Introduction:

With the aging US electrical grid and climate change underway, power outages are becoming increasingly common, and they pose substantial health risks to vulnerable populations such as older adults and people who use durable medical equipment. To study the health effects of outages, we assembled a dataset of nation-wide county-level power outages, used in Do et al. “Spatiotemporal distribution of power outages with climate events and social vulnerability in the USA”.

The study of power outage as a health-related exposure is new, and there are several complexities in measuring this exposure.

At the county level, we characterize power outages based on the proportion of people affected in the county and how long they last. However, unlike some environmental exposures such as air pollution or noise, in reality, power outages are a binary exposure. Power is either on or off at a person’s residence. At the individual level, we don’t know how long an outage needs to be to be clinically relevant. When creating population-level (in this case county-level) measures of power outage, we don’t know what proportion of people in the county need to be affected for the outage to influence health, or what county-level measures of outage length are clinically relevant.

Additionally, power outage data comes directly from utilities, and sometimes data is missing. It’s unclear if this missing data is informative, and how missing outage data could influence the results of an epidemiologic study of the effects of outages on health.

In this paper, we present the results of simulations we designed to test how different definitions of power outage exposure and missing data could influence the results of a health study of short-term exposure to power outages. Specifically, we employ a case-crossover design to estimate the effect of county-level power outage exposure on county-level Medicare hospitalization rates for cardiorespiratory concerns nationwide.

Researchers working with several kinds of environmental exposure data similar to power outage data could face challenges in defining exposure and handling missing data like the ones we addressed with simulations in this paper.

Data on lead exposure or radon exposure follow a similar structure to power outage data. Exposure to radon, or lead in paint, are really binary exposures at the individual level – either there is lead paint in someone’s residence, or there isn’t. Spatial unit level measures allow us to infer about the chances of an individual exposure (for example, one neighbourhood may have more homes with lead paint than another), but applying any area-level measure of exposure to individuals will always create some exposure misclassification. When conducting a population-level study, data custodians must decide on a spatial-unit-level threshold to define exposure – for example, a proportion of homes in a neighbourhood with lead paint. Here, we tested the sensitivity of estimated health effects of power outage to different definitions of power outage. We aimed to learn how much changing the area-level definitions of power outage could bias the results of our epidemiologic study where we evaluate the effect of power outage exposure on cardiorespiratory hospitalization counts by people insured with Medicare.

Water quality data at the county level are also made up of several smaller spatial units of local water systems, and may be missing water systems that cover a certain proportion of people in a county, or time series for certain water systems may be incomplete. Here, we also test how missing data of these kind could bias results of our study. We hope other environmental epidemiologists can design simulations similar to what we did here, to test questions about exposure definition and data coverage with datasets that follow a similar structure. Simulations like these could be used to assess potential bias in similar contexts.

Methods:

We got raw power outage data covering the continental US for the years 2018-2020 from poweroutages.us. This data is scraped in real time from utility company APIs and then aggregated. The raw data contains counts of customers without power all served by the same utility in a sub-county unit. (Utilities define ‘customers’ as grid connection, which correspond to households, apartment buildings, or businesses.) These sub-county units can be entire counties, cities (where there are possibly multiple cities in a county), or neighbourhoods (where there are possibly multiple neighbourhoods in a city, in turn nested in a county). All sub-county units were nested inside counties. Many utilities can serve the same location, so these units were not necessarily geographically distinct. Two houses next to each other might be in two different spatial units in the power outage data if they were served by different utilities.

The sub-county unit level data came in a time series of ten-minute intervals covering 3 years, where there was an estimate of the number of customers without power in a sub-county unit in every 10 minute interval. The counts of customers out did not necessarily track the same customers – if 10 people were out in two subsequent 10-minute periods, the data do not give us information about whether the same 10 households were out. They only tell us that 10 households were out in each interval.

Because all subcounty units were nested inside counties, we were able to aggregate 10-minute counts of customers without power to the county level. We also aggregated these counts to the hourly level, to reduce the size of the data and smooth out noise.

Simulation on exposure misclassification:

We needed to define a power outage: when was a county without power? Other studies have defined binary power outage exposures over a spatial unit (whether there was an outage, Y/N) by deciding that there is an outage when the percentage of customers out exceeds a f certain threshold. For example, Northrop et al. defined three kinds of outages when the percentage of customers in a power operating locality exceeded 10%, 20%, and 30% of customers. We reasoned that if percentages of customers without power were high, there was a higher probability that many households in the area were experiencing a longer outage, and, some households were likely to be without power continuously while the percentages of customers out remained high.

We used several percentage-based definitions, because this allowed us to test the relationship between different ‘strengths’ of exposure and health effect estimates. However, we used 0.1%, 0.5%, and 1% as cutoffs, as in Dominanni et al. and Do et al., since we were working with much larger spatial units with larger populations than Northrop et al., so these smaller percentages still meant a large number customers were without power.

We initially hypothesized that for an individual, being without power for 8 hours or longer could produce health effects, since the batteries of most durable medical equipment last 8 hours. When batteries die, vulnerable people may be without life-sustaining medical devices like oxygen tanks or at-home ventilators. At the county level, we initially marked days as exposed or unexposed only if the percentage of customers without power in the county was above the 0.5% threshold for 8 continuous hours.

However, we don’t have any evidence to support the claim that 8 hours is the true threshold after which a power outage becomes clinically relevant, at the individual level or as a measure of outage duration at the county level. So, we were concerned that if it was in fact shorter or longer outages that mattered for health, we could get really biased results in our study if we cleaned the data and marked days as exposed or unexposed based on 8 hour outages. Said differently, if it was instead 4 hour outages or 12 or outages that were clinically relevant and mattered for health, and we cleaned the data as if it were 8 hour outages that mattered, we would be introducing serious exposure misclassification.

How we did the simulation:

In this simulation we wanted to test how much we would bias our results if we defined a power outage incorrectly. What would happen if at the county level, 12-hour outages produce health effects, but we incorrectly assumed 8 hour outages mattered for health?

We generated 100 simulated counties. Then, we assigned each county a number of nested sub-county areas. We determined the number of sub-county areas by drawing from the empirical distribution of sub-county areas in the poweroutages.us dataset. We populated each sub-county area with a total number of simulated customers, again drawn from the empirical distribution of customers in sub-county areas in the real poweroutages.us dataset. We then generated ten-minute counts of customers without power for each sub-county unit for one year. The counts of customers without power at each 10-minute interval was also drawn from the empirical distribution of the proportion of customers without power in the poweroutage.us data, then multiplied by the customers served in the sub-county unit. We constructed our simulated data from the proportion of customers without power because of very different counts of customers served by sub-county unit nationwide.

We aggregated these sub-county 10-minute counts of customers without power to the hourly level. We also aggregated hourly counts of customers out to the county level, by summing all the customers served in each sub-county area and the customers out in each hour in each sub-county area.

We used a “0.5% out” definition to develop a county time series of hours exposed to power outage based on hourly-county counts of customers without power. We marked a county-hour exposed if the percentage of customers without power in county *i* during hour *j* exceeded 0.5% of the customers served in county *i*. Finally, we aggregated to the daily level: we considered a county-day as exposed if there were 8 consecutive hours of ‘power outage on’ (customers without power percentages > 0.5% of county) in that county on that day. Power outages could last more than 24 hours. We also considered a county-day as exposed if a power outages lasting longer than 8 hours ended on that county-day. This left us with a one-year time series indicating if there was a power outage on each day in each of the 100 simulated counties.

First, to establish a base case representing an unbiased scenario where exposure was defined correctly, we generated simulated outcome data of cardiorespiratory hospitalization counts by day and county based on the simulated county-day power outage exposure timeseries described in the previous paragraph. We drew cardiorespiratory hospitalization counts for each county-day based on the total number of customers living in a county from a Poisson distribution with a base rate of 0.1%. This base hospitalization rate *is the actual hospitalization rate in Medicare data.*  County-days that met the 8-hour outage definition received a 10% rate increase (for a total hospitalization rate of 0.11%).

To generate an estimate for the effect of an 8-hour power outage on daily county-level hospitalization counts, for each day exposed in a county, we chose control days by matching on the day of week and month. In a multi-year analysis, we would also match on year. This would control for long-term trends, seasonality, and weekday effects in hospitalization rate.

We used these case and control dates in a Poisson model, one for each county (for a total of 100 models) to model the relationship between the binary 8-hour power outage exposure and daily hospitalization counts, with an offset for the number of customers in a county. This model represented an unbiased scenario where exposure was defined correctly.

Flag major issue: a customer isn’t a person, and a person isn’t a medicare recipient. WHAT DO?

Answer: perhaps nothing? How many people are medicare recipients? About 20% of population? What is the avg hh size? Around 3 people? There are maybe 6 million commercial buildings in the us? Could we just ignore this?

Again, we wanted to test how much we would bias our results if we guessed wrong about what length of power outage would cause health effects. Here, we’re talking about outages defined at the county level, rather than the individual level. It might be that 12-hour outages actually cause health effects, but we incorrectly assumed 8-hour outages mattered for health. In this simulation, we created outcome data – hospitalization data – where we increased the hospitalization rate when a county-day was exposed to a 12-hour outage or a 4-hour outage instead of an 8-hour one. Then, we modelled the relationship between the outcome data based on the 12-hour exposures and 4-hour exposures, and the exposure data based on the 8-hour power outage definition.

To do this, we created two additional exposure datasets for each county, defining a county-day as exposed if there was either a 4-hour outage (customers without power counts > 0.5% of total customers for 4 consecutive hours) or a 12-hour outage (defined the same but for 12 hours) in line with our prior 8-hour definition. We generated two additional datasets of outcome data (daily county hospitalization counts) based on the same base hospitalization rate of 0.1%, and a 10% rate increase (for a total hospitalization rate of 0.11%) on days with a power outage, but we based these counts on 4-hour outage exposure data and 12-hour outage exposure data. We classified county-days as exposed given the 8-hour definition and chose control days as above. We ran two sets of 100 Poisson models, one for the 4-hour exposure and one for the 12-hour exposure as above.

This represented exposure misclassification, where the misclassification was happening because we defined the length of a clinically relevant power outage incorrectly. We ran models assuming a county-level 8-hour power outage was the clinically relevant exposure, when in fact the “true” relevant exposure was either 4 hours or 12 hours. In the case where 4 hours was the relevant exposure, we would have missed many exposed county-days in our analysis, whereas if 12 hours was the relevant exposure, we would have identified too many exposed county-days if used an 8-hour duration for analysis.

We assessed the bias in these models by plotting the estimated rate increases with power outage exposure from each model in boxplots in Figure 1, and coverage of 95% confidence intervals - could be Figure 1. Exactly what we’ll plot is TBD. We also repeated the simulations for 2 other effect sizes of power outage on hospitalization (could also be Figure 1).

The poweroutages.us dataset is incomplete. Data were missing in a variety of ways. Sub-county units covered as little as 20% of the total customers in a county because many sub-county units were missing from the dataset. Some utilities don’t have a website that poweroutages.us could scrape. In addition, many time series for a neighbourhood-utility were incomplete, missing anywhere from 0%-100% of ten-minute intervals in our three years of data. Missing sub-county units or incomplete time series lead to missing data at the county level.

For the purposes of this simulation, we’ve assumed that data is missing at random. This is probably untrue – for example, small, rural utilities are much less likely to have a website or API, meaning that many of those utilities may not be in the poweroutages.us dataset. It’s also possible that utility websites and APIs crash during extremely large outages, meaning that there may be missing data when counts of customers out are extremely high. In both these cases, the missingness may be related to the outcome of hospitalization. Rural counties may have different hospitalization rates than urban ones. Large power outages are often caused by natural disasters, which may also affect hospitalization rates. However, to evaluate the potential effects of missingness not at random, we would have to speculate about so many different parameters in a simulation that we don’t feel that a simulation would be informative. So, we’ve limited ourselves to an investigation of missingness that we assume is random.

Missing data from sub-county units that are not in the dataset or incomplete time series could mean most of a county’s data is missing. We need to know when we should exclude counties from the epidemiological analysis because there is so much missing data that they will bias the effect estimates rather than contribute information.

If no power outage exposure data was missing from a county, then the data should have information on all the customers served for all of the hours in a year. The total amount of data missing from a county can be represented as a percentage of that total customer-hours in county *i*. Customer-hours present in county *i* should be:

[customers served in county *i*] \* [number of hours in the year] = N customer-hours

We created three new exposure datasets by randomly removing first 10% of the customer-hours­ in each of the 100 counties, then 30%, and then 50%.

Data is often missing for an entire sub-county unit in the poweroutages.us dataset. When that happens, we have no information on that unit – no average value for customers out, and sometimes not even an estimate of the number of customers served by the utility in the sub-county unit area. For this reason, it’s very hard to infer or interpolate missing values. Therefore, in our simulation, we replaced missing values of counts of customers without power with zero. Zero is also by far the most common value, and the average value, of customers out in the non-missing subcounty data in poweroutages.us.

We generated outcome data from the complete dataset before we added any missingness as we did in the previous simulations, based on 8-hour power outages as the health-relevant exposure. We drew ‘cardiorespiratory hospitalization counts’ for each day in each county from a Poisson distribution based on the total number of customers living in that county, with a base hospitalization rate of 0.1%, and a 10% rate increase (for a total hospitalization rate of 0.11%) on days marked as exposed to a power outage based on the 8-hour and 0.5% threshold definition.

Then, we cleaned the exposure datasets from which we had removed observations (i.e., 10%, 30%, and 50% removed) and assigned exposure data to each county-day. Again, we treated missing exposure data on a county-day as equivalent to no power outage occurring. In the three datasets with missingness, we replicated our analysis as above.

And we assessed bias and plotted results etc.