We would like to thank the Editor and Reviewers for the thoughtful comments and suggestions. We have addressed these reviews and greatly improved our manuscript. Below please find our responses to the specific comments.

Comments from the reviewers:   
  
Reviewer #1: This is a study on a novel research question. Generally the study appears to be well conducted and reported. Particularly, Figure 2 is very clear and illustrated the results beautifully. Thank you! My methodological comments are:  
  
1. It appears the study used ecological case-cross over design, where the 'ecological' part is that the unit of analysis is based on county rather than individuals. If my interpretation is correct please be explicit about the study's ecological nature, and discuss the limitations arises from this, notably ecological fallacy and potentially individual-level time varying confounding.

This is a good point, thank you. We added the following sentences to first paragraph of the study limitations section, so it now reads:

In this study, we measured county-level power outages since no national finer-resolution power outage exposure data are available. This means our study was an ecological study, subject to ecological fallacy and potential individual-level time-varying confounding bias. It is possible that the relationship between power outage and hospitalizations that we observed at the county level does not hold at the individual level, despite its theoretical plausibility and community concerns about power outage and adverse health outcomes. At this time, no individual-level power outage data are available.

2. More details are required on the selection of 'control' days. It was reported as 'matching on county, day of week, and month'. In that case the control days would be at least 1 year apart from the 'case' days. How far apart are the case and control days are? Would it not be an issue that with year difference, demographic shift would play a non-negligible role?

Thank you for this comment. Multiple reviewers requested that we clarify how we selected control days. We have updated this paragraph and included some citations to justify our selection of control days:

For each case day (i.e., a county-day with a non-zero hospitalization count), we selected control days from the same county in the month and on the same day of the week. Since our study included only days from 2018, control days were also matched to the same year. This matching controlled for time-invariant confounders such as county-level socioeconomic factors that might influence both hospitalization and power outage rates, as well as seasonal and day-of-week effects. To check for overdispersion in the outcome, we ran a quasi-Poisson model but found no evidence of overdispersion.  
3. Consider splitting some text in Statistical Analysis into a separate section as 'Study design' for clarity.

Thank you for this suggestion. We have added a new section called ‘Study design’ which includes the following text:

We used a time-stratified case-crossover design with a conditional Poisson model (42) to analyze the association between daily county-level power outage exposure and CVD or respiratory hospitalization rates. We modeled CVD and respiratory hospitalizations separately because heat, cold, loss of power to medical devices, and dehydration may affect these outcomes differently (43–45). We selected control days for every case day (i.e., county-day with a non-zero hospitalization count) by matching on county, day of week, and month. This matching controlled for time-invariant confounders such as county-level socioeconomic characteristics, which could affect both hospitalization rates and power outage rates, as well as seasonal and day-of-week trends. We first ran a quasi-Poisson model to test for overdispersion of the outcome, but found no overdispersion.

4. Consider adding a negative control outcome (e.g. infectious disease?) - this could add confidence to the findings given the ecological nature.

Thank you for this suggestion. This is something we also considered when conceptualizing this paper. However, we have been unable to find a good negative control outcome for such an analysis that would add confidence in results. For negative control outcomes to add confidence to the results, the confounders that influence the primary outcome of interest and the negative control outcome must be the same and modelled correctly in both the primary analysis and negative control analysis. Our analysis could be confounded by severe weather and temperature, and so a negative control analysis could theoretically help us detect residual confounding from severe weather and temperature. We don’t believe that the relationships between temperature and infectious disease and temperature and CVD and respiratory hospitalizations are analogous enough that a negative control analysis would add to our results. For example, it would take a lot longer for temperature to affect infectious disease rates compared to cardiorespiratory hospitalizations. Infectious disease hospitalizations might be reduced due to severe weather in the same way that CVD and respiratory hospitalizations are, but daily temperature does not influence infectious disease hospitalizations in the same way that it influences CVD and respiratory hospitalizations, rather, seasonality has a stronger effect. Additionally, to correctly model the relationship between power outage and infectious disease, we’d need to use an entirely different modelling strategy to account for disease transmission. We would be open to conducting an analysis if we could find an appropriate negative control outcome.

Thank you for your comments Reviewer 1. We very much appreciate your feedback and contributions.

Reviewer #2: This paper describes the association between power outages and hospitalisations in the United States. I was asked for a methodological report and I interpret that to include all aspects of the design and conduct of the study.

Points of detail

Page 6 Using last observation carried forward seems rather arbitrary. My initial thought was that I would have tried to average some readings before and after the missing data period. If the missingness is related to power outages which occurred during the period of missingness imputing what happened before will be biased.

* We think that most missingness is random here, from utility website outages, or internet connectivity issues with the poweroutages.us not able to collect information
* Of course there is a chance that there are gaps in data reporting at the same time as large power outages, which would not be random. And so in that case yes, LOCF would be biased.
* However even if it IS biased…when we used LOCF to fill values where there were <4hrs missing, we only filled in less than .3% of observations. So we could average the values beforehand and after, but it seems unlikely that it would make any difference to the analysis.
* To fully address this comment, we’ve added a sensitivity analysis. We filled in these gaps with 0s, meaning no power outage, and then we also filled them in as if there had always been power outages during these times, to model the two extremes of how our effect estimates might change depending on how we chose to impute these values.
* Write about results here
* Give Vivian two datasets to rerun analyses with on FASSE.
* Decide on a cutoff for customers out that’s the threshold
* For the 4hrs missing, randomly add some customers out
* Draw customers out from a distribution around 1% with some variability around 0.05
* Rerun the main analyses and see how it changes things
* <4 hr chunks

Page 7 I assume the person had to be living in the same county on the control day as the power outage day. Can this be clarified?

The first reviewer also found our text about our selection of control days unclear. We clarified this with the following text:

For each case day (i.e., a county-day with a non-zero hospitalization count), we selected control days from the same county in the month and on the same day of the week. Since our study included only days from 2018, control days were also matched to the same year. This matching controlled for time-invariant confounders such as county-level socioeconomic factors that might influence both hospitalization and power outage rates, as well as seasonal and day-of-week effects. To check for overdispersion in the outcome, we ran a quasi-Poisson model but found no evidence of overdispersion.

Our analysis was at the county level, and we did pick control days (county-days) within county, day of week, and month strata. It’s possible that some people moved in or out of the counties in between case and control days, but because they are within the same month, movement was probably minimal.

Page 7 Could people enter the models twice if their county experienced more than one event? If they could how was this included in the models?

Our study was a case-crossover analysis, and we aggregated counts to the county level to use a conditional Poisson model for the analysis. This is statistically equivalent to doing a case-crossover analysis with a logistic regression model. Because it was a case-crossover, people could be in the models twice if they were hospitalized twice or more, and they might have also been included in the models if they were hospitalized on a selected control day. Their presence in the model was dependent on if they had the outcome rather than the exposure.

This method does account for county clustering, by including intercepts for county.

Pages 7 and 8 The authors have fitted slightly different models to cardiovascular disease (CVD) and to respiratory disease. I wonder whether this is optimal if there is going to be a comparison between the conditions.

We have indeed fit two different models for each outcome. Our goal in doing so was to correctly model the relationship between exposure, outcome, and confounders to best estimate the effect of power outage on CVD and respiratory outcomes. We hypothesized that power outage would cause cardiovascular and respiratory hospitalizations because of heat or cold exposure, loss of electricity to medical equipment, mobility devices and elevators, and refrigeration, water, and communication devices. There are different relationships between heat and cold, dehydration, and CVD and respiratory conditions. For example, the lagged effect of heat exposure on CVD is different than the lagged effect of heat on respiratory conditions. We chose to model these outcomes differently and separately because of these differing relationships and so we think it makes sense that the models are different. The other option would be modelling them the same. Not sure why we’d do that.

The third reviewer also asked us to address these differences in the discussion and so we’ve included the following text clarifying this:

We hypothesized that power outages may lead to CVD and respiratory hospitalizations in older adults due to increased heat exposure, cold exposure, stress, and loss of electricity to life-sustaining medical devices and mobility aids. Power outages may also cause changes in indoor air quality when dehumidifiers, air purifiers, and ventilation systems lose power. The lagged effects of CVD and respiratory hospitalizations due to power outage are likely different since the lagged effects of high or low temperature and air pollution differ for CVD versus respiratory disease (26–28).

Page 8 I am afraid I have very little knowledge about Medicare or Medicaid. Are these universal, that is, does everyone over age 65 become included?

Thank you for this comment. We can definitely clarify this both here and in the manuscript. Everyone in the US above 65 is covered by Medicare, however, people have the option to choose from two different kinds of Medicare coverage – Fee-For-Service, or Advantage. Our dataset only includes Fee-For-Service Medicare enrollees, which is only about half of the US population over 65, but is still 23 million older (65+) adults. More work is needed to figure out how Fee-For-Service enrollees differ from Medicare Advantage enrollees demographically (which we’re actually currently working on) but there have been a few non-academic publications about this. It seems Advantage enrollees may be more vulnerable to health effects from power outage compared to Fee-For-Service enrollees based on demographic differences reported in these white papers, such as income, education, measures of social risk, rates of chronic health conditions, and healthcare utilization. In summary, the people we’ve included here are likely less vulnerable to health effects from power outage compared to the other half of the population which is not represented. We’ve added a paragraph to the methods section about this population and how it compares to the overall population, and then a couple sentences to the discussion about how this may have influenced our results:

Methods:

Our study population included Medicare Fee-For-Service beneficiaries aged 65+ and enrolled for at least one month between January 1st, 2018, and December 31st, 2018. From the Medicare Beneficiary Summary File (MSBF), we obtained age, sex, Medicaid-eligibility status, and county of residence for all beneficiaries. Medicare Fee-For-Service beneficiaries make up about 50% of the US population age 65+, and may have fewer chronic health conditions, more education and income, and less healthcare utilization compared to Medicare Advantage beneficiaries.

Discussion:

“Next, power outages may be more common in high social vulnerability communities (5). If power outage exposure in our study was more common among more vulnerable older adults, the effects estimated here may be overestimates of the population-level effect. At the same time, our study population was comprised of Medicare Fee-For-Service beneficiaries, who may be less vulnerable to health effects from power outages due to higher SES and fewer chronic health conditions compared to Medicare Advantage beneficiaries, which may have led to underestimation of the population-level effect.”

Page 9 A mean of 7 with a standard deviation of 29 is crying out for expression also as quantiles given the obvious skew. I would suggests 5% and 95% as well as the usual quartiles.

Thank you, this point was bought up by another reviewer as well. We have changed this to report the IQR and 5th and 95th quantiles.

Include that here when we do it.

Page 10 Looking at Figure 2 and the confidence intervals there I find it difficult to agree about the effect at 6 days for CVD. Unless there is a strong theoretical argument for supposing that the effect returns after six days I would focus on the earlier period where the confidence intervals do not include the null unlike 6 days.

Thank you for bringing this up. There is some evidence in the literature of week-later delayed effects of heat on CVD hospitalizations, but we were already on the fence about how to interpret that last lag even before this comment. We think it’s fair to interpret it as no effect. We have removed that interpretation from our manuscript and it now reads:

In our main analysis we used a case-crossover design with a conditional Poisson model to test the association between 8+ hour power outage and emergency CVD hospitalization rates up to 1 week after power outage exposure. We found increases in CVD-related hospitalizations 1-3 after power outage exposure (**Figure 2**).

Page 10 Is the statement comparing the outage sizes based on a formal test? I do not see this as too important since the pattern of results over the early days and the dose–response effect are quite compelling.

Thank you for bringing this up, we think it’s a very good point. No, these comparisons are not based on a test. They are only based on the results from our models. It is true that the confidence intervals for the smaller vs. larger outages do overlap a lot, so it may not be fair to say that the effect is significantly higher with the larger outages. We’d be open to rephrasing this.

Page 11 A really minor point but I would call out Figure 2 immediately after the first respiratory results (as done for CVD).

Thank you – we added this.

Supplementary Figure 2 This seems to be to reinforce my comments about 6 days for CVD. Incidentally there is a typo cardiovascular not cardiovasular.

Thank you – we fixed this typo.

Points of more substance

Generalisability

I had to read page 9 several times to convince myself that I had understood just how many power outages there are in the United States. The map was helpful here. For context I cannot remember of the several decades of my adult life ever experiencing an outage of 8+ hours and I do not think that is a memory problem. My country does experience some long power outages but usually in only in rural areas in very severe winter weather. In an international context what is the authors’ message to our governments? Is it ‘This could happen to you?’.

Thank you for bringing our attention to this paragraph. Though I’m not sure we have a message for the international context, this sentence definitely needs clarifying. Counties were considered exposed to power outage if >1% of the population was without power for 8 consecutive hours or more. This means that only a very small proportion of the population actually experienced an 8+ hour outage. We have changed the text to clarify this:

We included 2,161 US counties, covering 71.1% of older adult Medicare Fee-For-Service beneficiaries age 65+ (N = 23,645,101). These counties experienced an average of 7 (standard deviation, SD=29) 8+ hour power outages affecting ≥1% of customers in 2018 (**Figure 1**). Because counties were considered exposed to an 8+ hour power outage if ≥1% of customers were without power for 8+ hours, it was not the case that individuals experienced an average of 7 8+ hour outages; rather, a small portion of customers in these counties were affected by 8+ hour outages.

DME use

There appears to be no mention in the text of effect modification for CVD To be fair Figure 3 does present a display. At the very least it should be discussed in the main text. This seems to me an opportunity for using CVD as a negative control.

In our results at the end of page 13, we had previously written:

For DME use, the association between power outages and respiratory hospitalizations appeared stronger in counties with smaller percentages of DME users (quartile 1 of DME use) compared to counties with larger percentages of DME users (quartile 4 of DME use). Respiratory hospitalizations remained elevated in counties with quartile 1 DME use for two days after power outage, while in counties with fourth quartile DME use, hospitalizations were elevated only on the day of power outage (**Figure 3**).

To directly address CVD rather than respiratory hospitalizations, we have added this sentence to the end of that paragraph:

“The effect of power outage on CVD hospitalizations was not modified by DME use.”

It’s not clear to us what it would mean to use CVD, one of our main outcomes, as a negative control. Could we clarify what the reviewer meant here? We did run the effect modification analyses for CVD as well.

For a discussion of use of negative control outcomes in our analysis, please see our response in comment 4.

Summary

Mostly points for clarification about the analysis. I have no issues with the core strategy.

Michael Dewey

Thank you for your clear and helpful review Dr. Dewey! We sincerely appreciate your feedback.

Reviewer #3: This study investigates the association between power outages and emergency hospitalizations for cardiovascular and respiratory diseases among older adults in the US. The authors assembled the first nationwide dataset of power outage exposure and linked it with Medicare claims data from 23 million beneficiaries aged 65 and older at county level. Using a case-crossover design with a conditional Poisson model, they analyzed the lagged effects (up to 1 week) of outages on hospitalization rates. Overall, the study addresses a critical gap in understanding the health impacts of power outages in the US, which are becoming more frequent due to climate change. However, I have several concerns regarding exposure assessment, confounding, and result interpretation.

(1) This study does not have exposure data at a fine spatial resolution such as census tract or ZIP code level. Power outage exposure is assessed at the county level, meaning many individuals within a county may not have actually experienced a power outage. The definition of exposure (≥1% of customers without power for 8+ hours) may have led to substantial exposure misclassification.

Thank you for this comment. Yes, this is true, and we share your concern. We agree that it would be great if we could assess exposure at a smaller spatial unit and not deal with this issue. Unfortunately, the best data available for the US are at the county level and no finer-grain data are available. In order to determine when a county was exposed to power outage, we needed to pick some threshold after which to consider a county exposed.

Because of the inherent exposure misclassification involved in our exposure assessment strategy, we also conducted secondary analyses where we used two other definitions of power outage exposure where we considered counties exposed if ≥3 and ≥5% of customers were without power for 8+ hours. We found larger effect sizes when we increased the threshold, which is what you’d expect for non-differential exposure misclassification.

We’ve changed the text in the statistical analysis section to highlight the reason for our secondary analyses of these larger power outages, and address this exposure misclassification:

In our primary analysis, we considered a county-day exposed to power outage if ≥1% of county customers were without power for 8+ consecutive hours, a definition that may have substantial exposure misclassification (up to 99% of customers may be unexposed). To address this, we conducted a secondary analysis where we assessed power outage exposure based on higher thresholds of customers (≥3%, ≥5%), scenarios with less exposure misclassification.

We’ve also included text on this issue in the study limitations:

In this study, we measured county-level power outages since no national finer-resolution power outage exposure data are available. In our main analysis, we considered a county-day exposed to power outage if ≥1% of county customers were without power for 8+ consecutive hours, a definition that may have substantial exposure misclassification (up to 99% of customers may be unexposed). This misclassification likely biased study results towards the null, but the magnitude of bias remains unknown. When we assessed power outage exposure based on higher thresholds of customers (≥3%, ≥5%), scenarios with less exposure misclassification, we observed larger effect estimates. Future studies could collaborate with utilities to obtain finer-resolution power outage data or use satellite imagery to identify exact outage boundaries of long-duration outages lasting into the night (36, 37) to address these exposure assessment issues.

(2) The [PowerOutage.us](http://PowerOutage.us) (POUS) dataset is incomplete, with many counties missing up to 50% of data. The authors used a simple method to impute county-days with missing 4 or fewer hours, but it is unclear how the remaining missing values were handled. Further discussion is needed on how missing data may have biased the results.

Thank you for this good point. We were also concerned about the effects of missing data on our results before we began this project, and so we published a study where we conducted simulations to test the effect of missing data on the results of this paper, which we attempted to include with our paper submitted here. It was not published at the time we submitted this paper, but it is now, and so may not have made it to this reviewer. It discusses in detail our strategy for handling missing data and the impacts that this strategy has on the results of our study:

<https://journals.lww.com/environepidem/fulltext/2025/08000/assessing_bias_in_measuring_power_outage_exposure.4.aspx>

In our methods section we referenced this study, and explained how we handled missing data and why:

Substantial exposure data were missing from the POUS dataset. The POUS dataset was created using web scraping, and some utilities did not have websites, or their websites were offline during part or all of the study period. Previously, we conducted a simulation study where we treated missing data as no power outage exposure, because no exposure was the median value in the dataset (38). Under this assumption, missing data biased results of a study like ours towards the null. When a small percentage (~ <15%) of hours within each county (county-hours) were missing exposure data, bias was minimal (37). When larger amounts (~ >50%) were missing, bias was substantial. To balance generalizability and bias, we excluded counties with >50% of county-hours missing in the POUS data (n=907 counties). On average, the remaining counties were missing data for 7% of county-hours. When included counties were missing 4 or fewer hours of consecutive exposure data, we carried forward the last observation to impute those hours. Our final analytic dataset included counties in 48 states (all counties in HI and AK were excluded) and 2,161 counties, covering 71.1% of 2018 Medicare beneficiaries (N = 23,622,770).

And we also referenced our simulation study in the study limitations section:

Finally, the POUS dataset we used to assess exposure is missing substantial data. We excluded counties missing more than 50% of customer-hours in 2018 to balance generalizability and bias from missing data based on our prior simulation study (38).

We also have done an additional sensitivity analysis where we only include counties with less than 20% of data missing, in response to your comment below, which we discuss more below.

(3) The study uses different sources for the number of affected customers and the total number of customers. This inconsistency may introduce bias in estimating the proportion of the population affected by power outages. The authors should justify this approach or explore alternative sources that ensure consistency.

We had previously written the following about our calculation of electrical customers served by county:

To assess power outage exposure for 2018, we used PowerOutage.us (POUS) (24) nationwide county-level data. These data included the number of customers without power every hour by county. ‘Customers’ referred to residential consumers, such as households or families, and non-residential consumers, such as businesses. Because county-level estimates of customers served from POUS were unreliable, we used Energy Information Administration (EIA) estimates of customers served by state (40) and 2013-2018 American Community Survey estimates of the number of households and establishments by county to determine the proportion of state customers in each county. We then allocated state-level customers to each county, estimating the number of customers served in each county.

We have added additional justification about why the POUS customer counts are extra bad:

Customer counts reported in POUS are often unreliable, as utilities rarely provide this data. POUS captures it when available, but inconsistencies and gaps mean estimates provided by POUS are often orders of magnitude different across months and years. Instead of relying on these unstable estimates, we used Energy Information Administration (EIA) estimates of customers served by state (40) and 2013-2018 American Community Survey estimates of the number of households and establishments by county to determine the proportion of state customers in each county. We then allocated state-level customers to each county, estimating the number of customers served in each county.

(4) If vulnerable individuals (e.g., people of lower SES) are overrepresented among the exposed group, while healthier individuals are not affected, the population-level effect may be overestimated. The authors should clarify how they accounted for this potential bias.

Thank you for bringing up this important point. This is something we considered when conducting this analysis. Our group conducted a study using the same power outage exposure data we use here, and we did find that lower SES communities were likely more exposed to power outage. So, it is likely true that the effect we estimate here is influenced by that fact. We’ve included the following sentence in our limitations which cites our study showing that vulnerable communities are likely more exposed:

“Next, power outages may be more common in high social vulnerability communities (5). If power outage exposure in our study was more common among more vulnerable older adults, the effects estimated here may be overestimates of the population-level effect.”

(5) Based on Figure 1, most excluded counties are in the Midwest and Western regions, raising concerns about regional representativeness. Please consider focusing on regions with more complete data (e.g., South and Northeast) and assessing whether included and excluded counties are comparable in main characteristics.

* We’re not quite sure what the reviewer means by ‘focus on regions with more complete data’
* We think they could be suggesting that we do a sensitivity analysis where we exclude regions with missing data and look only at the effect estimate in those regions with data
* However, health-related characteristics vary so much between regions of the US that if the effect estimates of such a sensitivity analysis are different, it wouldn’t really tell us anything about whether missing data was influencing those results, or if it was just regional differences in the effect
* There’s no reason to think that the effect would be the same across regions
* Another way we could interpret this comment is a request to compare county characteristics between regions that have less missing data and more missing data. However there will certainly be differences, so we’re not sure what that would accomplish.
* We’ve done the following to address this comment. We chose a few county characteristics that we think might modify the effect of power outage on hospitalizations in older adults – rurality, income, and home ownership
* Within the 9 climate regions of the US, we’ve compared these characteristics among excluded vs included counties.
* Within 9 climate regions, pick a few characteristics (rurality, % in poverty, svi)
* Look at within region how similar are the included vs excluded counties are
* And then add a sentence in the limitations
* Know that health-related char vary across the county
* Know they are different
* Even if we do a sens analysis and it showed that the results were different it wouldn’t tell us anything
* We agree that missingness is a concern
* But there are many other differences across regions (reasons for PO, healthcare access, population differences)
* We do agree that this is a problem, but if we stratify and the results are different

Hm.

(6) The Poisson regression model was adjusted for temperature, precipitation, and wind speed, but it did not account for air pollution and other extreme weather events, such as wildfires and ice storms, which may influence both power outages and hospitalizations. I suggest adjusting for air pollution and a more comprehensive set of meteorological covariates including extreme weather events in sensitivity analyses.

* Check cooccurrence btw wildfire and power outage
* And then add a sensitivity analysis with control for wildfire
* Control for wildfire-generated PM with marissa’s data
* Power outages can impact air pollution levels – travel less or more, gas generators,
* Mediator if anything
* Can’t think of cause or association of air pollution with power outage other than weather (wind or wildfire etc.) but we adjusted for weather

Air pollution isn’t a confounder.

We could control for wildfires but like idk most power outages are not wf related.

Ice storms should be controlled for with the variables that we did.

Other types of AP are mediators

(7) This study did not distinguish between planned and unplanned power outages. Planned outages may have less severe health effects as residents may be better prepared. If possible, the authors should analyze them separately.

This is a very important point – planned and unplanned outages do likely have very different effects on health. Unfortunately, there is no data available that allows us to distinguish between the two at the moment. We have included the following sentence in the limitations section addressing this issue:

“Further, we were unable to distinguish between planned and unplanned power outages, which may have different effects on hospitalizations.”

In CA we have access to planned power outage events and we have ongoing analyses related to their impact on health. We agree that it’s important.

Planned ones could be worse because they affect more people?

(8) The analyses were conducted at the county level given the exposure assessment method. It would be more appropriate to explore effect modification by county-level characteristics such as urban/rural, socioeconomic status, and geographic region. Since the temperature and season data are readily available, please consider implementing the effect modification by temperature and season in the current study.

We considered doing this and we’re happy to include these analyses.

Yep ok.

Above 95th percentile of temperature or below the 5th percentile

Subset those out

Best option available now is to stratify on day of temp

Fine to use same-day temp

Urban/rural collapse the ruca codes into two categories

(9) Since the approach to handling missing data is critical to the results, a sensitivity analysis should be conducted to assess the robustness of the findings when removing data with varying proportions of missingness.

I think we can do this. We can just remove even more missing data, and then redo the analyses. We’ll have sensitivity analyses coming out of our ears.

Remove counties missing more than 20% and get a stronger effect.   
  
Minor comments:  
(1) Please consider including the results for qAIC in supplementary materials for completeness.

Yes, thank you - we’ve added the qAICs to the supplement.

(2) In the main analysis results for CVD hospitalizations, the authors report increases in CVD-related hospitalizations 1-3 and 6 days after power outage exposure. However, Figure 2 suggests that significant increases are only observed on lag days 1 and 2. Additionally, Figure 2 indicates that power outages with a size ≥1% appear to have protective effects on CVD-related hospitalizations on lag days 4 and 5 (a similar pattern is observed in Supplemental Figure 1). Could the authors provide an explanation for these findings?

Thank you for bringing this up. There is some evidence in the literature of week-later delayed effects of heat on CVD hospitalizations, which is why we interpreted lag 6 as positive, but we were already on the fence about how to interpret that last lag even before comments from both this reviewer and Reviewer 1 about lag 6. We think it’s fair to interpret it as no effect. We have removed that interpretation from our manuscript and it now reads:

In our main analysis we used a case-crossover design with a conditional Poisson model to test the association between 8+ hour power outage and emergency CVD hospitalization rates up to 1 week after power outage exposure. We found increases in CVD-related hospitalizations 1-3 after power outage exposure (**Figure 2**).

The observed protective effect on hospitalizations 4–5 days after a power outage may be a harvesting effect. For larger outages, there seems to be no impact on hospitalizations, making it unclear to us whether any true protective effect exists. We’ve added the following text explaining our interpretation of these findings to the discussion:

The observed protective effect on hospitalizations 4–5 days after a power outage may be a harvesting effect. For larger outages, there seems to be no impact on hospitalizations, making it unclear whether any true protective effect exists.

(3) The discussion could include an explanation for the differing lag effects observed for CVD and respiratory-related hospitalizations.

Yes – we chose to analyze these two outcomes separately because we hypothesized that the effects of power outage on CVD and respiratory hospitalizations may be different. We’ve included the following text in the discussion:

We hypothesized that power outages may lead to CVD and respiratory hospitalizations in older adults due to increased heat exposure, cold exposure, stress, and loss of electricity to life-sustaining medical devices and mobility aids. Power outages may also cause changes in indoor air quality when dehumidifiers, air purifiers, and ventilation systems lose power. The lagged effects of CVD and respiratory hospitalizations due to power outage are likely different since the lagged effects of high or low temperature and air pollution differ for CVD versus respiratory disease (26–28).

Thank you for your comments Reviewer 3!

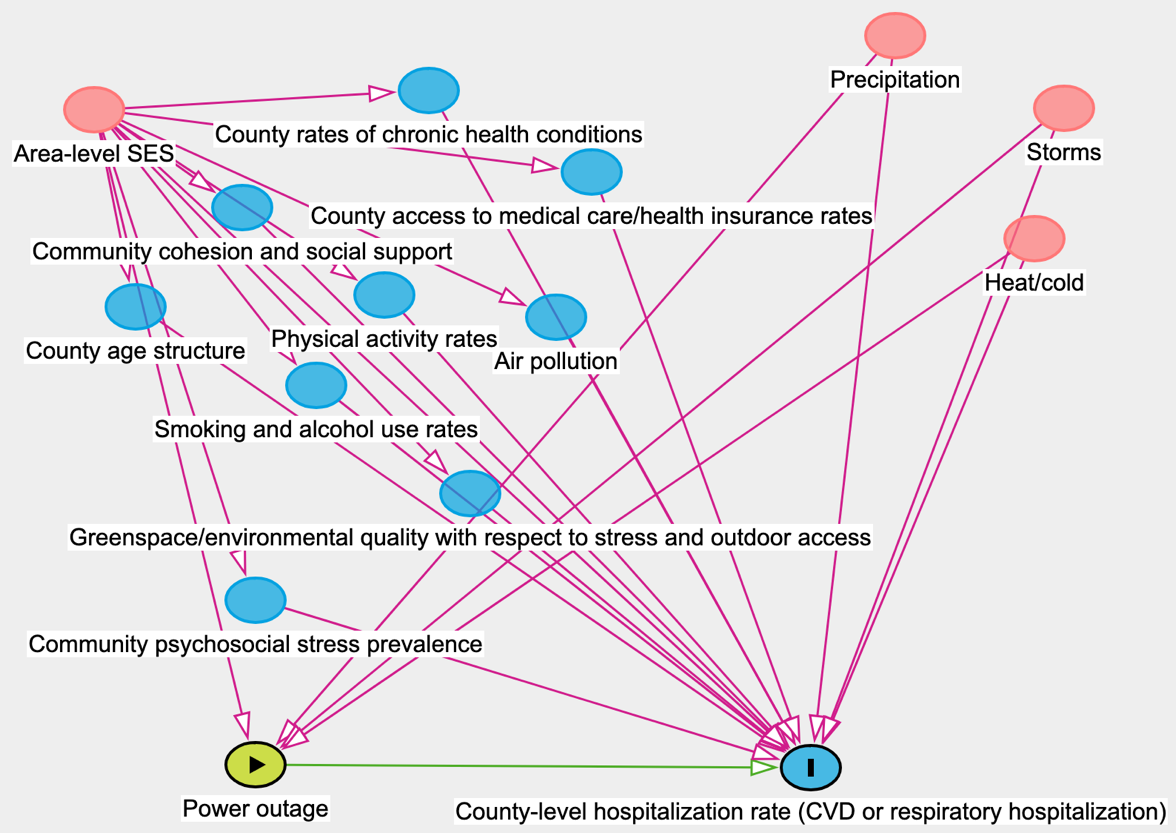
Editor comments:

1. Agree with R2 about improving the description of the total number of outages (7 outages, SD=29, of 8+ hours). The huge right skew renders mean/SD uninterpretable, so median/IQR is needed.

Yup we fixed that sorry.   
  
2. Also, even if removing the mean/SD, as R2 points out I just think the phrasing of this statistic is not easy for readers: "counties experienced an average of 7 (standard deviation, SD=29) 8+ hour power outages affecting ≥1% of customers in 2018." I have lived in rural U.S., and I had maybe 1 of these outages per year. So it does not resonate with me. The issue is definitional -- the reader's mind automatically interprets this to mean that each individual in the U.S> experiences 8 outages per year. But the 1% bit of the statistic is doing a lot of work here -- it means that only a tiny number of households, different each time, are affected. The more useful statistic is the median + IQR of power outages experienced by a given person/household per year.

Thanks for highlighting this point again. What you said makes a lot of sense. We also wanted to calculate the statistic that you mention at the end of that sentence, but we can’t calculate that from the data we have available. We did change what we’ve written to give the reader a more correct interpretation of what this means:

We included 2,161 US counties, covering 71.1% of older adult Medicare Fee-For-Service beneficiaries age 65+ (N = 23,645,101). These counties experienced an average of 7 (standard deviation, SD=29) 8+ hour power outages affecting ≥1% of customers in 2018 (**Figure 1**). Because counties were considered exposed to an 8+ hour power outage if ≥1% of customers were without power for 8+ hours, it was not the case that individuals experienced an average of 7 8+ hour outages; rather, a small portion of customers in these counties were affected by 8+ hour outages.   
3. Would a DAG be helpful in illuminating the assumptions in this causal inference paper? In particular, being transparent about which confounders are and are not controlled for?

Yes – we can make a DAG! Here it is: 

So we controlled for wind speed, temperature, and precipitation, which we think should capture severe weather and its impact on hospitalizations. Those are the variables in the top right. Our case-crossover study design automatically controls for the county-level characteristics in the top-left corner that could also confound the relationship of interest.

We added this to the supplemental content.   
  
4. Could the authors describe their results in absolute terms? The Poisson output of IRR is useful as a relative measure. But ultimately, it is hard to get a sense for the absolute impact as all of the IRRs are tiny even if statistically significant.

This is a great point. Need to use the data to get that though.

Thank you Editor!

Skip urban/rural

TODO for Vivian:

* First thing is rerun fasse\_data\_prep/03\_urgent\_hosp\_combine\_to\_analytic.R and run the part at the bottom, where it says descriptive stats and take a screenshot of that output. Need the median, IQR, 5th, and 95th percentiles of those outages
* Second thing: rerun the data cleaning pipeline (all scripts in fasse\_data\_prep) with the dataset containing hot/cold days (replace 'analytic\_exposure\_data\_2018.parquet' in 03\_urgent\_hosp with analytic\_exposure\_hot\_cold\_urban\_rural\_smoke.parquet) then run the hot\_cold\_analysis.R script, the urban\_rural\_analysis.R script, and the sens\_control\_for\_wf\_smokel.R script.
* Third thing: run these lines in 03\_urgent\_hosp\_combine…
* an\_dat\_low\_missingness <- an\_dat %>% filter(percent\_served >= 0.8 &
* !is.na(percent\_served))
* # VIVIAN
* # Need to know how many counties have >80% of data
* Need to know how many counties have high coverage
* And then run the new second half of the script that cleans a high coverage version of the dataset (below the section marker ‘high coverage’)
* And then run the analysis script that runs the low missingness analysis low\_missingness\_sens.R
* Fourth thing: need to find and add the qAICs to the supplement. Very hard to do this out of the server lol. Not sure if we need to rerun models, or if they’re saved, or what. I can’t look without logging on.
* Fifth thing: need to make a list of counties that are included/excluded in the main analysis. I added a line in 03\_urgent\_hosp that should do that.