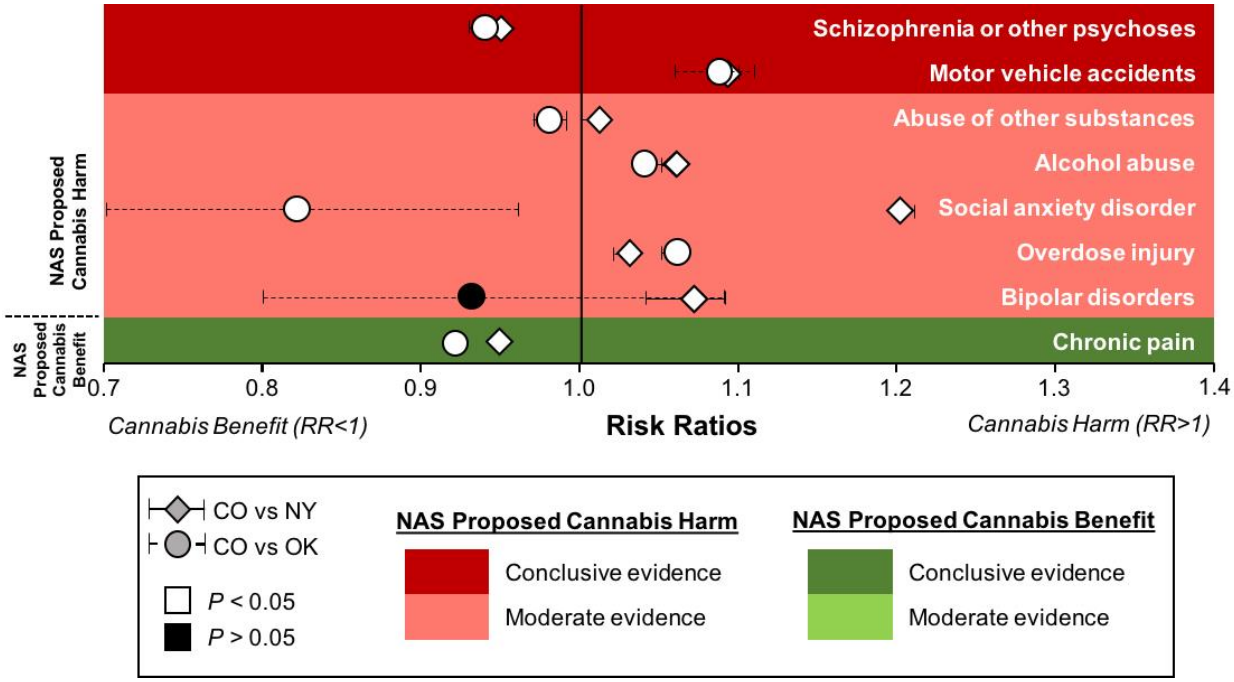
We appreciate this very insightful comment. Indeed Colorado had legalized medical cannabis many years before 2012 for individuals with Tourette's or PTSD or for nausea/vomiting in cancer patients. However, medical cannabis was not legal before 2012 in NY and OK (it became legal in 2016 in NY and 2018 in OK). As Reviewer #1 suggests, we could have selected control states that already had medicalized cannabis instead of NY and OK. However, based on the 2016 Health Affairs manuscript by Williams & Olfson referenced by reviewer, there are complex differences among medical legalization regimens among states. Hence, selecting control states based on a specific medical legalization regimen for a manuscript primarily focused on the effects of recreational (not medical) cannabis, is not ideal. Finally, we believe these effects would primarily reduce our power to detect a real difference rather than contribute to any false positive spurious results. We have addressed these points in the Discussion section, Strengths and Weaknesses (Page 18, Lines 368-375):

"We also acknowledge that CO had legalized medical cannabis prior to legalization of recreational cannabis on December 10, 2012. However, control states NY and OK had not. We could have selected control states that had already medicalized cannabis instead of NY and OK, but we acknowledge that complex differences exist among medical legalization regimens among states.[25] Hence, we believe selection of control states with diverse medical legalization regimens would have primarily reduced our power to detect a real difference rather than contribute to any false positive spurious results."

25. Williams AR, Olfson M, Kim JH, Martins SS, Kleber HD. Older, less regulated medical marijuana programs have much greater enrollment rates than newer 'medicalized' programs. Health Aff (Millwood). 2016;35(3):480-8.

3. Using two different controls make the presentation of results very confusing. For example Figure 3 presents *two* risk ratios for each outcome. There are no error bars or p-values, and so the reader has to refer to the legend to determine whether the two RRs were significantly different from each other and whether they were significant estimates in their own right.

We agree and have simplified the figure by reducing the number of diagnoses being evaluated for influence by recreational cannabis legalization. Specifically, we have only focused on substantial or moderate levels of evidence based on NAS review (as also suggested by Reviewer #2). A simplified figure 3 now allows inclusion of error bars.



4. The authors never describe which ARHQ data sets they are using. Are they national or state?

We are happy to better describe the datasets we employed. They are state databases and not national. We used the Healthcare Cost and Utilization Project (HCUP), one of the databases sponsored by ARHQ. We have added the following sentence to the Methods section (Page 6, Lines 93-96):

“HCUP is a state-wide database containing all-listed diagnoses and procedures, discharge status, patient demographics, and charges for all patients, regardless of payer (e.g., Medicare, Medicaid, private insurance, uninsured).”

And in the following paragraph (Page 6, Lines 101-104):

“We used separate HCUP databases for CO, NY and OK. Unlike the National Inpatient Sample (NIS), which is (by definition) a sample, these state-specific databases include data from every actual admission, providing direct and complete information regarding all healthcare utilization.”

5. A summary of the NAS review in the introduction would be helpful. At least a partial list of what conditions are going to be analyzed and the different levels of evidence would help orient the reader to the eventual presentation of results.

Thank you for pointing out the lack of an adequate summary of the NAS review on effects of cannabis. We have added the following sentences to the Introduction (Page 5, Lines 73-78):

“Recently, an extensive and rigorous summary of the current evidence on health effects of cannabis was developed by a committee of experts appointed by the US National Academy of Science (NAS) focusing on systematic reviews and high-quality primary research.[3] The health endpoints assessed in the NAS summary included oncologic, cardiometabolic, respiratory, immunologic, and psychiatric disorders as well as outcomes related to injury and death (i.e. motor-vehicle accidents).”

The following sentences were added to the Methods section (Page 8, Lines 141-145):

“For each health endpoint of interest in the NAS review, the weight of evidence regarding the statistical association of recreational cannabis with a specific health endpoint or the therapeutic use of cannabis had previously been categorized into substantial, moderate, limited or no evidence.[3] We opted to focus on NAS health endpoints with either substantial or moderate evidence.”

3. Committee on the Health Effects of Marijuana, Board on Population Health and Public Health Practice, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. The Health Effects of Cannabis and Cannabinoids : The Current State of Evidence and Recommendations for Research. Washington, DC: The National Academies Press; 2017.

6. The choice to apply a Bonferroni correction is questionable. Presentation of 95% confidence intervals would be preferable.

We agree that 95% confidence intervals are a useful addition that we are happy to include in the revision. Confidence intervals for each health outcome are now shown on a modified figure 3. Regarding the use of Bonferroni correction, extremely stringent standards were initially employed to address possible false positives due to multiple hypotheses testing. However, we acknowledge that we may have been excessively conservative. Hence, we are no longer using Bonferroni correction. This is made straightforward to justify given other revisions (such as ones mentioned above), where we have less comparisons (limited only to moderate or substantial NAS evidence of example). In the revised manuscript, we are using the following criteria to define a ‘positive’ comparison in the Methods section, which we think is more intuitive and more appropriate than our previous Bonferroni-corrected approach (Page 10, Line 192-198):

“In order to address possible false positives due to multiple hypotheses testing in assessing the diagnoses with moderate or substantial evidence described in the NAS document, we employed stringent standards: to be considered ‘positive,’ comparisons between CO and each of the other 2 states needed to be in the direction expected from the NAS report, each exhibiting statistical significance using a two-tailed alpha of 0.05. A ‘positive’ designation also required absence of a statistically significant difference between NY versus OK using a two-tailed alpha of 0.05.”

7. The “validation analysis” showing in increase in cannabis abuse diagnoses is interesting, but should be presented with the caveat that detection bias is a possibility. That is, patients may be more likely to disclose cannabis use, clinicians may be more likely to ask for / test for it after legalization.

This is a good point we agree should be acknowledged. We have added the following sentence to the Discussion section, Strengths and Weaknesses (Page 17, Lines 351-355):

“Patients may be more likely to disclose cannabis use following legalization of recreational cannabis and clinicians may be more inclined to test for it. Hence, we cannot completely exclude detection bias in the validation analysis showing an increase in cannabis abuse diagnosis following legalization of recreational cannabis”.

8. The authors need to clarify when they are using individual demographic variables from the AHRQ data and when they are using geographic-level variables. Also where do the “lifestyle” variables (alcohol use and smoking) come from?

We agree with the Reviewer that the distinction between individual demographic variables and geographic-level variables should be more clear. Initially, all models were adjusted only for race, age, and sex (variables obtained at the population level from census data). In HCUP, only basic demographic data (such as age and sex) and diagnoses based on ICD-9 code are available. Therefore, alcohol abuse is available from alcohol abuse codes and smoking is available from medical coding for smoking. We have included a list of the specific codes in eTable 2. Of note, we have employed these methods in several previous publications.[19-21, 24] We describe this specifically in the Methods section (Page 8, Lines 138-141):

“We used International Classification of Diseases-9th Edition (ICD-9) codes (eTable 2) to assess changes in healthcare utilization specific to various medical diagnoses potentially treated by or exacerbated by cannabis use as identified by the NAS summary.”

We also refer to the accuracy of these diagnoses from previous validation studies, but acknowledge reliance on this coding in the Limitations section of the manuscript (Page 17, Lines 345-348):

“The HCUP database relies on physician coding; however, such coding for several medical diagnoses listed in table 1 have been shown to be highly specific with variable sensitivity[17-20] and HCUP has proven to be a powerful tool in large population studies.[20-24]”

And in the following paragraph (Page 17, Lines 348-351):

“Some potentially important mediators are not captured by ICD-9 codes, such as quantity of cannabis used or formulation (oral versus other), although these may be more relevant to identifying mechanisms and their absence would likely not lead to false positive results.”

Based on the comment from Reviewer #1 as well as recommendations from Reviewer #2, we have added the following covariates to the models: 1) income level 2) proportion of HCUP individuals living in rural vs urban setting 3) with tobacco use, with 4) alcohol abuse, or 5) psychiatric disorders in each age, race and sex population strata. Description of these additional covariates and whether they were obtained at the individual vs population level is further clarified in the Methods (Page 7, Lines 115-121):

“In addition to the population-level variables age, sex and race, we obtained the following covariates at the individual level for each age, sex and race population strata: 1) income level 2) proportion of HCUP individuals living in a rural versus urban setting, 3) with tobacco use, 4) alcohol abuse, and 5) psychiatric disorders. The latter 3 covariates were available from ICD-9 codes. Income level was categorized by quartiles using the median household income for each patient’s ZIP code.”

17. Kim HM, Smith EG, Stano CM, et al. Validation of key behaviourally based mental health diagnoses in administrative data: suicide attempt, alcohol abuse, illicit drug abuse and tobacco use. BMC Health Serv Res. 2012;12:18.

18. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. PLoS One. 2014;9(3):e92286.

19. Dewland TA, Glidden DV, Marcus GM. Healthcare Utilization and Clinical Outcomes after Catheter Ablation of Atrial Flutter. PLoS ONE. 2014;9(7):e100509.

20. Dukes JW, Dewland TA, Vittinghoff E, et al. Access to alcohol and heart disease among patients in hospital: observational cohort study using differences in alcohol sales laws. BMJ. 2016;353:i2714.

21. Dewland TA, Olgin JE, Vittinghoff E, Marcus GM. Incident atrial fibrillation among Asians, Hispanics, blacks, and whites. Circulation. 2013;128(23):2470-7.

22. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. N Engl J Med. 2002;346(15):1128-37.

23. Gialdini G, Nearing K, Bhave PD, et al. Perioperative atrial fibrillation and the long-term risk of ischemic stroke. JAMA. 2014;312(6):616-22.

24. Whitman IR, Agarwal V, Nah G, et al. Alcohol Abuse and Cardiac Disease. J Am Coll Cardiol. 2017;69(1):13-24.

9. In sum, the consideration of outcomes and statistical modeling are very well thought out. The control strategy needs rethinking and the presentation needs improvement.

We appreciate the positive feedback and hope that our edits and revisions in the current manuscript are satisfactory. Again, we want to thank the reviewer for the insightful comments. We very much appreciate how they have helped to improve the manuscript.

Reviewer #2:

The authors have performed an extensive secondary analysis of data from the US Healthcare Cost and Utilization Project (HCUP) with the goal of evaluating whether the December 2012 legalization of recreational cannabis use in Colorado (CO) influenced inpatient healthcare utilization rates, both overall and for specific conditions (ICD-9 diagnoses) that might be influenced by cannabis use (either positively or negatively). To accomplish this, they compared utilization rates before (2010-2012) and after (2013-2014) legalization in Colorado with rates in the same time periods in two comparison states: New York (NY) and Oklahoma (OK). The question addressed is certainly of great public health and public significance. This would be the first study of which I am aware to evaluate the influence of recreational cannabis legalization (as opposed to medical cannabis legalization).

We appreciate the kind and supportive words.

1. However, I have serious concerns about how the authors conducted their analysis, making me doubt the validity of their findings. Most concerning is their use of 2013 as the dividing date for recreational cannabis becoming legal in CO. It is true that the referendum passed Nov. 6, 2012 and was certified as official on Dec. 10, 2012 (as stated in the manuscript). However, the law was not implemented (i.e., retail sales begun) until Jan., 2014. Thus, half the period treated as during legalization was actually before effective legalization. This blurring may account for the weak and inconsistent findings. The authors should redo their analyses using more appropriate time periods, e.g., 2011-2013 as pre-legalization and 2014-2015 as post-legalization.

We thank the reviewer for the insightful comment. We understand and appreciate this perspective regarding the optimal “dividing” date to select, but we deliberately chose 1/2013 as the date that recreational cannabis first became more readily available in CO. Specifically, after December 10, 2012 (when the “Colorado Amendment 64” became official), adults aged 21 or older could grow cannabis plants privately, legally possess all cannabis from these plants, and give cannabis as a gift to other adults aged 21 or older. Therefore, although legal retail sales in stores began in 2014, our working hypothesis is that recreational cannabis increased after 2012 when cannabis started to become available in private homes and in the streets. In order to validate that hypothesis, we performed a specific analysis highlighted in figure 1, demonstrating a jump in cannabis admissions prior to 2014 (which interestingly was not observed in either New York or Oklahoma at that same time).

However, we cannot deny that the additional effect related to availability in retail stores may yet be important. Hence, in response to this suggestion, we have now performed and include in our revision sensitivity analyses for all health care utilization (total number of admissions, length of stay, costs) and NAS-evidence based outcomes investigations utilizing January 1, 2014 as the change of policy date. Of note, the results did not meaningfully change in the sensitivity analyses for health care utilization, again highlighting the overall neutral effects of recreational cannabis legalization with regard to total admissions, length of stay and costs. Similarly, the main findings highlighted in the primary NAS outcomes analysis using stringent criteria (increase of alcohol abuse, motor vehicle accidents, overdose injury and decrease of chronic pain – further detail in response to comment #3), again met stringent criteria in all NAS outcomes sensitivity analyses except for overdose injury and

chronic pain (the latter meeting 2 of 3 criteria). Relative risks, confidence intervals and p values for all NAS sensitivity analyses with change of policy date to January 2014 are now included in eTable 3 in the manuscript supplement.

In order to address the distinction between these two important dates and highlight these additional analyses, we have added the following sentences to the Introduction section (Page 5, Lines 85-86):

"After January 1, 2014, recreational cannabis could be legally purchased in retail stores.[5]"

To the Methods section (Page 8, Lines 150-153):

"Finally, to take into account the additional effects related to cannabis availability in retail stores, we performed sensitivity analyses for all health care utilization (total number of admissions, length of stay, costs) and NAS-evidence based outcomes investigations utilizing January 1, 2014 as the change of policy date."

To the "Statistical Analysis" subsection of the Methods section (Page 10, Lines 184-186):

"The same model was used in the sensitivity analyses using January 1, 2014 (date of legal retail cannabis sales) as change of policy date."

To the "Number of Admissions" subsection of the Results section (Page 12, Lines 228-230):

"Results were similar in sensitivity analysis using January 1, 2014 as change of policy date (CO versus OK: RR 0.96; 95% CI 0.96 to 0.96; P<0.0005 and CO versus NY: RR 0.99; 95% CI 0.96 to 1.02; P=0.59)."

To the "Length of Stay" subsection of the Results section (Page 12, Lines 238-240):

"Length of stay did not change significantly in CO compared with the control states when we used January 1, 2014 as change of policy date."

To the "Healthcare Costs" subsection of the Results section (Page 12, Lines 246-249):

"No meaningful differences of healthcare costs were observed in the sensitivity analyses with January 1, 2014 as change of policy date or when we assumed the cost-to-charge ratio for OK was equal to the average of the other states."

To the "Changes in Specific Diagnoses Highlighted by the National Academy of Sciences" subsection in the Results section (Page 14, Lines 275-280):

"In the sensitivity analysis using January 1, 2014 as change of policy date (eTable 3), the main findings highlighted in the primary analysis using our stringent criteria (increase of alcohol abuse, motor vehicle accidents, overdose injury and decrease of chronic pain) again met the same criteria in favor of significant associations, except for overdose injury and chronic pain (the latter meeting 2 of the 3 criteria described in the Methods)."

And to the Discussion section (Page 16, Lines 326-333):

"In our primary analysis, we used December 10, 2012 as the date of recreational cannabis legalization. Following this date, private possession and growth of cannabis became legal in Colorado. After January 1, 2014, recreational cannabis could be legally purchased in retail stores.[5] To take the additional effect of retail sales of recreational cannabis into account, we performed sensitivity analyses using January 2014 as the 'change of policy' date. We did not observe any meaningful

difference in the majority of the outcomes studied, suggesting that the effect of recreational cannabis on healthcare utilization was independent of availability in stores."

5. Wark, J. Up Early and in Line for a Marijuana Milestone in Colorado. The New York Times [Internet]. 2014 Jan 1 [cited 2019 Feb 14];U.S.:[about 2 p.]. Available from: https://www.nytimes.com/2014/01/02/us/colorado-stores-throw-open-their-doors-to-pot-buyers.html

2. Another major concern is the authors' failure to use all the data available in the HCUP data set, which increases the possibility of confounding by unaccounted for factors. Potential confounding by variables associated with cannabis use (other than the identified "independent" variable, in this case, legalization) is clearly an issue in this type of population-level (epidemiological) analysis. This argues for adjusting for as many variables as possible in the statistical analyses (in other terms, using as many covariates as possible).

a) The authors adjust only for age, race, and sex in their analyses, ignoring other relevant variables for which they have data, e.g., income level, "tobacco use," and "alcohol abuse" (Table 1). In addition, the HCUP database records other relevant variables that the authors do not present, e.g., urban/rural location of hospital and presence/absence of psychiatric or alcohol abuse comorbidity. The authors should include all these variables in their analyses.

Thank you for this suggestion, which we believe has helped to strengthen our observations. We have added the following covariates to all multivariate models: 1) income level 2) proportion of HCUP individuals living in rural versus urban setting 3) tobacco use 4) alcohol abuse, and 5) psychiatric disorders in each age, race and sex population strata. Description of these additional covariates is further clarified in the Methods section (Page 7, Lines 115-122):

"In addition to the population-level variables age, sex and race, we obtained the following covariates at the individual level for each age, sex and race population strata: 1) income level 2) proportion of HCUP individuals living in a rural versus urban setting, 3) with tobacco use, 4) alcohol abuse, and 5) psychiatric disorders. The latter 3 covariates were available from ICD-9 codes. Income level was categorized by quartiles using the median household income for each patient's ZIP code. Income level was not available for OK."

After adjusting for the additional covariates, results of healthcare utilization (cost, length of stay, total admissions) remained neutral. However, results of our initial NAS outcomes analysis changed substantially when taking into account new covariates. Post-legalization changes most consistent with NAS-based evidence included an increase of alcohol abuse, motor vehicle accidents, overdose injury and a decrease of chronic pain. The latter 3 are all consistent with the NAS review and were not highlighted by the original analysis. We believe these findings may prove to be especially important and thank the reviewer for substantially improving our manuscript in this regard.

We have modified RR, CI and p values according to new covariates in the healthcare utilization results for the "Number of admissions" subsection (Page 11, Lines 224-228):

"After adjusting for covariates, there was a reduction of number of admissions following cannabis legalization in CO when compared to OK (RR 0.97; 95% CI 0.96 to 0.98; P<0.0005). The point estimate was similar when comparing CO to NY, but did not reach statistical significance (RR 0.99; 95% CI 0.98 to 1.01; P=0.47)."

For the "Length of stay" subsection (Page 12, Lines 234-238):

"After adjusting for covariates, length of inpatient stay did not change significantly in CO following the law change when compared with each of the 2 control states (1.75% annual reduction; 95% CI -12.25% to 10.01% in CO versus NY, and 3.46% annual reduction, 95% CI -16.48% to 37.90% versus OK; P= 0.30 and P=0.20, respectively)."

For the “Healthcare Costs” subsection (Page 12, Lines 224-246):

“In multivariate analyses, healthcare costs remained similar in state comparisons (2.99%; 95% CI -7.55% to 14.74%; P=0.18 versus NY and 3.45%; 95% CI -7.31% to 15.46%; P=0.16 versus OK) after cannabis legalization.”

We have modified the Discussion section accordingly (Page 14, Lines 292-297):

“Effects of recreational cannabis legalization on healthcare utilization appeared to be overall neutral. There was no evidence that either the length of stay or healthcare costs changed following liberalization of recreational cannabis. There was a reduction of overall hospitalizations in CO when compared to OK but not compared to NY, potentially because unaccounted state-level characteristics may have driven overall admissions differently.”

Further details about the NAS outcomes analysis using the new covariates and focusing on more compelling NAS evidence can be found in the response to comment #3.

b) The authors use only two states as comparators for CO. They picked NY as the most populous state with data available for the years under consideration, and OK as a contrasting largely rural state adjoining CO. The differences in pairwise comparisons between CO vs. NY and CO vs. OK suggest that the findings are sensitive to unaccounted for differences in state-level characteristics (e.g., urbanicity). They also tend to confirm the authors’ concern that “secular trends…might differ across different populations” (p. 8) [or states]. All the more reason to compare CO with all other states which have data available for the relevant years, not just with 2 states. The larger sample size would hopefully average out any state-level differences in potentially confounding variables.

We agree that in an ideal world we would be able to compare Colorado to every other state in the US. As the reviewer may be well aware, most states do not have data up until the end of 2014, each requires a separate HCUP application and processes, and each carries costs (some can be very expensive, as much as more than $10,000 per state). We were fortunate to have the means to purchase data from Colorado over this time period, and selected New York and Oklahoma as two of the few states with 2014 data available, given that we did/ do not have unlimited funds or resources to pursue and pay for 47 more applications. As the reviewer accurately points out, we did have a specific a priori rationale as well, selecting an urban geographically-disparate state and a largely rural state that was ‘next-door.’ We of course agree and acknowledge that there may be secular trends that might influence results. However, as we have attempted to communicate throughout the paper, those limitations would tend to reduce our sensitivity and would unlikely lead to spurious false positives. In other words, if something changed substantially in Colorado before and after legalization of cannabis and a similar pattern was not observed in either of these two very different states over that same time period, this would provide reasonable evidence that those changes may have been related to the legalization of cannabis. As described above, we now do indeed adjust for other state-level differences (such as urbanicity), which we believe has helped to strengthen the paper.

We have better acknowledged this limitation and its potential consequences in the revision by adding the following sentences to the Discussion section (Page 18, Lines 363-368):

“We did not select all US states as controls. Most states do not have HCUP data up until the end of 2014, each requires a separate HCUP application, and each carries significant costs. We acknowledge this represents a limitation as we could not fully account for secular trends that might influence results. However, we have adjusted all analyses for urbanicity and other state-level differences in order to minimize the impact of secular trends.”

c) For similar reasons, the authors should include more than one index legalization state in their analyses. Why not also use Washington State (effective legalization July, 2014, similar to CO), Oregon (2015), and Alaska (2015).

We agree that it would be interesting to further validate our findings, but those states do not yet have that data available (requiring at least up until the end of 2015 to provide a year of "post-legalization" data in Washington and until the end 2016 for Oregon and Alaska). And even if we had the funds to also complete the required applications from each state and pay for those states' data and/ or used them when they become available, we would be concerned about yet additional changes related to intervening secular trends and the need then for perhaps additional years within control states, which would likely lead to more confusion and make the paper particularly unwieldy. We are excited to be able to provide the first report of these changes in a state that has taken this step to legalize cannabis and hope the current paper can be a useful starting place for subsequent similar studies.

3. The inconsistent pattern of findings, as shown in Fig. 3, also raises my concern about validity. If I interpret the figure correctly, the majority of significant differences between CO and NY/OK are in the opposite direction from those the authors predict based on the health effects of cannabis summarized in the National Academy of Sciences (NAS) 2017 review. For only 3 of more than 2 dozen diagnoses did the authors find significant benefit (in terms of lower inpatient hospitalization rates) from recreational cannabis legalization, consistent with the NAS review. For several other diagnoses, the authors find significant benefit, while the NAS review suggests harm. The authors should discuss these discrepancies in terms of possible sources and its implications for their findings. Given that the NAS review is comprehensive and highly regarded in the field, these discrepancies cast doubt on the authors' analyses more than on the validity of the NAS report.

This is a good point we agree should be acknowledged. However, with several revisions suggested by both reviewers, we actually believe many of these inconsistencies have either been addressed or are less relevant. In the original figure 3 we included diagnoses with limited evidence of an effect of cannabis based on NAS review. This may explain why there were discrepancies between our findings and those of the NAS report. Following the recommendations of Reviewer #1 we have focused on NAS outcomes with moderate or substantial evidence of a cannabis effect. As suggested and as discussed above, we have also adjusted for other covariates that better capture state-level characteristics such as urbanicity, income level, tobacco and alcohol use. In doing so we have not only improved and simplified figure 3 but have also identified a higher and as such more compelling number of diagnoses where cannabis effect goes in the same direction of what was previously highlighted in the NAS reports. When maintaining strict criteria (1. comparison of NY vs OK $p > 0.05$, 2. both comparisons CO vs NY and CO vs OK with significant values and 3. RR with same direction of NAS review), we demonstrate that after legalization of recreational cannabis there is an increase of admissions for motor vehicle accidents, alcohol abuse, overdose injury and a decrease of chronic pain. When using less stringent criteria (2 of 3, one being NY vs OK > 0.05 and the other being that both CO comparisons with NY and OK are significant at $p < 0.05$ – even if only one state comparison goes in the same direction as NAS), admissions for abuse of other substances and social anxiety disorder also increase with cannabis legalization.

The following sentences were added to the Methods section (Page 10, Lines 192-197):

"In order to address possible false positives due to multiple hypotheses testing in assessing the diagnoses with moderate or substantial evidence described in the NAS document, we employed stringent standards: to be considered 'positive,' the comparisons between CO and each of the other 2 states needed to be in the direction expected from the NAS report, each exhibiting statistical significance using a two-tailed alpha of 0.05. A 'positive' designation also required absence of a statistically significant difference between NY versus OK using a two-tailed alpha of 0.05."

To the Results section (Page 13, Lines 258-272):

“In the NY versus OK P>0.05 group (figure 3), changes in rates of diagnoses after cannabis legalization reflected NAS-based evidence for most health outcomes. Among the diagnoses most consistent with NAS-based evidence, there was an increase in motor vehicle accidents, alcohol abuse, overdose injury and a decrease of chronic pain after recreational cannabis legalization (each meeting statistical significance for both comparisons with control states) . When using less stringent criteria (P<0.05 for both comparisons with NY and OK but where the relative risk for only one comparison was in the same direction as NAS), admissions for abuse of other substances and social anxiety disorder also increased with recreational cannabis legalization (figure 3). Effects of cannabis liberalization on psychiatric outcomes (schizophrenia and bipolar disorders) were either not consistent or weakly consistent with NAS (figure 3). When differences between the two control states (NY and OK) were also significantly different (suggesting effects potentially related to secular trends)(eFigure 1), changes in diagnoses were mostly not concordant with NAS findings or one of the comparisons between CO and either NY or OK did not exhibit statistical significance.”

And to the Discussion section (Page 15, Lines 298-321):

“Following legalization of recreational cannabis, changes in rates of medical diagnoses reflected NAS-based evidence for most health outcomes. The absence of statistically significant differences in these outcomes between the two control states (NY and OK) over the same time period provides some evidence that these observed differences were less likely related to broader secular trends. After legalization of recreational cannabis, there was an increase in motor vehicle accidents, alcohol abuse, overdose injury and a decrease of chronic pain. In addition to information provided in the NAS summary,[3] the association of cannabis with motor vehicle accidents has been highlighted by recent literature.[7, 8] Consistent with our findings, a substantial bidirectional comorbidity between cannabis use and alcohol use has been previously demonstrated, resulting in a moderate level of evidence designation for this relationship per the NAS.[9] Also compatible with the NAS summary, we report that recreational cannabis legalization was associated with an increasing number of admissions for overdose injury. These included overdose of analgesics, barbiturates, sedative, hypnotics and psychotropic drugs (supplementary eTable 2). This finding underlines the association of cannabis use with other behaviors of drug addiction as highlighted by a recent cross-sectional survey of over 30 thousand community-living US adults.[10] Finally, reduction of chronic pain, especially neuropathic pain, in cannabis users is well known, although prior literature has focused on medical rather than recreational cannabis.[11, 12]

Except for social anxiety, the effects of recreational cannabis legalization on psychiatric outcomes (schizophrenia, bipolar disorder, depression) were weakly or not consistent with prior NAS-evidence. This finding may be related to residual confounding by state-level characteristics, unaccounted secular trends, or insufficient longitudinal follow-up.”

3. Committee on the Health Effects of Marijuana, Board on Population Health and Public Health Practice, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. The Health Effects of Cannabis and Cannabinoids : The Current State of Evidence and Recommendations for Research. Washington, DC: The National Academies Press; 2017.

7. Rogeberg O. A meta-analysis of the crash risk of cannabis-positive drivers in culpability studies--Avoiding interpretational bias. Accid Anal Prev. 2019;123:69-78.

8. Hostiuc S, Moldoveanu A, Negoi I, Drima E. The association of unfavorable traffic events and cannabis usage: a meta-analysis. Front Pharmacol. 2018;9:99.

9. Blanco C, Hasin DS, Wall MM, et al. Cannabis Use and Risk of Psychiatric Disorders: Prospective Evidence From a US National Longitudinal Study. JAMA Psychiatry. 2016;73(4):388-95.

10. Hayley AC, Stough C, Downey LA. DSM-5 cannabus use disorder, substance use and DSM-5 specific substance-use disoreders: Evalyating comorbidity in a population-based sample. Eur Neuropsychopharmacol. 2017;27(8):732-743.

11. Whiting PF, Wolff RF, Deshpande S, et al. Cannabinoids for medical use: A systematic review and meta-analysis. JAMA. 2015;313(24):2456-73.

12. Hill KP. Medical marijuana for treatment of chronic pain and other medical and psychiatric problems: A clinical review. JAMA. 2015;313(24)2474-83.

4. Other changes would improve the manuscript:

More information should be provided about use of the HCUP database.

a) Which specific HCUP database was the source of the data analyzed? Was it the National Inpatient Sample (NIS)?

The source of the data analyzed was not the National Inpatient Sample (NIS) as this sample does not contain state-specific information. We used separate state HCUP databases for CO, NY and OK. Unlike the NIS, which is (by definition) a sample, these state-specific databases include data from every actual admission, providing direct and complete information regarding all healthcare utilization. We believe this is a major strength of the paper and are happy to have the opportunity to better clarify this important point. We have better clarified the source of the data analyzed in the Methods section (Page 6, Lines 101-104):

“We used separate HCUP databases for CO, NY and OK. Unlike the National Inpatient Sample (NIS), which is (by definition) a sample, these state-specific databases include data from every actual admission, providing direct and complete information regarding all healthcare utilization.”

b) How were diagnoses extracted from the HCUP database? Was only the primary diagnosis for each admission used? If so, this disregards some information, as multiple diagnoses may be coded for each inpatient admission: up to 30 in CO, 25 in NY, and 16 in OK.

Both primary and secondary diagnoses were extracted from the HCUP database for each admission used. Therefore, we believe we were able to capture all diagnoses and did not disregard any crucial information that could influence our results. Further details about how diagnoses were extracted from the HCUP database can be found in the following sentence in the Methods section (Page 6, Lines 104-106):

“Both primary and secondary diagnoses were extracted from each HCUP database for each admission used: up to 30 ICD-9 codes in CO, 25 in NY, and 16 in OK were provided for each encounter.”

c) Several of the substance use disorder diagnoses listed in eTable 2 are “in remission” diagnoses, raising the question of whether they were incidental to the diagnosis actually responsible for the hospitalization. Were these “in remission” diagnoses included in the statistical analyses?

We thank the reviewer for the insightful comment. “In remission” diagnoses were included in the original analysis for both “Abuse of other substances” (defined by ICD9 codes for alcohol, opioids, sedatives, anxiolytics, hypnotics, cocaine, amphetamines, hallucinogens and antidepressants) and “alcohol abuse” alone. We have run sensitivity analyses excluding “in remission” diagnoses from both “abuse of other substances” and “alcohol abuse”. “Alcohol abuse”, which increased in the original analysis based on stringent criteria, was again noted to increase after recreational cannabis

legalization using our most conservative criteria in the sensitivity analysis without “in remission” diagnoses. For “abuse of other substances” which only met 2 of our stringent criteria in the original analysis, results did not change in the sensitivity analysis where “in remission” diagnoses were excluded. These findings overall suggest that when the substance use disorder was in remission, it likely did not play a significant role in the overall results.

We have added the following sentence to the Methods section (Page 8, Lines 146-149):

“Because diagnoses of alcohol or other substance abuse ‘in remission’ may have been incidental and not directly responsible for the hospitalization, we performed a sensitivity analysis where we removed ‘in remission’ diagnoses from the group of ICD-9 codes defining alcohol abuse and other substance abuse (eTable 2).”

To the Results section (Page 14, Lines 280-281):

“In sensitivity analyses removing ‘in remission’ diagnoses, no meaningful changes were observed (relevant to alcohol abuse and abuse of other substances).”

And Discussion section, (Page 16, Lines 322-325):

“We demonstrate that abuse of alcohol or other substances remains higher after recreational cannabis legalization in a sensitivity analysis without ‘in remission’ diagnoses. Overall, these findings suggest that when the substance use disorder is in remission, it does not play a significant role in overall results.”

5. The authors should provide specific hypotheses regarding the association between recreational cannabis legalization and hospitalization rates, based on the NAS review. They can then use their statistical findings to confirm or dispute the specific hypotheses. This is the expected structure of a typical scientific paper. I suggest that the authors limit their hypotheses to putative cannabis health effects (either beneficial or harmful) for which the NAS review found either “conclusive,” “substantial,” or “moderate” evidence of an association. This greatly reduces the number of diagnoses being evaluated for influence by recreational cannabis legalization and lessens the possibility that failure to observe a significant influence was a false negative finding, i.e., due to absence of a significant association with cannabis, rather than true absence of an influence by legalization. I further suggest that the authors structure their discussion by the strength of evidence for the cannabis health effect.

We fully agree with the reviewer and again believe that our incorporating this suggestion has strengthened the manuscript. In the revision we have only focused on NAS diagnoses with moderate or substantial evidence. As mentioned above in the reply to comment #3, by only including higher NAS evidence diagnoses and adjusting for state-related characteristics such as % urban, income level, alcohol, tobacco abuse and psychiatric disorders, results of our initial analysis have changed substantially. We thank the reviewers for the opportunity to improve our manuscript. Specifically, we now demonstrate that after legalization of recreational cannabis there is an increase in admissions for motor vehicle accidents, alcohol abuse, overdose injury and a decrease of chronic pain (effects also highlighted by the NAS review). When using less stringent criteria, admissions for abuse of other substances and social anxiety disorder also increase with cannabis legalization.

We have clarified our hypothesis related to change of health outcomes as requiring this higher level of NAS evidence with cannabis legalization in the Introduction section (Page 5, Lines 82-88):

“On December 10, 2012, Colorado enacted Colorado Amendment 64, legalizing recreational cannabis. Following this date, adults aged 21 or older could grow cannabis plants privately, legally possess all cannabis from these plants, and give cannabis as a gift to other adults aged 21 or older.[4] After January 1, 2014, recreational cannabis could be legally purchased in retail stores.[5] We

hypothesized that changes in healthcare utilization and diagnoses most consistent with NAS-based evidence occur when access to recreational cannabis becomes liberalized."

4. Colorado Marijuana Legalization Initiative, Amendment 64 2012 (Colo.)

5. Wark, J. Up Early and in Line for a Marijuana Milestone in Colorado. The New York Times [Internet]. 2014 Jan 1 [cited 2019 Feb 14];U.S.:[about 2 p.]. Available from: https://www.nytimes.com/2014/01/02/us/colorado-stores-throw-open-their-doors-to-pot-buyers.html

We have added the following sentences in the Methods section (Page 8, Lines 141-144):

"For each health endpoint of interest in the NAS review, the weight of evidence regarding the statistical association of recreational cannabis with a specific health endpoint or the therapeutic use of cannabis had previously been categorized into substantial, moderate, limited or no evidence.[3] We opted to focus on NAS health endpoints with either substantial or moderate evidence."

3. Committee on the Health Effects of Marijuana, Board on Population Health and Public Health Practice, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. The Health Effects of Cannabis and Cannabinoids : The Current State of Evidence and Recommendations for Research. Washington, DC: The National Academies Press; 2017.

Results and Discussion have now been structured based on the strength of evidence for the cannabis health effect as shown in the answers to comment #3.

6. The authors should be more precise in some of their terminology and chains of reasoning.

a. The term "abuse" is used to refer to both "abuse" and "dependence" diagnoses (legend to Table 1). This will be confusing for readers and is contrary to the terminology of both DSM-IV and DSM-5. I suggest the authors use the term "substance use disorder," which encompasses both abuse and dependence and is the term used in DSM-5.

We agree and have substituted the words "substance abuse" with "substance use disorder" in both body and legend of table 1.

b. The manuscript frequently mentions "Amendment 64" in isolation. Although this term is defined in the Introduction (but first mentioned earlier in "Strengths and limitations"), it will not be readily familiar to readers outside of CO. More informative would be to mention what it stands for, i.e., "2012 legalization of recreational cannabis."

We now refer to "Amendment 64" only in the Introduction (Page 5, Line 82), but use the words "legalization of recreational cannabis" elsewhere in the manuscript.

7. The authors are rightly concerned that secular trends might differ across states (p. 11; see item #3 above). However, a single pair-wise comparison between CO and another state does not alleviate this concern, as the concern is generated by the possibility that multiple states vary in secular trends. The manuscript contains evidence that this is the case: pairwise comparisons between CO-NY, CO-OK, and NY-OK are inconsistent for several diagnoses. The solution is to include as many states as possible in the analyses.

As mentioned above in our reply to comment 2b, we agree that comparing CO with all other states represents the "gold standard" in managing secular trends across different populations. However, many states do not have HCUP data up until 2014 or do not have it at all. We do wish to point out that this analysis could have been restricted to changes within Colorado alone and the substantial effort and resources expended to at least include not only one other state but two intentionally disparate states as elucidated above and in the manuscript. Among those states that have HCUP data, some

come at great cost. We have mentioned the lack of comparison with all states among the Limitations of our study (Page 18, Lines 363-368):

“We did not select all US states as controls. Most states do not have HCUP data up until the end of 2014, each requires a separate HCUP application, and each carries significant costs. We acknowledge this represents a limitation as we could not fully account for secular trends that might influence results. However, we have adjusted all analyses for urbanicity and other state-level differences in order to minimize the impact of secular trends.”

8. The authors state that the “increased frequency of cannabis abuse diagnoses in CO helps validate…that legalization would result in greater use” (Discussion, p. 12). In fact, they observed an increased frequency of hospitalizations for cannabis abuse. This does not necessarily imply greater prevalence of cannabis use or even of abuse. The association could reflect increased severity of cannabis abuse, resulting in a higher rate of hospitalization for treatment (rather than, e.g., outpatient treatment).

We thank the Reviewer for raising this point. We agree and have added the following sentence to the Discussion section, Strengths and Weaknesses subsection (Page 17, Lines 355-357):

“Moreover, we cannot exclude that the increased frequency of cannabis abuse hospitalizations is simply related to severity of cannabis abuse rather than to a greater prevalence of cannabis use.”

9. The manuscript frequently mentions “cannabis legalization” without clarifying that it is referring to legalization of recreational cannabis. CO already had legalized medical cannabis throughout the period of this study.

We have followed the Reviewer’s recommendation and now only refer to cannabis legalization as “recreational cannabis legalization” throughout the manuscript.

10. The authors should put their findings regarding hospitalizations for cannabis and alcohol abuse in the context of large, multi-state, population-based analyses of medical cannabis legalization. For example, some studies find legalization of medical cannabis associated with increased cannabis use but not increased cannabis use disorder (e.g., Williams et al., Addiction, 2017). Could legalization of recreational cannabis have different effects than legalization of medical cannabis? What would explain any such differences?

We thank the Reviewer for this insightful comment and appreciate the reference provided (Williams et al., Addiction, 2017). We agree legalization of recreational cannabis may not have the same effects as legalization of medical cannabis at a population-level.

We hypothesize why such differences may occur in the Discussion section (Page 18, Lines 375-384):

“In our study, we demonstrate that legalization of recreational cannabis is associated with more cannabis abuse. Prior large, multi-state, population-based analyses have shown that legalization of medical cannabis is associated with increased cannabis use, but not increased cannabis use disorder.[26] One could speculate that medical cannabis consumption is more controlled and rationed when compared to recreational cannabis, the latter available in both retail stores and private homes and therefore perhaps more easily abused. Further studies are needed to quantify the amount of cannabis consumption (in grams) in states where only medical cannabis is legal versus states where both medical and recreational cannabis are allowed.”

26. Williams AR, Santaella-Tenorio J, Mauro CM, Levin FR, Martins SS. Loose regulation of medical marijuana programs associated with higher rates of adult marijuana use but not cannabis use disorder. Addiction. 2017;112(11):1985-1991.

11. Literature cited should be updated in some areas.

a. Reference #10 (O'Neill et al., 2017), cited to support cannabis treatment for PTSD symptoms, actually concludes that "Evidence is insufficient to draw conclusions about the benefits and harms of plant-based cannabis preparations in patients with PTSD."

We agree the evidence demonstrating any benefits or harms of plant-based cannabis in patients with PTSD is limited. Hence, we have removed PTSD and relative references from our NAS outcomes analysis.

b. Reference #12 (Li et al., 2012), cited for the association of cannabis with motor vehicle accidents, should be replaced with a more recent review, e.g., Rogeberg, Accid Anal Prevention, 2018; Hostiuc et al., Front Pharmacol, 2018.

We have substituted reference #12 with the more recent review suggested by the Reviewer.

12. The Abstract should include a brief description of statistical methods.

We have now added a brief description of statistical methods in the Abstract (Page 2, Lines 38-39):

"Negative binomial models were used to compare rates of admissions between states".

13. There are currently 31 (not 29) US state with legal medical cannabis (Introduction, p. 5, l. 15).

We thank the Reviewer for pointing this out, which may have been a product of changes since we initially drafted the manuscript. We have now corrected the number of states with legal medical cannabis in the Introduction section (Page 5, Lines 69-70):

"Thirty states and the District of Columbia now allow cannabis for the treatment of medical conditions.[2]"

2. State Medical Marijuana Laws 2016. National Conference of State Legislatures. http://www.ncsl.org/research/health/state-medical-marijuana-laws.aspx. Updated June 27, 2018. Accessed December 1, 2017.

14. Some headings in Table 1 should be clarified.

a. Title mentions "baseline" characteristics of the 3 study states, while the legend states that "All variables pertain to the 2010-2014 period," i.e., the entire study period, not just 2010.

We have changed the title of table 1 (Page 28):

"Clinical characteristics of the 3 study states over the 2010-2014 period".

b. The heading "NAS Diagnoses" actually refers to the hospital admission diagnoses from HCUP. The table legend should explain that the admission diagnoses listed were based on those found influenced by cannabis in the 2017 NAS review.

The sentence "Admission diagnoses listed were based on those found influenced by cannabis in the 2017 NAS review" was added to the body of table 1 legend (Page 30).

c. The NAS review should be explicitly cited in the table legend.

We have now explicitly cited the NAS review reference in the table 1 legend (Page 30).

VERSION 2 – REVIEW

REVIEWER	David A. Gorelick,D MD, PhD University of Maryland School of Medicine USA
REVIEW RETURNED	10-Mar-2019

GENERAL COMMENTS	The authors have satisfactorily addressed almost all my concerns, resulting in a much improved manuscript. However, wording changes must be made to make the manuscript more accurate and clarify some key concepts. 1) Wording in the Abstract (l. 41) and many places throughout the text mistakenly states that the authors studied the prevalence of certain conditions and diagnoses. For example, on page 15, the authors describe “changes in rates of diagnoses” (l. 280); on page 22, they state that “recreational cannabis legalization is associated with an expected increase in cannabis abuse” (l. 442-443). This is misleading for readers, as the authors actually studied the prevalence of hospital admissions for certain conditions and diagnosies, not the prevalence of individuals with these conditions and diagnoses. Many individuals with a condition or diagnosis are not hospitalized for their condition or diagnosis. The terms “admission,” “hospital admission,” or “hospitalization” should be inserted (as appropriate) everywhere the authors mention a condition or diagnosis as an outcome variable in their study. 2) Similarly, when describing covariates in their statistical models (p. 10), the authors should refer to “diagnosis of alcohol abuse” (l. 183) and “diagnosis of psychiatric disorders” (l. 184). 3) The authors acknowledge as a limitation of their study the use of inpatient hospitalization data, which does not include outpatient treatment. This is a valid limitation, but the authors’ discussion does not go far enough. a. Even inclusion of outpatient data would not capture individuals with the condition or diagnosis who are not in formal treatment. The authors should acknowledge that treatment databases do not capture all individuals in the population. This gap may be especially large for substance use disorders. Many individuals with substance use disorders never enter treatment; even those who do more likely get outpatient treatment, rather than hospitalization, so would not show up in the authors’ study. b. In the US healthcare system, hospitalized individuals tend to be sicker, with severer forms of their illness, than outpatients. The authors should mention this “bias” and discuss how it might influence their conclusions about overall healthcare costs. 4) The listing of study strengths and limitations at the beginning of the text needs improvement. a. The use of only 2 comparison states is a major limitation to which the authors devote an entire page of text. This warrants its own bullet in this listing. b. To make room for the above limitation, I would drop mention of the statistical control for multiple hypothesis testing. c. Several bullets should be reworded to convey the actual data used. i. 1st bullet should mention hospital admissions, rather than healthcare utilization. Hospital admissions are only a portion of all healthcare utilization ii. 2nd bullet should mention comprehensive state-wide hospital admissions data, rather than HCUP data (which term will be meaningless to most readers).

	iii. 4th bullet should mention recreational cannabis legalization, rather than Amendment 64 (which the authors have already done elsewhere in manuscript). 5) The manuscript incorrectly states that 30 states currently have legalized medical cannabis (p. 6, l. 77), citing a 2016 web report that was "Updated June 27, 2018" but "Accessed December 1, 2017" (reference #2). According to the web site cited in reference #2 (accessed March 10, 2019), 33 states have legalized medical cannabis. The text and reference #2 should be corrected. 6) "abuse" should be replaced by "use disorder" (p. 22, l. 443), as the authors did elsewhere in the manuscript. 7) Evidence is "previously published," not "previously established" (p. 22, l. 448).

VERSION 2 – AUTHOR RESPONSE

Reviewer:

The authors have satisfactorily addressed almost all my concerns, resulting in a much improved manuscript. However, wording changes must be made to make the manuscript more accurate and clarify some key concepts.

We appreciate the Reviewer's positive feedback and the constructive suggestions, which we have addressed as follows to strengthen the paper:

1. Wording in the Abstract (l. 41) and many places throughout the text mistakenly states that the authors studied the prevalence of certain conditions and diagnoses. For example, on page 15, the authors describe "changes in rates of diagnoses" (l. 280); on page 22, they state that "recreational cannabis legalization is associated with an expected increase in cannabis abuse" (l. 442-443). This is misleading for readers, as the authors actually studied the prevalence of hospital admissions for certain conditions and diagnoses, not the prevalence of individuals with these conditions and diagnoses. Many individuals with a condition or diagnosis are not hospitalized for their condition or diagnosis. The terms "admission," "hospital admission," or "hospitalization" should be inserted (as appropriate) everywhere the authors mention a condition or diagnosis as an outcome variable in their study.

We understand the Reviewer's concerns. We now use the terms "admission", "hospital admission", or "hospitalization" in place of conditions as outcome variables throughout the manuscript.

2. Similarly, when describing covariates in their statistical models (p. 10), the authors should refer to "diagnosis of alcohol abuse" (l. 183) and "diagnosis of psychiatric disorders" (l. 184).

We agree with the Reviewer and have added the term "diagnosis of" to the Statistical Analysis section (Page 9, Lines 169 and 170).

3. The authors acknowledge as a limitation of their study the use of inpatient hospitalization data, which does not include outpatient treatment. This is a valid limitation, but the authors' discussion does not go far enough.

We appreciate the Reviewer's insightful comments and have modified the Discussion section as detailed below.

a. Even inclusion of outpatient data would not capture individuals with the condition or diagnosis who are not in formal treatment. The authors should acknowledge that treatment databases do not capture all individuals in the population. This gap may be especially large for substance use disorders. Many

individuals with substance use disorders never enter treatment; even those who do more likely get outpatient treatment, rather than hospitalization, so would not show up in the authors' study.

We have added the following sentence to the Discussion section (Page 19, Lines 394-400):

"However, even inclusion of outpatient data would not capture individuals with a specific condition or diagnosis who are not in formal treatment. Indeed, treatment databases do not capture all individuals in the population. This gap may be especially large for substance use disorders. Many individuals with substance use disorders may never enter treatment; even those who do more likely receive outpatient treatment, rather than a hospitalization, so they may not have been included in our study."

b. In the US healthcare system, hospitalized individuals tend to be sicker, with severer forms of their illness, than outpatients. The authors should mention this "bias" and discuss how it might influence their conclusions about overall healthcare costs.

We thank the Reviewer for making this point. We have added the following sentence to the Discussion section (Page 15, Lines 299-305):

"Moreover, in the US healthcare system, hospitalized individuals tend to be sicker, with more severe forms of specific illnesses and higher inpatient costs compared to outpatients. Therefore, our conclusions on overall neutral effects of recreational cannabis on healthcare costs only reflect higher inpatient costs and not overall costs. Outpatient costs may include use of specific medications or substance detoxification programs that are not captured in our study and may be more cost-effective compared to inpatient treatments."

4. The listing of study strengths and limitations at the beginning of the text needs improvement.

We agree and have modified the Strengths and Limitations Section (page 4) as follows:

a. The use of only 2 comparison states is a major limitation to which the authors devote an entire page of text. This warrants its own bullet in this listing.

We have added a separate bullet mentioning the limitation of having only 2 comparison states, OK and NY (Lines 58-60).

b. To make room for the above limitation, I would drop mention of the statistical control for multiple hypothesis testing.

We have dropped mention of the statistical control for multiple hypothesis testing to make room for the additional bullet mentioned above.

c. Several bullets should be reworded to convey the actual data used.

i. 1st bullet should mention hospital admissions, rather than healthcare utilization. Hospital admissions are only a portion of all healthcare utilization

ii. 2nd bullet should mention comprehensive state-wide hospital admissions data, rather than HCUP data (which term will be meaningless to most readers).

iii. 4th bullet should mention recreational cannabis legalization, rather than Amendment 64 (which the authors have already done elsewhere in manuscript).

We thank the Reviewer for greatly improving the opening "Strengths and Limitations" section of the manuscript. We have made the changes as suggested.

5. The manuscript incorrectly states that 30 states currently have legalized medical cannabis (p. 6, l. 77), citing a 2016 web report that was “Updated June 27, 2018” but “Accessed December 1, 2017” (reference #2). According to the web site cited in reference #2 (accessed March 10, 2019), 33 states have legalized medical cannabis. The text and reference #2 should be corrected.

We are grateful to the Reviewer for noticing this mistake. We have changed the number 30 to 33 in the Introduction section (Page 5, Line 70). We have also modified reference #2 to better reflect the most recent access date of the web report mentioned (Page 23, Line 453).

6. “Abuse” should be replaced by “use disorder” (p. 22, l. 443), as the authors did elsewhere in the manuscript.

We agree and have replaced the word “abuse” with “use disorder” in the Conclusions section (Page 20, Line 409).

7. Evidence is “previously published,” not “previously established” (p. 22, l. 448).

We appreciate this comment and have substituted the words “previously established” with the words “previously published” in the Conclusions section, (Page 20, Line 414).

VERSION 3 - REVIEW

REVIEWER	David A. Gorelick, MD, PhD University of Maryland School of Medicine USA
REVIEW RETURNED	03-Apr-2019

GENERAL COMMENTS	The authors have satisfactorily addressed all my remaining concerns. I have no further comments.