We thank the reviewers for reading our paper and for their comments. We found them very helpful, and we believe that they will improve our paper and make it clearer.

Reviewer #1: This is a time-series study assessing the association between wildfire exposures and health care usages among a Southern California population who use durable medical equipment.

Major Comments:

1. Why the study population was limited to population age 45 and older? Have the authors considered to perform modification analysis by different age group?

We aimed to assess the effects of wildfire exposure on people whose durable medical equipment use might indicate that they were medically or socially vulnerable. When people rent DME, they are renting Bilevel Positive Airway Pressure (BiPAP) machines, enteral feeding machines, hospital beds, infusion pumps, oxygen equipment, suction pumps, ventilators, and wheelchairs, all of which assist with medical conditions indicating some level of disability. We wanted to examine the effects of wildfire in this potentially vulnerable population. However, DME rental statistics also include people who rent breast pumps, which may not indicate vulnerability and may in fact indicate a certain level of health. Excluding DME using people under 45 excludes nearly all people renting breast pumps, leaving those we hypothesize as the most vulnerable to wildfire exposure. We consider older age to be part of this vulnerability. We have characterized this population in Casey et al. 2021.

Unfortunately, we weren’t able to get the additional demographic data that would make modification analyses possible due to resource limitations, so we haven’t performed any modification analyses.

We added additional sentences to the description of the study population:

*“We used electronic health record data from KPSC to identify all individuals who were 45 or older as of October 28th, 2019 and had rented DME in the year prior. We excluded younger DME renters in order to focus on socially and medically vulnerable older adults, but also to exclude breast pump users, a healthy subgroup of the otherwise vulnerable DME using population, who we did not hypothesize to be particularly vulnerable to wildfire exposure.”*

*Page 3 of the main text, under “Study population”.*

1. The reviewer understand that this is not a traditional cohort study. However, could the author present a traditional Table 1-like summary table to present the study population characteristics for those who had at least one event during the study period? This would be very useful to compare and contrast your study with others' for similar/different populations.

Unfortunately, as we mentioned above, resource limitations preclude us from getting detailed demographic variables in this dataset. We did describe the larger DME population (not only those who had an event) in Casey et al. 2021, if that helps.

1. How DME usage in year prior is an indicator of current vulnerability? Are there any previous studies assessed the proportion of patients who stopped DME usage within one-year of usage initiation?

We selected this population because we anticipate that people using DME are more vulnerable to wildfire exposures than people who don’t use DME. We aren’t interested in whether healthcare visits made by DME users are for the functioning of DME, or are made for conditions associated with DME, rather, we hypothesize that DME may be an indicator of social and medical vulnerability. We think DME may indicate vulnerability because DME is usually rented to address a disability or medical issue and is associated with some other indicators of social and medical vulnerability, such as older age and Medicaid insurance (Casey et al, 2021).

The average length of DME rental in our population was a year (Casey et al 2021). We aren’t sure that people were renting DME at the time of their healthcare visit, but again, we are using DME rental as a proxy for potential vulnerability, and we therefore think that even if someone wasn’t using DME at the time of exposure, we would still be capturing people who are potentially vulnerable. Again, we are not interested in healthcare visits specifically for DME equipment or conditions that the equipment is rented to manage.

We describe here why DME users may be vulnerable – in short, because they are more likely to have health conditions that may be exacerbated by wildfire exposure, may have less mobility than people who don’t use DME, and may be older than people who don’t use DME (a risk factor in itself).

*People who use durable medical equipment may be particularly vulnerable to both wildfire PM2.5 exposure and stress from wildfire proximity or evacuation. DME use is common among older adults and is associated with respiratory illness and other disabilities.29 A prior study among Kaiser Permanente Southern California (KPSC) members found increasing prevalence of DME rentals from 2008-2018 and the highest prevalence of use among older adults.30 DME types included bilevel positive airway pressure (BiPAP) machines, enteral feeding machines, infusion pumps, oxygen equipment, suction pumps, ventilators, and wheelchairs.30*

*This group may face unique challenges during wildfire events. Prior studies have found elevated effect estimates between wildfire smoke exposure and respiratory and cardiovascular disease outcomes among older adults compared to younger populations.16,31 Further, people using DME may have co-occurring medical conditions such as cardiovascular disease that make them more vulnerable to both the effects of wildfire PM2.5 and wildfire-related stressors beyond wildfire smoke (e.g., threatened or actual evacuation). Limited mobility or need for electricity access may result in increased difficulty evacuating disaster zones.30,32*

*Page 2 of main text, in “Introduction”.*

1. The quality of the wildfire PM2.5 is not guaranteed with a published peer-reviewed manuscript. Could the author provides more information to assure the quality of their wildfire PM2.5 estimates which is the key to avoid exposure misclassification.

ADD TARIK’S REPSONSE.

1. The quality of fire proximity exposure assessment is also not assured. Please clarify what are "final fire perimeters". Those fire perimeters were "final" in relevant to what? How fire activity is defined? For the 20km buffer, have the authors performed any sensitivity analysis on other buffer size? If not, sensitivity analysis is strongly recommended.

Fire activity is logged by CALFIRE in their database where they create and update geographic fire perimeter files. They timestamp these fire perimeters, so there are usually several perimeters associated with several dates during which a given fire was burning. Usually, fires expand and the perimeters get progressively bigger.

CALFIRE does not always record a perimeter every day, so there are sometimes no boundaries or only one boundary for a given fire. There are several recorded boundaries for the Getty fire, and one for the Woolsey fire. We used the last recorded fire perimeters, which we called final fire perimeters, which represented the largest burned area in the CALFIRE dataset.

We have revised the description to make it clearer:

*We obtained shapefiles of the total areas burned during the Getty and Woolsey fires from the CALFIRE Fire and Resource Assessment Program.39 Fire boundaries expanded while the fires were active, but fire perimeters recorded during the fires did not differ significantly from the total burned areas of either fire, since dynamic boundary data available did not include perimeters from very early in either fire.40 We therefore used last recorded fire perimeters in the CALFIRE dataset to define exposure. We considered ZCTAs exposed if they were within 20km of the last recorded fire perimeter on days that a fire was active. We hypothesized that living within 20km of a fire perimeter could elicit a stress response, similar to effects described in previous studies*.7,8,41

*Page 4 of the main text, under ‘Proximity to wildfire’.*

TODO: sensitivity analyses on the buffer, and interpretation of results.

1. There appears to be problem with zero-inflated data. However, based on the reviewers understanding, negative binomial model should solve this problem. Please explain why the authors still aggregated to ZCTA groups to solve this issue? Have the authors considered other statistical models (i.e. two-stage negative binomial model by ZCTA, conditional Poisson model)?

Though the negative binomial model may have still produced relatively unbiased estimates of association with reasonable confidence intervals on zero-inflated data, because most of the observed daily visit counts in all categories were 0, we wanted to increase our power and get closer to meeting negative binomial distributional assumptions by aggregating to the weekly level. We would have aggregated visit counts to the weekly level in both the proximity and PM 2.5 analyses, but we had reason to believe (from the literature on PM2.5) that there may be lagged effects of PM 2.5 exposure only detectable at the daily level, so we chose to aggregate to larger spatial groupings instead, to keep measurements at the daily level. We have edited the manuscript to reflect the importance of statistical power in our decision:

*Daily visit counts by ZCTA were low and often zero (median outpatient visits = 1, IQR = 3, median ED and inpatient visits = 0, IQR = 0). For the wildfire PM2.5 analyses, to avoid zero-inflation in our models, and to increase statistical power, we could have aggregated ZCTA counts to the weekly level. However, prior studies of wildfire smoke exposure have found associations between same-day air pollution and healthcare visits over the course of the following week.18–21 To evaluate a lagged effect in our data, we required daily healthcare visit counts, therefore, we opted to aggregate our data into higher-level spatial groupings of several ZCTAs based on spatial proximity (hereafter ‘ZCTA groupings’; grouping method described in the eAppendix).*

*Main text page 4.*

1. For the proximity and evacuation analysis, the analysis is done on the weekly level, which could be problematic as the study participants would have already left the affect area and use health care in other locations within a week, and this potential movement could lead to biased effect estimates (likely non-differential bias). Have the authors looked at the movement pattern right after a big fire and/or evacuation order. Also, have the authors tried to perform the analysis on the daily level?

We are using KPSC administrative data, which means that even if the study participants left the affected area, and sought care at another KPSC location, they would still be included in our dataset. It is absolutely true, however, that if they did not seek care within the KPSC system, they would not be included in the dataset. However, participants would be motivated to get care at a KPSC location because they are insured there. We do discuss this as a limitation in our limitations section. However, additionally, we consider evacuation to be part of our exposure/an effect modifier. If the participants had left the affected area and sought care, we would want to capture that.

We do not know the movement pattern before big fires or during evacuation orders, and this is an open research question that we’re interested in, and has been explored in the literature.

Comment?

<https://www.sciencedirect.com/science/article/abs/pii/S136192092200102X>

<https://escholarship.org/uc/item/5w85z07g>

<https://www.cambridge.org/core/journals/disaster-medicine-and-public-health-preparedness/article/abs/prevalence-of-unmet-health-care-needs-and-description-of-health-careseeking-behavior-among-displaced-people-after-the-2007-california-wildfires/1571035D1E31DF828A4500B2BBA0EAC0>

We could have performed the analysis at the daily level, but we opted to analyze at the weekly level for three reasons. First, as stated above, though the negative binomial models we are using are reasonable when distributional assumptions are violated and there is zero-inflation, aggregating our visit counts still increases our power to detect associations, since most of our daily visit counts were 0. Second, we do not have detailed information about where fire boundaries were located at the daily level – for the two-week Woolsey Fire, only one fire boundary was reported in CALFIRE. We didn’t have the detailed daily exposure information to do analyses at the daily level. Third, we had no reason to believe from the existing literature that there would be lagged effects that we could only disentangle at the daily level. We have edited the manuscript to say:

*For proximity and evacuation analyses, we used ZCTA level daily visit counts aggregated to the weekly level. Because our exposure data was not as granular as that in the PM2.5 analyses, as we used last recorded fire boundaries and last recorded evacuation zones rather than daily data, we evaluated relationships at the weekly level. This aggregation also removed weekend-weekday patterns in outpatient visits, increased power, and reduced zero inflation. We considered a week exposed if the Woolsey or Getty fire burned any day that week.*

*Page 5, ‘Outcome definition’.*

1. The author controlled for ZCTA grouping level covariates like income, home ownership, poverty, age structure, etc. Are those covariates varied over time? If they are not, by this study design, those variables should have been controlled automatically.

We agree with the reviewer that these non-time-varying covariates are controlled for by design. We include them because they can increase the precision of association estimates in the absence of a random intercept controlling for spatial unit, which we couldn’t include because of computational limitations.

ADD CITATION, maybe edit text?

1. There is concern for temporality (outcome may occur preceding the exposure) for the same week (week 0) wildfire PM2.5 exposure analysis.

Yes – thank you for this comment, we missed this. We will exclude week 0 from this sensitivity analysis.

TODO: do this. When we are excluding, do we exclude it from the analysis itself, or just from the reporting?

1. For the results of proximity and evacuation exposures, most of the RRs are null with 95% CI including the null value. Could the author please acknowledge this while reporting and interpreting the results?

We added acknowledgement of this:

### *Woolsey Fire proximity and evacuation exposure*

*During the Woolsey Fire, all visit types increased across the study area, regardless of fire or evacuation exposure. Woolsey Fire proximity during the fire was significantly associated with decreased outpatient visits, and associated with increased inpatient admissions for cardiorespiratory disease, though not significantly (Figure 3). We observed similar associations between Woolsey Fire evacuation exposure and healthcare visits. Evacuation from the Woolsey Fire was also significantly associated with increased inpatient admissions for cardiorespiratory disease, and associated with decreased outpatient visits, though the confidence interval included the null (Figure 3).*

### *Getty Fire proximity and evacuation exposure*

*During the Getty Fire, outpatient visits, ED visits, and ED visits for cardiorespiratory problems increased across the entire study area. We observed reduced risks of all visits types among proximity exposed ZCTAs during the Getty Fire, but confidence intervals were very wide and included the null (Figure 3). We observed similar, if somewhat attenuated, associations among evacuation exposed ZCTAs. Confidence intervals also included the null.*

*None of our results were sensitive to spline flexibility.*

*Page 8.*

Minor Comments:

1. There appears to be discrepancies between what is reported in the main text vs. the abstract. Could the author clarify where are those RRs in "proximity RR =1.48, 95% CI:1.01, 2.17, evacuation RR = 1.76, 95% CI: 1.02, 3.05)" presented in the main text? To the reviewer, no RR reported in Figure 3 has a upper 95% CI limit that over 3.

The RRs reported in the figure are actually log RRs, and we made an error and did not include ‘log RR’ in the axes titles. We apologize and have corrected this error – see the next minor comment, Minor Comment 2, for an updated Figure 3.

1. Figure 3 is confusing with contradictory x-axis label and figure legends. Please consider create separate panels for proximity vs. evacuation analyses and specific health care usage outcomes.

TODO: separate this into 2 figures.

1. Reference 38 is not completed.

This was a citation manager issue – thank you for pointing it out. We have fixed it!

Reviewer #2: Summary:

This paper evaluates whether people who use durable medical equipment have increased inpatient, outpatient, or ED visits related to exposure to PM2.5 during wildfire events or if they live in proximity to the wildfire or its evacuation zone. In general, this is a very interesting take on understanding the health impacts of wildfire by focusing on a group who may be more affected, however, the paper could use some editing to keep the reader clearer about what was done and what the findings are and who was impacted.

One general point throughout the paper is that it should be made clear that this study is only focused on the durable medical equipment (DME) user population and not comparing this group to those who do not use durable medical equipment. We cannot say, from this study, that the DME population is more affected by wildfires but can just say that some, but not all outcomes, were associated with some, but not all, of these exposures among this population. In the results and discussion sections, many of the results are not stated clearly that it is just among the DME population. A reader could take these sentences out of context and infer these findings apply to all people in the study area, which is not the case.

Similarly, there are many other places that require more precise language, as noted in the specific points below.

Another main point is that in the introduction and the discussion, I do not think that the authors have done a sufficient review of the literature on the health impacts of wildfire smoke. It would be advised that the authors do some more reading on these topics and/or change the language - again, see specific points related to this below.

Because the reviewer addresses these specific points again below, we address them and document changes below.

Specific Points:

Introduction:

1. Lines 28-33 may be missing a good amount of literature on the stress and mental health impacts of wildfire exposure not just due to smoke. Many studies comment on the fact that some of the pathways by which wildfires affect a given health outcome could be both through stress or air pollution exposure pathways - for example see (Holstius et al. 2012) or (Cohen et al. 2022) or (Murphy et al. 2021). There is a also a recent literature review on the mental health impacts of wildfire that reviewed 60 studies on the mental health impacts of wildfire exposure (To et al. 2021) and many other papers have come out since that review on these topics. For just a few examples, look at the following: (Obuobi-Donkor et al. 2022; Usher et al. 2022).

I can add these citations, and change our intro accordingly.

Methods:

1. The subsection titled "Study population and outcome data" describes where the data came from, but not much about the people in the dataset nor specifics about the outcomes were collected in those data. The authors state that daily healthcare visits were collected. Were these separated into visiting a physician versus an ED visit?

It seems that both reviewers were confused by the naming of this section, so it must have been confusing! We have changed the name of the section ‘study population and outcome data’ to just ‘study population’. There is another section called ‘outcome definition’ where we do give all these details.

*We obtained daily counts of all-cause outpatient visits, all-cause inpatient admissions, and all-cause emergency department (ED) visits, as well as inpatient admissions and ED visits specifically for circulatory or respiratory disease outcomes made by KPSC members 45 and older who rented DME. Causes were identified using International Classification of Diseases 10 codes I00-I99 and J00-J99, respectively. We included visits from January 1st, 2016 to March 15th, 2020.*

Page 5, ‘Outcome definition’.

1. What about if the visit had nothing to do with the person's use of DME?

As we mentioned in the responses to reviewer 1’s comments, we selected this population because we anticipate that people using DME are more vulnerable to wildfire exposures than people who don’t use DME. We aren’t interested in whether healthcare visits made by DME users are for the functioning of DME, or are made for conditions associated with DME, rather, we hypothesize that DME may be an indicator of social and medical vulnerability. We think DME may indicate vulnerability because DME is usually rented to address a disability or medical issue (Casey et al, 2021).

Visits in our paper may or may not be for DME use. We have added specific language to indicate that:

*We used electronic health record data from KPSC to identify all individuals who were 45 or older as of October 28th, 2019 and had rented DME in the year prior. We excluded younger DME renters in order to focus on socially and medically vulnerable older adults, but also to exclude breast pump users, a healthy subgroup of the otherwise vulnerable DME using population, who we did not hypothesize to be particularly vulnerable to wildfire exposure. We obtained daily counts of healthcare visits by this population at the Zip Code Tabulation Area (ZCTA) level, in seven counties in Southern California from January 1st, 2016 to March 15th, 2020.* ***These visits were not necessarily related to DME use.*** *236,732 DME patients lived in the study area, which covered most of San Bernardino, Orange, Los Angeles, Riverside, San Diego, Ventura, and Kern counties (Figure 1). The area consisted of 582 ZCTAs, each containing 1-1773 patients. During 2018 and 2019, these seven counties experienced 23 wildfires that each burned over 3 km2 in California,34,36 contributing to wildfire smoke in the area.*

*Page 3, Methods, study population.*

1. It would be good to know what the outcome is defined as and if any demographic data were collected on the patients.

The outcome definition is covered in the “outcome definition” section below, quoted here:

(This may have been unclear due to our confusing naming of the sections, which we have now corrected.)

## **Outcome Definition**

*We obtained daily counts of all-cause outpatient visits, all-cause inpatient admissions, and all-cause emergency department (ED) visits, as well as inpatient admissions and ED visits specifically for circulatory or respiratory disease outcomes made by KPSC members 45 and older who rented DME. Causes were identified using International Classification of Diseases 10 codes I00-I99 and J00-J99, respectively. We included visits from January 1st, 2016 to March 15th, 2020.*

*Daily visit counts by ZCTA were low and often zero (median outpatient visits = 1, IQR = 3, median ED and inpatient visits = 0, IQR = 0). For the wildfire PM2.5 analyses, to avoid zero-inflation in our models, and to increase statistical power, we could have aggregated ZCTA counts to the weekly level. However, prior studies of wildfire smoke exposure have found associations between same-day air pollution and healthcare visits over the course of the following week.18–21 To evaluate a lagged effect in our data, we required daily healthcare visit counts, therefore, we opted to aggregate our data into higher-level spatial groupings of several ZCTAs based on spatial proximity (hereafter ‘ZCTA groupings’; grouping method described in the eAppendix).*

*For proximity and evacuation analyses, we used ZCTA level daily visit counts aggregated to the weekly level. Because our exposure data was not as granular as that in the PM2.5 analyses, as we used last recorded fire boundaries and last recorded evacuation zones rather than daily data, we evaluated relationships at the weekly level. This aggregation also removed weekend-weekday patterns in outpatient visits, increased power, and reduced zero inflation. We considered a week exposed if the Woolsey or Getty fire burned any day that week.*

*Page 5 of main text.*

We have addressed the absence of demographic data in our response to reviewer 1’s comment #2 – unfortunately, we were not able to obtain demographic data due to resource limitations, but we do describe the study population in detail in Casey et al 2021.

1. Was the data de-identified before sending to the authors and that is how it was considered not human subjects research?

Yes – thank you for pointing this out, we have added this fact to the main text.

*The KPSC Institutional Review Board (IRB) approved this study, and the Columbia IRB did not consider it human subjects research, since the data were de-identified before the authors received them.*

*Page 3, Methods, Study population.*

1. Lines 54-56 in "Wildfire PM2.5" section- what data were used in the imputation models for the counterfactual PM2.5? Presumably, some monitoring data only? Were monitoring data used in the machine learning models for total PM2.5?

TARIK’S RESPONSE.

1. Lines 56-59 in "Wildfire PM2.5" section: I am a bit confused by this sentence: "We finally estimated the difference between such counterfactual values to observed values during an exposure to wildfire smoke to estimate daily/ZCTA levels of wildfire smoke PM2.5." What do the authors mean by "During an exposure to wildfire smoke"? Were the estimates only made during specific time periods during which the air quality was affected by wildfire smoke?

TARIK’S RESPONSE.

1. Line 59 in "Wildfire PM2.5" section - How did the authors "obtain daily/ZCTA levels of non-wildfire PM2.5"? Is this from their imputation models or did someone else provide it and thus it was obtained?

This was from the imputation models. We’ve added that to the text:

*We finally estimated the difference between such counterfactual values to observed values during an exposure to wildfire smoke to estimate daily/ZCTA levels of wildfire smoke PM2.5, and daily/ZCTA levels of non-wildfire PM2.5.*

1. Lines 4-6 in in "Wildfire PM2.5" section (on next page): I don't see an explanation of the "higher-level ZCTA groupings" in the outcome definition section (presuming that this is referring to the section titled "Study population and outcome data".

Now that I see the Outcome Definition section, I suggest renaming the first section so that it does not refer to outcome data and putting the outcome definition before the exposure section so that the reader understands what is meant by 'ZCTA groupings" when first mentioned.

We have renamed the ‘study population and outcome data’ section to just ‘study population’. This confused both reviewers! We’re sorry.

However, when we mention these ZCTA groupings in the study population section, we say ‘see the outcome definition section below for a description of what these are’, and then the next time we mention them, we define them. We feel that reorganizing the rest of the methods section here would disrupt the flow of the paper, but we are open to discussing it more.

1. Suggestion to mention the Thomas Fire in the study period explanation just before the methods section. Additionally, if "most" of the Thomas fire burned outside of the study area, that implies that "some" of the fire was inside the study area and thus some people in the study population may have been within a region to be considered exposed. By excluding this exposure, those people's exposure would be misclassified.

Please see reviewer 2’s comment #14 for an explanation of how and why we mentioned to Thomas Fire, and a revised sentence to make this clearer. See reviewer 2’s comment #12 for an explanation of how control regions were chosen to avoid exposure misclassification.

1. I looked at the three cited papers for the 20km threshold of affected by a fire and didn't see any of them mention 20km. Please clarify how this distance was chosen.

What we wrote was: “We hypothesized that living within 20km of a fire perimeter could elicit a stress response, similar to effects described in previous studies.7,8,41”

We have no particular reason for choosing 20km beyond the fact that it is ‘close’ – unfortunately, we haven’t found anything in the literature that’s more precise than that. We know that being close to a fire and being in the actual burned area are stressful, and no study we know of has evaluated how that stress changes quantitatively with distance. These studies all show that being close to a fire is stressful, and don’t mention 20 km.

We have revised this statement to: *We hypothesized that living within 20km of a fire perimeter could elicit a stress response, since stress responses have been described in previous studies at various distances from wildfires.7,8,41*

*Page 4, Proximity to wildfire.*

1. This sentence, "The DID estimators subtracted the change in visit frequency during a fire among control ZCTAs (difference 1) from the change in visit frequency during a fire among ZCTAs exposed to the fire or evacuation zone (difference 2)." could be stated more clearly, especially for someone who is unfamiliar with DID. The first thing that could be stated clearer is how control ZCTAs are defined (this could go into the section on assigning exposure to wildfire or evacuation) and stating that the "DID estimators subtracted the change in visit frequency during a fire compared to when there was not a fire (difference 1)" and the same for difference 2. Additionally, the authors need to explain how the time periods of "not a fire" were defined, if indeed, I am interpreting what was done correctly. If I was not, then this should be revised to be clearer about the difference.

I think our revised explanation of how controls were chosen addresses this issue:

*To avoid bias in our analyses, we excluded some observations from some ZCTAs from the pool of controls. If a ZCTA was exposed to the Getty and Woolsey Fires, or exposed to any other large fire (>500 km2) during the study period, we excluded observations from that ZCTA after the fire exposure. Since almost all ZCTAs in the area would have been exposed to fire at some point, and we felt that ZCTAs exposed to other fires would serve as ideal comparison groups, but only prior to their exposure to those fires. We used a CALFIRE fire perimeter data36 to identify all fires > 500 km2, and excluded data from ZCTAs within 20 km of any of these fire boundaries, from the fire ignition date onward*.

*Page 6, proximity to wildfire and evacuation.*

Joan – what should we say about the DID thing being unclear to people who don’t’ know DID? a) everyone does, and b) if they don’t we have no hope explaining it in a paragraph LOL.

1. When the authors say that they performed 20 regression analyses in this section due to evaluating each relationship separately for each fire and each type of healthcare visit, I would presume that they also evaluated them separately for exposure to the fire as well as exposure to evacuation orders, given that these were two separate exposures. Is this correct? If so, please revise that to be clear.

Thank you – we have revised this:

*To evaluate proximity to and evacuation from wildfire, we used a difference-in-differences (DID) analysis with negative binomial regression to estimate the associations between wildfire proximity and evacuation and weekly ZCTA-level healthcare visit counts. We evaluated each relationship separately for each fire, for evacuation and proximity, and for each type of healthcare visit.*

1. Additionally, please clarify if there were two of three fires assessed as the paper in some places only refers to the Getty and Woolsey fires and in other places also refers to the Thomas fire.

We have revised the text to more accurately describe how the Thomas Fire is involved. We mention the Thomas Fire once, because it was a huge fire that burned during our study period, and we wanted to explain why we did not choose to include it in our proximity/evacuation analysis. We do not mention it again, and the only mention of it in this (now-revised) excerpt:

*“Notably, The Thomas Fire also burned over 1100 km2 during our study period.38 However, most of the fire burned in the rural northern corner of Ventura County and outside the study area. Therefore, we did not include the Thomas Fire in the proximity analyses, since very few participants would have been exposed to it. Still, smoke from this fire contributed significantly to wildfire PM2.5 in Ventura County in December 2017, and therefore was included in our PM2.5 analyses (Figure 2).”*

1. Lines 38-39 in the "proximity to wildfire and evacuation" section - when the authors say that they excluded all ZCRAs exposure to other large fires, is this just other large fires within the study period or all large fires ever?

See comment #12 for our revised description of this.

1. Additionally, this sentence, "Therefore, we excluded observations from these ZCTAs made during and after fire exposures." is unclear - during and after which fire exposures?

See comment #12 for our revised description of this again – we removed this sentence.

1. Given the study area has been exposed to fires and evacuations over many decades, it would prove hard to ensure that control ZCTAs had never in the tenancy of people living in those ZCTAs been exposed to a fire or an evacuation zone.

Yes, we agree with the reviewer here - this is why we only excluded ZCTAs that were exposed during the study period.

1. And in the last sentence of this paragraph, the fire ignition date of which fire?

Again, we revised this – see comment #12 for our revised version.

Results:

1. Line 28 of "PM2.5 exposure" section - what USEPA limit are the authors referring to? (Presuming the daily PM2.5 NAAQS standard (which is not technically a "limit"), but this should be clarified).

Revised to say *“just under the annual USEPA NAAQS standard of 12 (Figure 2a).”*

1. For all results, to be clearer, the authors should state the population being studied. For example, I suggest adding "among DME users" after "outpatient visits" to the sentence "In adjusted negative binomial models, a daily 10 /3 increase in wildfire PM2.5 was associated with a decrease in risk of outpatient visits one day later (RR = 0.96, 95% CI: 0.94, 0.99),". The reason for this suggestion is so that someone doesn't read this sentence out of context and think that all outpatient visits declined in the study area on days with higher wildfire PM2.5 concentration. Similarly, the term "among DME users" should be added to the second sentence in this paragraph.

This makes sense – we have revised both the sentences suggested here, and the discussion in general, to better reflect the study population.

1. It is also interesting that the authors decided to highlight in the writing the one day (lag 1) with a protective effect for wildfire PM2.5 when on lag days 2, 3, 5, and 6, they found a significant adverse effect (and a null finding on lag days 0 and 4). In my interpretation of Table 1a, there are more significant effects for all-cause outpatient visits to increase due to wildfire PM2.5 than decrease and additionally, a distributed lag estimate across all of these days would be informative, rather than just each lag day separately.

We interpret this comment to mean that we should highlight the significance adverse effect more than the initial decrease. We have edited the sentence in question to do so:

*We observed an initial next-day decrease in all-cause outpatient visits, and then a positive association between wildfire PM2.5 and all-cause outpatient visits among DME users for the two weeks following exposure, suggesting that there is overall a significant increase in all-cause outpatient visits among DME users following exposure. These findings are consistent with much of the literature.*

We also did perform a sensitivity analysis that looked at the cumulative effect – joan, do you think this is what she means/is this relevant to her question?

1. Lines 43-45 of the "PM2.5 exposure" section should state that "weekly wildfire PM2.5" was not associated with frequency of other visits (emphasis added).

We revised the line in question: *“Weekly wildfire PM2.5 was not associated with the frequency of any other visits.”*

1. Throughout, I would suggest not referring to the EPA PM2.5 NAAQS as a "recommended daily limit" because the NAAQS is more of a regulatory standard than it is a recommended limit. Although it is supposed to be based on health, we know there are health impacts below that threshold. Additionally, with the NAAQS, areas are allowed to exceed the NAAQS on 2 percent of days, essentially. This is a minor point, but I think helps to keep what the EPA threshold values are through the NAAQS clear to a reader. The WHO Air Quality guidelines are more of a "recommended" daily and annual limit as they are not regulatory in nature and are recommendations that have no 'teeth' to them and are more in-line with what is known about health impacts of PM2.5.

We think this is a great point and have changed our language throughout so as to not refer to the limit, but rather the standard.

CHANGE FIGURE CAPTIONS?

1. Figure 3 - the title should be clarified - the change in frequency of visits during a fire compared to what?

In our figure description, we do already provide a reference group:

*“We used negative binomial regression to evaluate the effect of wildfire evacuation or proximity during an active fire. The DID estimators subtracted the change in visit frequency during a fire among ZCTAs far from the fire or evacuation zone (difference 1) from the change in visit frequency during a fire among ZCTAs exposed to the fire or evacuation zone (difference 2).”*

*Figure 3 description.*

1. Figure 3 - the two orangey/yellow colors are hard to discern. Given that these outcomes are not a scale related to each other, there isn't really a need to have a color scheme that increases from one shade to another. I would suggest a color scheme where every color is very different from each other or a vertical axis where each healthcare visit type is written.

We have reworked this figure not to include colours and hopefully to be clearer to everyone, since we also got feedback from reviewer 1 on this. The updated figure is included in our resubmission.

1. Lines 7-8 in the "Woolsey Fire proximity and evacuation exposure" section - the outpatient visits remained the same during the fire compared to what? The same goes for the next sentence - compared to what? And "outpatient visits" should be "all-cause outpatient visits".

We revised this to: “*Residence in an evacuation zone of the Woolsey Fire during the fire was also significantly associated with increased inpatient admissions for cardiorespiratory disease compared residence outside of it, and associated with decreased all-cause outpatient visits, though the confidence interval included the null (Figure 3).”*

1. The paragraphs about the increase in visits during the Woolsey Fire need to be prefaced with the confidence intervals, as the cardiorespiratory emergency and cardiorespiratory inpatient are really null based on these confidence intervals.

We have revised this paragraph to include confidence intervals and better descriptions of which results are null:

*Woolsey Fire proximity during the fire was significantly associated with decreased outpatient visits (RR = 0.89, 95% CI: 0.79, 1.00), and associated with increased inpatient admissions for cardiorespiratory disease, though not significantly (RR = 1.48, 95% CI: 1.01, 2.17) (Figure 3). We observed similar associations between Woolsey Fire evacuation exposure and healthcare visits. Residence in an evacuation zone of the Woolsey Fire during the fire was also significantly associated with increased inpatient admissions for cardiorespiratory disease compared residence outside of it, and associated with decreased all-cause outpatient visits, though the confidence interval included the null (RR = 1.76, 95% CI: 1.02, 3.05, RR = 0.87, 95% CI: 0.73, 1.04 respectively) (Figure 3).*

1. I suggest rewording the findings for the Getty Fire. Someone who is not an epidemiologist would read this as all of these visits increased during the fire and miss the nuance that there was an increase everywhere and therefore it was not associated with proximity the fire or being in the evacuation zone and that all of those findings were null.

We’ve removed the sentence about increasing visits completely, since on second thought, we didn’t think it was important to say and comes close to being a Table II fallacy. citation

Discussion:

1. I would change the first sentence of the discussion to have a "but" instead of an "and". It is very interesting that next day visits decreased but then for almost all other days that week and the following week there were increased visits.

We think this is a good point and have changed the sentence in question: *Using electronic health data describing 236,732 Kaiser Permanente DME-using patients from 2016-2020, we found that an increase in wildfire PM2.5 concentration was associated with next-day decreases in all-cause outpatient visits, but increases in all-cause outpatient visits up to two weeks later.*

1. The first paragraph of the discussion should mention that proximity and evacuation during the Getty Fire were not associated with any visits among this population group.

We have added this sentence:

*Using electronic health data describing 236,732 Kaiser Permanente DME-using patients from 2016-2020, we found that an increase in wildfire PM2.5 concentration was associated with next-day decreases in all-cause outpatient visits, but increases in all-cause outpatient visits up to two weeks later. Increases in wildfire PM2.5 were not associated with the frequency of ED or inpatient visits among DME users. Residential proximity to the large Woolsey Fire was also associated with fewer all-cause outpatient visits, as well as more cardiorespiratory inpatient visits, and evacuation from the Woolsey Fire was associated with increased cardiorespiratory inpatient visits among DME users.* ***Getty Fire evacuation or proximity was not significantly associated with frequency any kind of healthcare visit.*** *Our study was unique in that we included inpatient, ED, and outpatient visits, evaluated healthcare utilization among DME users, a group hypothesized to be susceptible to disaster and wildfire smoke exposures, and examined residence near a wildfire or an evacuation zone.*

1. Second paragraph of the discussion: I disagree that there is a strong relationship between wildfire smoke exposure and cardiorespiratory health in the literature. There is a strong relationship for respiratory health impacts of wildfire smoke and growing, but mixed, evidence of impacts of wildfire smoke on cardiovascular health.

There are some studies that have found that wildfire PM2.5 affects people who are not using rescue medication usage for asthma more than those who are - see a discussion of this hypothesis in (Reid et al. 2016) and (Lipner et al. 2019).

We have revised this sentence to: *The literature describes a strong relationship between wildfire smoke exposure and respiratory health20, and a strong relationship between PM 2.5 exposure and cardiovascular health, though the effects of wildfire smoke on cardiovascular health are still being characterized.*

1. I find the statement "These findings are consistent with much of the literature" to be strange, and in my opinion, not true. I do not know of any study that has found significant impacts of wildfire PM2.5 on health outcomes beyond the first week and the paragraph that this is part of mentions that very few studies have even investigated whether wildfire PM2.5 is associated with outpatient visits at all (which I agree is the case).

We have revised this to: *These findings are consistent with much of the literature in that they indicate increased healthcare utilization following smoke exposure.*

1. In the limitation section about using the Kaiser data - I wonder whether people, due to evacuation, may have been farther from the clinic they normally use and may have gone elsewhere?

We are using KPSC administrative data, which means that even if the study participants left the affected area, and sought care at another KPSC location, they would still be included in our dataset. It is absolutely true, however, that if they did not seek care within the KPSC system, they would not be included in the dataset. They would be motivated to go to Kaiser because they are insured there. We do discuss this, but we can make it clearer that other Kaiser clinics would have been included in our data. We also discuss this in our response to reviewer 1’s comment #7.

1. Additionally, the authors should mention that this population is likely different from the non-Kaiser population in significant ways such as they all have health insurance likely through their employer.

**Yeah I mean some of these people have Medicaid, right? Joan is there a citation we can use that evaluates Kaiser’s representativeness?**

1. Can the authors defend the statement that "all visits were infrequent during the study period"? Were the visits more infrequent than normal?

I just meant that a lot of days were 0 visits. Will revise.

Related to this, was an assessment of power calculated to defend the statement that the models may have been underpowered. This is a big deal if the analyses were underpowered to even have detected an association. For the statement about the differences in findings by proximity and evacuation to the Woolsey Fire, wasn't it true that evacuation zones are essentially subsets of proximity areas? Thus, it is not surprising that the confidence interval was larger for the evacuation zone.

**No power analyses…. And the last part is true. Not sure what to say here.**

* Didn’t’ do power analyses cause we worked with the whole population
* We only meant that we observed wide Cis
* Future studies could look at larger populaitons
* No idea about effect size.

The authors should say more about whether they think that adjustment for the ZCTA-level SES variables sufficiently dealt with spatial confounding. If the authors are truly concerned about spatial confounding, they should assess for spatial autocorrelation in the residuals of their models and if they find some, they should then run a spatial error model.

**Maybe we can chat about this?**

The paper needs a better concluding paragraph and not end on a statement about a limitation.

Will add.

References:

Cohen O, Shapira S, Furman E. 2022. Long-Term Health Impacts of Wildfire Exposure: A Retrospective Study Exploring Hospitalization Dynamics Following the 2016 Wave of Fires in Israel. Int J Environ Res Public Health 19:5012; doi:10.3390/ijerph19095012.

Holstius DM, Reid CE, Jesdale BM, Morello-Frosch R. 2012. Birth weight following pregnancy during the 2003 Southern California wildfires. Environmental health perspectives 120:1340-5; doi:10.1289/ehp.1104515.

Lipner EM, O'Dell K, Brey SJ, Ford B, Pierce JR, Fischer EV, et al. 2019. The Associations Between Clinical Respiratory Outcomes and Ambient Wildfire Smoke Exposure Among Pediatric Asthma Patients at National Jewish Health, 2012-2015. GeoHealth 3:146-159; doi:10.1029/2018GH000142.

Murphy VE, Karmaus W, Mattes J, Brew BK, Collison A, Holliday E, et al. 2021. Exposure to Stress and Air Pollution from Bushfires during Pregnancy: Could Epigenetic Changes Explain Effects on the Offspring? Int J Environ Res Public Health 18:7465; doi:10.3390/ijerph18147465.

Obuobi-Donkor G, Eboreime E, Shalaby R, Agyapong B, Oluwasina F, Adu M, et al. 2022. Evaluating the Prevalence and Predictors of Moderate to Severe Depression in Fort McMurray, Canada during the COVID-19 Pandemic. Int J Environ Res Public Health 19:7090; doi:10.3390/ijerph19127090.

Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT. 2016. Critical Review of Health Impacts of Wildfire Smoke Exposure. Environmental health perspectives 124:1334-43; doi:10.1289/ehp.1409277.

To P, Eboreime E, Agyapong VIO. 2021. The Impact of Wildfires on Mental Health: A Scoping Review. Behav Sci (Basel) 11:126; doi:10.3390/bs11090126.

Usher K, Durkin J, Douglas L, Coffey Y, Bhullar N. 2022. Coping styles and mental health outcomes of community members affected by black summer 2019-20 bushfires in Australia. Int J Ment Health Nurs; doi:10.1111/inm.13035.

\* \* \* \* \*

Preparing a revision

1. For estimates of causal effects, we strongly discourage the use of categorized P-values and language referring to statistical significance, including whether a confidence interval covers the null. We prefer instead interval estimation, which conveys the precision of the estimate with respect to sampling variability. We are more open to testing with respect to modeling decisions, such as for tests of interaction and for tests for trend.

2. We do not permit acronyms unless they are generally recognized by epidemiologists (e.g. HIV is okay, but LVA is not). When in doubt, we recommend that you spell out.

3. Please do not include uninformative precision (excessive decimal places). For example, percents should be rounded to nn%, n.n%, or 0.0n% and risk ratios should be rounded to nn, n.n, or 0.nn unless clarity of the presentation and the sample size justify more significant digits.

4. Please be sure to include explicit information about approval of human subjects research by an independent review board. If no such review was required, include an explicit statement about why the requirement for review was waived.

5. Do not include public health policy recommendations in Brief Reports or Original Articles that present new research findings.

6. Data appearing in the abstract must also be cited in the main text, not just in tables or figures.

7. Resubmissions must adhere to word limits. The word limits for main text (generally the introduction, methods, results, and discussion) are 1500 words for Brief Reports (plus 150 words for its abstract), 4000 words for Original Articles (plus 250 words for its abstract), 5000 words for reviews (plus 250 words for its abstract), 2000 words for Commentaries (no abstract), 600 words for Research Letters (no abstract), and 400 words for Letters to the Editor (no abstract).

8. We advise that total word counts for Original Articles should not exceed 7500 words and for Brief Reports should not exceed 3500 words. The total word count includes main text (introduction, methods, results, and discussion), bibliography, figure legends, tables, and figures (250 words per figure, including each figure in a panel). The title page, abstract, acknowledgments, and funding information do not count in the total word count.

9. Figure labels: Make font size as large as possible, so as to be legible when figures are reduced for publication (typically one column [8.5cm] in width).

10. Footnotes to tables and figures should use superscript lowercase letters to link content to the footnote, not symbols or numerals.

11. Do not use parenthetical phrases like “(data not shown), (results not shown), or (available from the authors upon request).” In these circumstances, the data or results should be provided in Supplementary Digital Content.

12. Additional details regarding submission requirements can be found in the Instructions for Authors, which are posted at http://edmgr.ovid.com/epid/accounts/ifauth.htm .

Preparing for resubmission

13. Prepare a response document for the Editor that responds point-by-point to the reviewers' comments (presenting each comment followed by your response). Give the page number where revised text can be found and, where practical, paste revised text directly into the reply document.

14. Submit versions of the manuscript with and without your changes displayed.

15. Supplementary Digital Content should be submitted as a single PDF file, and you should use our convention - e.g. eFigure 1, eAppendix 2 - to label and refer to online content.

16. Authors should submit copies of any closely related manuscripts (published, in press, or under review).

17. Please revisit information about page charges and color printing charges available in the Instructions for Authors, which are posted at http://edmgr.ovid.com/epid/accounts/ifauth.htm .

18. We request that the complete revised manuscript (with all tables and figures) be completed by 23 Nov 2022. If you are not able to meet this deadline, please notify the editorial office.

Resubmitting via Editorial Manager