My original question: 6.2. Second, the county-level small area models strongly assume that all the county-level contextual variables (covariates) have universal impact on smoking outcomes across the whole US and thus could suffer substantial ecological bias in model prediction. Thus, the validation of county-level model-based estimates is critical and necessary. Although the BRFSS was not designed for county-level estimates, the county-level model-based estimates of smoking outcomes, when aggregated to state level, should be consistent with the state-level direct survey estimates from BRFSS. BRFSS were designed for reliable and accurate state-level estimates. If there are significant differences between model-based and direct survey estimates of smoking outcomes at state-level, the county-level small area models might be misspecified. The model validation implemented in this paper did not fully meet this goal.

Authors’ response: We thank the reviewer. Our findings suggest a huge variation within a state and that in many times counties are affected by neighboring counties irrespective of the border line imposed by a state. Therefore, we felt that we should not impose on counties within a state a model that will force the estimates for the counties to add to that of a state. Further, we have examined state level trends directly from the BRFSS and these show marked variation overtime. Given the large swings in the data at the state level overtime and the issue that populations may share things with neighboring counties in other states, we feel that this is a better approach in such analyses in the US where states have open borders and lots of interactions between counties.

My comments: If a county-level small area model specification is appropriate, an aggregated state estimate from all county-level model-based estimates should be close to BRFSS direct estimate. It is an internal validation of the small area model. In addition, based on authors’ comments, county-level model-based estimates are not stable sometime and might be affected by some uncertainties or interactions between counties. Top 10 and bottom 10 ranking might also be not stable or consist with other approaches. So, the authors should be very cautious about these ranks. If the authors insist in posting these ranks, they should point out the limitations of these ranks.

My original question: 6.3. Third, National Cancer Institute (NCI) combined BRFSS and National Health Interview Survey (NHIS) to generate county level small area estimates of current smoking outcomes by sex for the time periods
1997-1999 and 2000-2003 (http://sae.cancer.gov/estimates/current.html)[1]. It could be a good data source for an external validation of county-level model-based small area estimates of current smoking in this paper. The comparison between NCI estimates and the estimates from this paper may be able to evaluate the bias from the missing wireless-only population and even the population without phone in BRFSS.

Authors’ response: We thank the reviewer. We are aware of the NCI work (one of the co-authors of this manuscript received a joint statistical award for this work with NCI from the BRFSS side). We did not include this work in our manuscript since it was based on data in years where cell phones were still a relatively minor issue.

My comments: NCI’s estimates for 1997-1999 were based on the data with landline but were consist with the data (landline only) the small area model here used for 1996-2000. If the small area models are appropriate, we should expect both estimates are close.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.