Dear Dr. Merino:
Thank you, the editorial team and the reviewers for very thoughtful and helpful comments and suggestions on our paper. We have revised the manuscript in the ways suggested and provide the results of multiple new sensitivity analyses and robustness checks that you have recommended. We think the paper is substantially stronger for these changes. We provide the details of all changes and our responses to comments below.

Reviewer’s comments are in bold, responses in normal type.

Editors:

1. **Clarify the topic for an international audience and provide greater background information. Why is the experience in Massachusetts so important?**

   We have tried to demonstrate the relevance for an international audience in the revised first paragraph of the introduction (page 5) and on page 8.

**Why did you select NY, NJ and PA as comparators?**

We initially investigated the possibility of using several contiguous and nearby states to Massachusetts in New England as comparator states. However, most of these states had very small black and Hispanic populations and many had incomplete collection of data on race/ethnicity of admitted patients, a key focus of our paper. Thus, we excluded these as comparators. Unfortunately, incomplete data collection on race/ethnicity is a feature of many state inpatient databases in the US. We ultimately selected three comparator states that balanced geographic proximity to Massachusetts with the need for adequate sample sizes of black and Hispanic residents and that had near complete data on race and ethnicity of admitted patients. All three selected controls have sizeable black and Hispanic populations and are in the “Mid-Atlantic” group of states that is adjacent to New England, and NY is actually contiguous with Massachusetts. For these reasons we (in two recent BMJ papers-Lasser, et. al, 2014 and Hanchate, et. al., in-press)[1 2] and others studying Massachusetts health reform (cites by us in the paper) have used this same set of comparator states in previous papers.

We limited the number of comparator states both because of our perception that the selected states could serve as adequate controls and for feasibility reasons. To obtain analyzable data from each state (separately) is a fairly lengthy and complex process involving a unique application requesting inpatient data from the relevant authorities in the state, signing contracts regarding data use, paying substantial user fees and learning and cleaning each unique database and receiving permission to publish each manuscript. Addition of more states to our analysis at this point would likely take an additional 6-9 months and thus we would prefer
not to pursue this option; we would be willing if deemed essential by the editors.

As can be seen from table 1 of the paper, there are baseline differences in several state characteristics between Massachusetts and control states. As we described, in all analyses we control for these baseline characteristics in multivariable regression models. We also perform a sensitivity analysis in which each state is sequentially removed from the group of 3 control and the models re-estimated to make sure that our results were not driven by an outlier state. We retain both of these analyses from the paper originally submitted.

In addition, as per the recommendation of one reviewer, we re-analyzed our data using a propensity score-matching scheme to identify and use only counties in control states that had characteristics matched to those in Massachusetts. This did not alter our findings (See new Appendix of the manuscript). We believe that the control states selected are, therefore, appropriate. We can think of no other states that, a priori, would be superior controls. We elaborate on these points in the limitations section of the manuscript (page 20).

In order to test whether pre-reform trends in admissions were similar in Massachusetts and control states, we conducted two pre-reform trend comparisons of ACSC admission rates (for overall composite, acute composite and chronic composite measures) and find no evidence for diverging trends pre-reform. Details of these analyses and the results are provided in the Appendix.

2. **Do you have more recent data with longer follow up? There was a lot of chaos for several years after reform went through in Massachusetts.**
   We recognize that the time period following implementation of the reform is not likely to capture the results of a mature system. Thus, we excluded the first two years after the start of the reform in all analyses (unchanged in this revision). However, as requested, we also conducted a sensitivity analysis in which we added an additional 12 months of post reform data, extending to the 3rd quarter of 2010. Adding these later months of observation did not change our results (see Appendix Table 1). Although we have no data beyond the 3rd quarter of 2010 (four years after the start of implementation), there is good reason not to use data up to the present: in 2012, Massachusetts passed a new cost control and payment reform bill (so called “Chapter 224”) that could well have influenced access to care (worsened or improved) in the state and thus would make it difficult to disentangle the effects of the main health reform law of 2006 and the subsequent 2012 law.

3. **You need to take into account the fact that there are a lot of variables other than insurance status that affect hospitalizations, including access to primary care, culture, availability of specialists, etc. Massachusetts may have a very high bias toward specialty care rather than primary care which is not the**
case in many other part of the country. The authors need to address the differences in health care cultures between states.

We agree that a number of variables, such as physician supply, could be associated with hospitalization rates and could vary between Massachusetts and control states. Specialist physicians per population does substantially vary among states, however, the differences between Massachusetts (2.59/1000 residents) and control states used in our analyses, NY (2.18/1000), NJ (1.58/1000) and PA (1.89/1000) are not as dramatic. In Massachusetts and all control states there is a substantial tertiary care and teaching hospital presence. To directly address concerns about physician supply, we controlled for an additional variable in all our models that captures primary care physician supply in the area of residence of cohorts in our study. This variable, Health Professionals Shortage Area (HPSA) designation is calculated by the Health Resource Services Administration of the US government and is commonly used in research on health services utilization. Beyond this, we are unable to capture “medical culture” in our models; however, we have added an acknowledgement that this could affect the results in the limitations section (Page 20, paragraph 1).

4. As a result of the reform, a new set of previously uninsured patients may have gained access to specialty care and an increase in hospitalizations. But those patients may now be healthier. Hospitalizations are not a health outcome and may be problematic metric. One of the reviewers mentions that mortality went down after the reform, which suggests better care. Do you we have mortality data to address?

In the US multi-payer system, most of those age <65 are covered by a variety of private (commercial) insurances and some under public insurances. Therefore, there is no way to obtain complete claims data on utilization of outpatient care at the community or population level that could directly indicate changes in access to (utilization of) care over the relevant time period, although such data are now beginning to be produced in some states (All Payer Claims Data in Massachusetts stated in 2009). Also, unlike in countries with national health systems, there also no regional or national authorities that collect or pool data across outpatient medical practices that could be used to assess changes in population-wide access. In this context, the most well validated and established method for detecting changes in population-wide access to care, that may occur due to a population-wide policy change, uses changes in preventable hospitalizations as an indirect but objective metric.

Underlying this concept is the expectation that timely outpatient care of some medical conditions can reduce the risk of hospitalizations certain conditions. It is true that hospitalization rates for certain conditions that are not clearly improved by timely outpatient care or procedures that are discretionary (hip replacement, CABG) might be expected to rise following new access to physicians on gaining
new insurance. However, in the present study, we examine only conditions for which outpatient treatment has been demonstrated to reduce the risk of hospitalization (Ambulatory Care Sensitive Conditions), such as urinary tract infection or congestive heart failure. The specific ACSCs that we used were each identified by US Agency for Health Care Quality and Research as having a strong evidence base supporting their use for this purpose and they are routinely used in US government assessments of changes in access to care.

A very large literature has accumulated over the past 25 years that has validated this approach (we cite only a fraction of these in the paper) and it continues to be commonly used in health services research of the type we conducted. While we cannot exclude the possibility that admission rates for ACSCs would increase for some residents, a substantive effect of this sort would not be expected based on this substantial prior literature. We have added this caveat to the revised paper (page 21, starting last paragraph).

A very well conducted study did recently demonstrate that state-wide mortality declined in Massachusetts following reform, in relation to controls. The reduction in mortality observed may well be due to increased insurance coverage and access to care; however, despite the robust statistical finding of a large reduction in mortality associated with reform, corroborative evidence from other objective markers of healthcare utilization – outpatient care, inpatient care, ED care – is limited, thereby leading to a lack of clarity in the potential pathways for a large reduction in mortality. In addition, this study was based on the entire population whereas ours focuses on a subset, those more likely to have acute and chronic medical problems, which could influence the results in the two studies. It seems unlikely that any single study will definitively answer the question of whether access improved following reform; more likely, the story will be told by the totality of disparate studies on this topic.

We have data on inpatient mortality only, which we do not think can be useful to address this question.

We have now cited this mortality study in the introduction (page 6) and discussion sections (page 23).

5. There were significant sociodemographic differences between Massachusetts and the other states at baseline. Based on the data presented, it seems there was less inequality in access to care before the reform was implemented. This raises the question whether the findings in Massachusetts can be generalized to other states. Can you please comment on these differences in the manuscript? The disparity problem seems much smaller and the numbers affected by the intervention in the intervention region compared to the control regions, so only a small affect would be anticipated.

We have bolstered the description of this issue in the third limitation in the
6. Do you have information about hospitalizations by type of hospital (safety net vs. not safety net; for profit vs. non-for profit; tertiary vs. community; academic vs. not academic; etc.)?

We have added information on safety net status, profit status and teaching status to our Table 2 (page 37-38).

7. The study did not find a difference over time. But there are some possible explanations that you should explore. You are looking at an indirect outcome, inpatient hospitalizations, as a proxy of the quality of outpatient care. If the outcomes are not identical (hospitalizations may actually reflect better care for some patients, for example, as discussed above) then the differences may hide any effect. Can you comment on the issue of using hospitalizations as a proxy and the possibility that the choice of proxy may hide the effect?

Please see response to point number 4 above.

8. With all studies like this you also need to ask whether there are other temporal or geographical factors which would confound the association. Can you please comment?

We very much agree. In our multivariable models, we control for age, gender, race/ethnicity, county median income, county median unemployment, quarterly (temporal) effects and now, as mentioned, a measure of regional physician supply (HPSA designation). Unfortunately, we do not have data available on other plausible confounders that we can incorporate in to models. However, in most prior analyses of the Massachusetts health care reform (including two of ours in the BMJ using these same data), this is the constellation of variables typically used as control variables. It is possible that other variables that may be confounders exist. We know of no large scale state policy changes or health care system infrastructure or market changes (other than an economic downturn which is captured by income and unemployment variables in our model) occurring during the study period in Massachusetts or control states that would be obvious candidates.

We have emphasized this caveat in the revised draft (page 20).

Reviewer #1

1. This is an important policy question and adds to a large body of evidence on the MA reform experience.

2. The general study design is appropriate, though the authors have not done enough to demonstrate that:
a. The choice of control states was appropriate, especially given the markedly different racial/ethnic make-up and income in these states compared to Massachusetts.  

Please see response to Editors’ comment 1.

b. Did they test whether the pre-expansion trends were similar?  

Yes. Please see response to Editors’ comment 1 and details in new Appendix to the paper.

c. They should also seek to broaden their control group or find states with more similar demographic features, especially since disparities are an important focus of their study design.  

Please see response to Editors’ comment 1.

d. How sensitive are their findings to the choice of functional form and how they model their standard errors.  

At the suggestion of the reviewers, we undertook multiple sensitivity analyses to test the robustness of our findings to alternative regression model specifications. Our baseline line model is a Poisson regression. We re-estimated the regression using 1. baseline model specifying a county-level fixed effects regression structure (with cluster-adjusted standard errors) to capture nesting of cohorts within county, 2. alternate clustering-adjusted standard errors using bootstrap, 3. linear probability models; 4.interrupted time series approach and 5. negative binomial models. We found no change from our baseline model for any of these alternative models. Details of the modeling and the results are shown in the Appendix.

e. The “DDD” model based on county uninsured rates isn’t really a ideal approach, since even low-uninsured counties aren’t a control group, as the DDD framework suggests.  

We agree that this is not technically a DDD analysis and have removed this description of the approach. However, we have retained this analysis because of its intuitive appeal: irrespective of changes going on in control states, if a change in ACSC admission rates occurred following reform in Massachusetts and this change occurred via an insurance expansion, one would expect the change to be greater in counties with the highest rates of uninsurance pre-reform. This is the question we address. Our analysis is identical to a recently published analyses of pre-post reform ACSC admission rates in the Medicare population.[4]
3. Assuming that the results are similar even after accounting for these concerns, the authors need to be much more circumspect about their findings. The paper states numerous times that the MA reform has produced “modest changes” – this seems to be out of whack with many of the changes documented in the literature. More broadly, the authors make a very strong conclusion that access to care in the state didn’t improve, when they are looking at one fairly narrow outcome.

Our characterization of the changes that have occurred after the reform as “modest” is accurate from the point of view of state-wide changes—most studies have shown 1-5% state wide improvement on measures of self reported access, for example. The reviewer correctly points out, however, that since only 5% of the population acquired new insurance, it is likely that there were substantial effects in that 5% of the population to produce 1-5% changes in measures of access in the whole population. Thus, whether the change is judged as large or small depends on whether one is looking at aggregate state-wide effects or inferring the effects on those gaining insurance. Thus, we have modified our characterization of the changes due to reform demonstrated in prior literature throughout the manuscript to better describe the impact on the newly insured, most often simply by not using terms of judgment such as “Modest”.

We agree that ACSC admissions is but one measure of access and thus have modified our conclusions to better reflect this view (page 18 last 2 sentences of the first paragraph): “Most but not all prior studies have shown improvements in multiple measures of access to care with less evidence that racial and ethnic disparities improved. Using the specific hospital-related metric of preventable hospitalizations, we were not able to demonstrate significant improvements in access or disparities in access to outpatient care following implementation of Massachusetts health care reform”.

4. INTRO: The authors describe the gains in access to care under MA reform from prior studies as “modest,” which I think is not really accurate and understates the policy’s impact. The fact that MA already had high coverage rates at baseline may obscure these relative changes, but the authors should not minimize this impact on those who were previously uninsured.

We agree. We have deleted “modest” from our description of insurance coverage gains after reform.

5. INTRO: Your literature review also doesn’t mention the recent Annals paper (Sommers, Long, & Baicker) looking at mortality (all-cause and health-care related) before and after the reform, which would certainly qualify as an “objective” health measure.

We now include references to this paper 3 times in the revision.
6. **Methods:** How did you select a 21 month pre & post timeframe? What happens if you add extra data? Some changes from coverage expansion may take longer, and this would also increase your power. The sensitivity analysis with a shorter time frame is less helpful.

The rational for the 21 month time periods and the results of analyses that extend the post reform period of observation are presented in the Appendix. An additional year of post reform data resulted in an increase in overall ACSC admissions in Massachusetts but no narrowing of black-white or Hispanic-white disparities. As mentioned in the manuscript (page 11): “This definition of the pre and post reform intervals, used in prior studies of the Massachusetts reform (refs given), balances the benefits of having a large enough window of observation to establish pre and post reform ACSC admission levels and sample size adequacy and avoiding time periods more distantly removed from the reform in which ACSC admission rates could be influenced by factors unrelated to the reform (substantial decline in US health care expenditures accelerating in 2010, e.g.).(ref given).

As recommended, we deleted the sensitivity analysis using a shorter interval from the revised manuscript.

7. **Methods:** If you have the county-level unemployment and income measures, just use those measures as the covariates.

We believe that defining county median income and unemployment as categorical variables in our models has more appeal that using the continuous variable. Dividing counties by tertile of median income captures income in the county relative to other counties in the same state which is appropriate because what is important is income relative to costs (purchasing power) in your area (state). Using a continuous measure of median income does not do this—a county in Massachusetts and a county in Pennsylvania may have the same incomes but if costs of living are higher in Massachusetts, the effective income would be different; revising income measures adjusted for cost-of-living differences is likely to be imprecise due to the complexity of factors affecting actual cost of living across areas. Thus, we retained the categorical definition of these two variables in the revised manuscript, consistent with methods in our two prior studies using these data in the BMJ.[1 2]

8. **Methods:** Do the data fit a Poisson distribution well? If there is evidence of overdispersion, a negative binomial model would be more appropriate.

We found evidence of overdispersion in our baseline models and thus controlled for it in our analyses. In addition, we re-estimated models using a negative binomial distribution. Results from these analyses did not differ significantly from
our baseline model specification (Poisson). Please see description of modeling and results in the Appendix.

9. **Methods:** Also, it’s generally become standard in the econ literature to prefer linear models to non-linear models for diffs-in-diffs analysis (see Norton and Karaca-Mandic) to allow for straightforward interpretation of the coefficients, so this would be a helpful sensitivity analysis.

We conducted sensitivity analyses using linear probability models and found no significant difference from results from our baseline model. Please see description of modeling and results in the Appendix (Tables 1 and 2).

10. **Methods:** The comparison between the high and low counties within Massachusetts doesn’t really make sense as a diffs-in-diffs-in-diffs model, since neither is a control group (they both were impacted by the policy, though potentially to a different degree). What would be more meaningful would be to stratify your control group in the other states similarly – compare high uninsured counties in MA to high uninsured counties elsewhere, and low uninsured counties in MA to low uninsured counties elsewhere.

Please see response to reviewer 1 comment 2, part e.

11. **Methods:** The authors provide no details on how they modeled the standard errors, which is an important issue in any diffs-in-diffs analysis.

We have added details on modeling of standard errors in our baseline models and sensitivity models. Please see response to reviewer 1 comment 2, part d and the Appendix.

12. **Methods:** Massachusetts is a referral center for New England more generally. If patients are coming to MA hospitals from outside the state, they will show in your dataset, yes – even though they aren’t eligible for the state’s health reform. This will bias your study towards not finding any impact.

We have deleted the very small fraction of out of state admissions from all analyses.

13. **Results:** The descriptive statistics suggest that MA looked quite different from your control states, particularly for race and ethnicity. How were these control states chosen? Why not broaden to other states with closer demographic match, and also to increase the precision of your estimates?

Please see response to Editors’ comment 1.
14. **Results:** Did the authors formally test whether the pre-expansion trends were similar in your outcomes for treatment and control states which is the underlying assumption for the D-in-D model?

   Yes. Please see response to Editors’ comment 1 and the Appendix for details of two tests of this.

15. **Results:** Table 4 shows pretty large point estimates for ACSC changes among blacks compared to whites (-4 to -6%), though p-values were non-significant. Given that the policy extended coverage to 5-7% of the adult population, this could represent a really large impact on disparities, but the precision of the estimates are quite broad. Also, technically, I wouldn’t call this a DDD – whites in MA aren’t really a control group, are they?

   We now say: “The lower bound of the 95% CIs for our composite overall ACSC measure is compatible with an overall reduction of 1.6 percentage points, and black-white and Hispanic-white reductions of 8.5 and 7.5 percentage points respectively, which would be substantial if true”. (Page 21).

   We also no longer refer to the race/ethnicity analyses as DDD analyses.

16. **Results:** Linear models would be very useful here, since the Poisson models provide relative change – but in linear models measuring absolute changes in ACSC, these might be statistically significant and look quite different from the Poisson results.

   Please see response to reviewer 1 comment 9, above. Results from linear models did not differ from baseline model results.

17. **Discussion:** “Hence, our results using this measure do not suggest that overall access to outpatient care, or racial and ethnic disparities in access to such care, improved to a significant degree following implementation of Massachusetts health care reform.”

   This is way too strong a statement. The abundance of evidence already shows that outpatient access to care under the state reform improved significantly – more people with a primary care provider, more able to afford their medical care, more with a preventive visit. I think a more accurate way to frame your findings would be to say, “For one specific hospital-related measure of outpatient access we don’t find any evidence of change after MA reform, unlike numerous previous studies of other access measures such as having a PCP, being able to afford care, and having a preventive health visit.”

   We have modified our language along the lines the reviewer suggests (last sentences of first paragraph of page 19).
18. Limitations: If your study is underpowered (wide confidence intervals on the D-in-D that can’t rule out relevant population-level changes), then dropping individual control states only makes the problem worse, not better. You come back to this point eventually on p. 20, but given this concern – that there may actually have been large worthwhile changes included in your confidence intervals (and perhaps even your point estimate). Accordingly, much more tempered conclusions would appear to be in order.

As indicated above, the purpose of dropping individual states in sensitivity analyses was to evaluate whether a single control state was driving our result point estimates (not confidence intervals)-none was. As indicated, we state that wide confidence intervals for some analyses can’t exclude meaningful changes (page 21, mid paragraph) and tempered our conclusions, as indicated above.

19. Also, the survey-based studies being cited are done with gold-standard national health interview data and similar sources; I’m hard pressed to imagine a source of recall bias large enough to explain all the evidence on MA reform’s improvements in access to care.

We agree that the surveys used in the prior literature are the gold standard and are done as well as could be done. It remains true, however, that surveys asking individuals about subjective perceptions (particularly those with modest response rates as these have) have the potential for bias, as has been discussed in some of the studies using these data have discussed.[5] We are less concerned about recall bias. However, social desirability bias (a cognitive bias) could occur since pre-reform samples had no knowledge of health reform but post reform samples knew that they were providing answers in the context of a new and much touted health reform. It is possible that post reform samples were influenced by the perceived social desirability of providing more positive answers regarding access, for example. It is plausible that this could account for much of the 1-5% of the respondents that reported improvements after the reform. We have left reference to potential bias, therefore. We will remove these if deemed advisable.

20. p. 20: The absolute changes being cited seem small only when you ignore the denominator. These correspond to cutting rates of people without a PCP by 30-40% - that’s a big deal. As the authors mention, the ACA itself is “only” going to increase coverage by 5-10%. Do the authors contend that coverage changes of that size are not particularly important?

Agree. Please see response to reviewer 1, comment 3.

21. p. 22: cutting the uninsured rate in half is hardly modest.

Agree. Please see response to reviewer 1, comment 3.
22. p. 24: the authors cite a study showing mortality reductions and access to care changes after Medicaid expansion. They do not mention the similar study in Annals of Int Med showing similar findings in Massachusetts, which would be worth comparing to the current results.

We now cite this paper in the discussion.

Reviewer: 2

1. This is a well-executed and well-written study showing that Massachusetts health reform did not improve rates of potentially preventable hospitalizations or disparities in these rates. The impact of Massachusetts health reform and of the Affordable Care Act on health outcomes has received too little attention and this study provides important evidence with a rigorous study design.

2. Although the study design is strong, and the difference-in-differences analysis is robust, this analytic approach could be missing a post-reform *upward* trend in avoidable hospitalizations in Massachusetts, a potentially important finding. An aggregate interrupted time series analysis on a "differenced" plot (intervention group rate minus control group rate) would be stronger and would have the ability to detect such a trend.

We have now conducted an aggregated interrupted time series analysis (intervention group rate minus control group rate), as suggested as a sensitivity analysis. Results did not vary from our results using our baseline model. Please see Appendix.

3. The authors should consider better tailoring the implications of this study to an international audience.

We have placed our study in a broader international context in the introduction (page’s 5 and 8) and discussion (end of paragraph on page 8).

5. It would be beneficial for the authors to state their hypothesis in the Introduction. A plausible hypothesis could be that the influx of newly insured could cause poor/delayed outpatient access (survey data notwithstanding) and therefore cause increasing preventable hospitalizations.

We have stated a hypothesis in the introduction (page 8).

6. Figure 1b shows a (slight but clear) increase in the Massachusetts' rate relative to the secular trend in the follow-up period, and it seems important to determine if this is a statistically significant trend increase. The study design is well-suited for using an aggregate interrupted time series approach (see Wagner AK et al. Segmented regression analysis of interrupted time series studies in medication use research.
As mentioned, our interrupted time series models did not show any statistically significant trend increase for any of the three ACSC measures. Please see appendix for details.

7. Presumably the authors used a 3-way interaction term in their regression models to generate their estimates of impacts on disparities in "difference-in-differences-in-differences" analyses. This is acceptable and correct but it is very difficult to obtain "statistically significant" results with 3-way interaction terms, and this approach could be missing important effects among key subgroups. Stratifying results by e.g. race/ethnicity might be a better option. This would not directly test whether disparities changed, but would detect the impact of Massachusetts health reform on these subgroups.

We have added adjusted D in D changes for each race separately as requested and now present these result in our revised table 4 (page 42). There were no significant changes for any race.

8. While this study is of substantial importance to a US audience and US policymakers, the interest to an international audience or the generalizability to international settings is not as clear. The authors discuss this briefly but should enhance their discussion of implications for non-US settings.

Please see response to reviewer 2, comment 3.

9. The authors had "insurance type" as a patient-level variable, presumably implying that they could detect rates of uninsurance at each hospitalization in addition to major types of insurance such as commercial/Medicare/Medicaid. This is important information that should be shown at baseline in Table 2 and potentially displayed as a before-after trend.

As requested, we now present these data in Table 2 (page 37-38).

10. The authors should describe their creation of a rolling average in more detail. Presumably this means that, in a given quarter, the number reported represents that quarter averaged with 3 other surrounding quarters. Were those quarters before, after, or surrounding the quarter reported? If I understand their approach correctly, Figures 1a and 1b should *not* have corresponding x-axis intervals because the "raw" Figure 1a results will either include quarters that preceded or followed the 21 month baseline period. It would also be important to know whether the moving averages include quarters that preceded the 21-month baseline or fell within the phase-in period.

We have clarified this description in the text of the manuscript (page 12, last paragraph).
11. A prominent paper found that Massachusetts health reform was associated with lower mortality rates. Such a finding could be mediated by a reduction in preventable hospitalizations, yet the present study did not detect this pattern. This might be worth a brief mention.

We now acknowledge these results in the paper introduction and discussion, as mentioned.

12. Did the authors control for repeated events within individuals? Although somewhat unlikely, this could bias results if a given state is better than another at preventing readmissions. Perhaps the authors do have a unique patient identifiers across hospitals? If so, please state.

We previously published a paper in the BMJ[1] using the same data as the present study to show that there were no significant changes in readmission rates in Massachusetts pre-post-reform, in comparison with these control states. Thus, we did not control for repeated events in our analyses for the current paper.

13. The authors presumably had the zip code of their subjects, but they do not list zip code in their list of variables that they obtained directly from hospital records.

We did have zip code information and have made the requested change (page 9).

14. Presumably patient zip codes were used to link patients to their county. Please explicitly state.

Correct. We have made the requested change (page 9).

15. The authors should note higher up in the text (e.g. in Data Sources rather than in Analytic Data Structure) that many of their variables are assigned at the county level because this is the only level that has an annual denominator and because some key variables such as SAHIE and Area Resource File are at the county level.

We have made the requested change (page 10)

16. Please also briefly note the limitations (and strengths) of geographically assigned characteristics, including county level versus more granular levels.

We have made this requested change (Page 21)

17. Standardizing rates by age and sex is not the strongest approach and leaves something to be desired because it is adjusting for so few characteristics. I understand that bin sizes limit the ability to further standardize. While I do not think it is crucial, the authors could potentially use marginal predictions to generate rates that are fully adjusted for all covariates, not simply age and sex. Also, it would
be helpful to have one sentence that clearly states where the standardize rates were used (plots and semi-adjusted results). Presumably the adjusted regression analyses did not use standardized rates as the dependent variable (because these could be adjusted for age and sex)? Please clarify.

As many epidemiologic studies do, we provide age and sex-standardized rates in the pre and post reform periods to provide a sense of comparability on basic patient characteristics but not adjusted for community-level factors that we hypothesized would influence ACSC admission rates. We capture full adjustment in the main D in D results presented in the same tables, which we retain from the original submission.

We think this is clear that adjusted regression analyses do not use standardized rates but will state again if deemed necessary.

18. Abstract, Conclusions: While understanding that the finding of no change in preventable hospitalizations after Massachusetts health reform is not completely novel, I believe this important, top-level finding should be reported in the abstract conclusion.

We currently state this (page 4).

19. Abstract, Conclusions: "Our findings do not suggest..." is a bit awkward. How about, "Our findings suggest that.... did not significantly lower..."

We have modified the language (page 18)

20. Abstract, Conclusions, last sentence: The end of this sentence is not clear. Putting the word "reduce" in front of "preventable hospitalizations" would clarify.

We have made this requested change (page 4)


We have made this requested change (page 6)

22. Use of the term "trends" in several places in the manuscript is confusing and might give the impression that the analysis is actually controlling for baseline trends (differing slopes between the groups) or analyzing changes in trends.

We have removed references to trends in all places.

22. Page 10, line 34: Should read "...(short-term diabetes complications, long-term diabetes complications, chronic obstructive ... etc)"

We have made this requested change (page 10)
23. Did the authors have hypotheses about how health reform would affect acute versus chronic preventable hospitalizations?

Due to space considerations, we did not elaborate these.

24. "Difference-in-differences" with the "s" on the end should be used consistently throughout.

We have made this requested change throughout.

25. How did the authors deal with out-of-state residents? Did they consider the possibility that Massachusetts health reform could influence in-state hospitalizations among out-of-state residents or vice-versa?

As described, these have now been deleted from all analyses.

26. Stryjewski has a recent HSR paper showing no impact of Massachusetts health reform on chronic disease outcomes.

We now reference this paper.

References