Impact of Wildfire Smoke Events on Seattle Children's Pediatric Patient Outcomes, 2006 - 2020

Daaniya Iyaz

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2022

Committee:

Tania Busch Isaksen

Coralynn Sack

Program Authorized to Offer Degree:

Environmental and Occupational Health Sciences

School of Public Health

University of Washington

Abstract

Impact of Wildfire Smoke Events on Seattle Children's Pediatric Patient Outcomes, 2006 – 2020

Daaniya Iyaz

Chair of the Supervisory Committee:

Tania Busch Isaksen

Department of Environmental and Occupational Health Sciences

Pacific Northwest wildfire smoke events have been increasing in prevalence and severity over the past three decades, resulting in documented negative health outcomes in adults. However, when examining the effect of wildfire smoke on pediatric health, the literature is scarce. We utilized a case crossover study to determine the odds of pediatric emergency department (ED) visits and hospital admission at Seattle Children's Hospital on smoke days versus non-smoke days during wildfire season (June to September) from 2006 to 2020. Our analysis used a conditional logistic regression model that controls for temperature and relative humidity to evaluate the association between daily average $PM_{2.5}$ concentrations and pediatric hospital encounters for all causes and specific conditions. The results indicate a 0.0% (95% CI: -3.0% - 3.0%) change in odds for all-cause same-day emergency department (ED) visits on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex and a 7.0% (95% CI: 3.0% – 12.0%) increase in odds for all-cause inpatient/observational hospital admissions. When stratifying by health outcomes, we found a 9.0% (95% CI: 1.0% – 17.0%) and a 11.0% (95% CI:

1.0% – 21.0%) increase in odds for being admitted to the ER for a respiratory and respiratory infection-related concern, respectively. We also observed statistically significant increases for inpatient/observational admissions odds, ranging from 5.0% - 8.0%, for lagged exposure on days 1 – 4 when using lag models that adjusted for Humidex. When evaluating inpatient/observational cases, we found a 44.0% (95% CI: 3.0% – 102.0%) increase in odds for trauma-related cases. Our findings add to the current understanding by observing increased odds of respiratory-related ED visits and all-cause hospital admissions for Seattle Children's patients on smoke days versus non-smoke days when adjusting for Humidex. The risk remained elevated with exposure to wildfire smoke up to 4 days prior to inpatient and observational hospital admission.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

95% CI: 95% confidence interval

AIRPACT: Air Indicator Report for Public Awareness and Community Tracking

APR-DRG: All Patient Refined Diagnostic Related Group

CO: Carbon monoxide

COPD: Chronic Obstructive Pulmonary Disease

ED: Emergency Department

EPA: United States Environmental Protection Agency

HAPs: Hazardous air pollutants

ICD-9: 9th revision of the International Statistical Classification of Diseases and Related Health Problems

ICD-10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems

IRB: Institutional Review Board

NAAQS: National Ambient Air Quality Standards

NEHA: National Environmental Health Association

NO_2: Nitrogen dioxide

OR: Odds ratio

PAHs: Polycyclic aromatic hydrocarbons

PHIS: Pediatric Health Information System

PM: Particulate matter

PM_{10}: Particulate matter less than 10 microns in diameter

$PM_{2.5}$: Particulate matter less than 2.5 microns in diameter

PTSD: Post-traumatic stress disorder

RR: Relative risk

SD: Standard deviation

SO_2: Sulfur dioxide

TRAP: Traffic-related air pollutants

VOCs: Volatile organic compounds

ACKNOWLEDGEMENTS

I have many people to thank for supporting me through every step of this project as well as my graduate program. First, I would like to express my sincere gratitude to my thesis advisor and committee chair, Dr. Tania Busch Isaksen. Tania worked with me from the beginning to understand what my interests and goals were, and it was through her guidance that I became a more diligent researcher, a semi-adept coder, and a rising public health professional. I am incredibly grateful for the hours she has spent answering questions, providing feedback, refining products, and ensuring me that I will, in fact, graduate. I would also like to acknowledge Dr. Coralynn Sack, my second committee member who spent a significant amount of time contributing her expertise to this project and answering my questions about epidemiology studies and medical conditions. Cora's astute feedback strengthened the analyses in this study, and her supportive nature gave me the confidence in my findings and in myself to share out the results.

I also have a vast amount of gratitude for our Seattle Children's Hospital collaborators. A huge thank you to Dr. Jon Cogen and Dr. Alexandra Perkins for demystifying the medical terminology and answering my questions ranging from pediatric anatomy to "what does APR-DRG stand for". Jon's encouragement and quick replies helped me power through any obstacles that came up. Alee's enthusiasm and handy email forwards provided the basis for much of the work in this project. I also wanted to thank Matt Hall who provided the dataset for us and had to answer questions at least twice a month long after his initial involvement with the study.

Much of the data analysis would not have been possible without the infinite genius and wisdom of the two PhD students that worked on this project with me, Annie Doubleday and María Valdez. Annie and María contributed an immense amount of time answering my innumerable questions and checking over code that I had been puzzling over at 3 AM. I would likely still be coding right now if it was not for their support.

And finally, my heartfelt gratitude goes out to my friends and family who supported me during time. Thank you to my mom, dad, and older sister who reminded me to hydrate and sleep throughout this project and thank you to my friends Carly and Claudia who have probably received over 100 texts about this project to date. Another round of thanks to my friends Joanne and Anna who also provided helpful feedback and solidarity while we all finished our programs together. This project was a team effort that helped me grow in several ways, and I am grateful for this opportunity as well as the people it connected me with.

LITERATURE REVIEW

WILDFIRE OCCURRENCE AND PREVALENCE

In the Pacific Northwest, changes in extreme heat, precipitation, and wind patterns have led to an increase in wildfire smoke occurrence (Xu et al., 2020). Globally, factors such as forest management practices, deforestation, and increases in the interface between human residence and densely forested areas have contributed to wildfire occurrence (Adetona et al., 2016). However, multiple studies have demonstrated that climate change is a key factor in increasing wildfire occurrence and severity (Abatzoglou & Williams, 2016; Turco et al., 2018; Xu et al., 2020). Significantly altered climate patterns, such as shifting rainfall concentration to winter, increases in soil evapotranspiration, more frequent and intense heat waves, and faster global surface wind speeds, create dry, oxygen-rich environmental conditions that set the ideal stage for wildfire formation (Bowman et al., 2020; Karnauskas et al., 2018; Sun et al., 2019; Xu et al., 2020). Due to climate change, resulting from human behaviour, forest fires consumed an estimated 4.2 million hectares of area during the years of 1984 – 2015, and this trend is expected to persevere in the future (Abatzoglou & Williams, 2016). The symptoms of anthropogenic climate change are documented in an increase of area burned, fire-season duration, and the number of large fires seen in the western United States, and Seattle is not an exception (Abatzoglou & Williams, 2016; Flannigan et al., 2013). Large fires are almost five times more frequent in the western US on an annual level when compared to fifty years ago, and these longer-lasting fires expose people to higher levels of smoke than historically experienced (Balmes, 2018). Climate change projections estimate a drastic increase in high-pollution smoke days resulting from wildfires, and these projections pose the greatest health impacts in

Washington, southern California, and Central Colorado (Liu, Mickley, Sulprizio, Yue, et al., 2016). As fire seasons become more severe (Flannigan et al., 2013), investigating the health outcomes resulting from wildfire smoke events has become a priority, especially since exposure is estimated to cost $11 – 20 billion in health care costs per year in the United States (Reid & Maestas, 2019).

WILDFIRE COMPOSITION AND $PM_{2.5}$

Wildfire smoke contains multiple pollutants in the form of gaseous emissions, including criteria air pollutants such as carbon monoxide (caused by incomplete combustion from smoldering fuel), nitrogen dioxide, small to ultrafine particulate matter, ozone (through secondary photochemical reactions), and sulfur dioxide as well as Hazardous Air Pollutants (HAPs), such as benzene and mercury (Balmes, 2018; Evans et al., 1977; Finlay et al., 2012; Naeher et al., 2007). The smoke composition also includes carcinogenic compounds such as aldehydes, Polycyclic Aromatic Hydrocarbons (PAHs), and Volatile Organic Compounds (VOCs), and the latter two can also be found on the surface of the particulate matter due to adsorption (Balmes, 2018; Naeher et al., 2007; W. Roberts, 2021). The emitted pollutant composition and quantity varies depending upon atmospheric conditions, the fuel source of the fire, and the flame temperature. Dispersion factors, such as surface wind speeds and unstable atmospheric conditions, influence smoke concentration; for example, high wind speeds and unstable atmospheric conditions would disperse a smoke plume over a larger area at a lower concentration than lower wind speeds with a temperature inversion (Balmes, 2018; Xu et al., 2020). Though emitted toxins such as carbon monoxide would not likely not require immediate clinical attention as they become less concentrated due to dispersion, long-term health effects from particulate matter and carcinogenic compounds found in wildfire smoke remain a cause of a concern (Adetona et al., 2016; Finlay et al., 2012; Naeher et al., 2007).

With increasing wildfires, air quality is expected to worsen during wildfire seasons. The primary pollutant of concern within wildfire smoke is a particulate less than 2.5 microns in diameter ($PM_{2.5}$) that can penetrate deep into the lungs, cross epithelial borders and enter systemic circulation for transportation throughout the body due to its size (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Reid & Maestas, 2019). Though the exact composition of $PM_{2.5}$ depends upon the emission source and atmospheric conditions, the particulates are composed of multiple chemical species, such as nitrates, organic carbon, and sulfates, and it can be coated with other chemicals, such as PAHs and VOCs (Liu & Peng, 2019; W. Roberts, 2021). Wildfire smoke events can increase the amount of ambient $PM_{2.5}$ in incident areas considerably, often exceeding the 35 μg/m^3 National Ambient Air Quality Standard (NAAQS) set for daily average $PM_{2.5}$ (Liu et al., 2015). A study observing smoke events in the Western US from 2004 to 2009 noted that over 70% of total $PM_{2.5}$ emissions on days exceeding the regulatory standards can be attributed to wildfires, indicating the significant pollution contribution by these events (Liu, Mickley, Sulprizio, Dominici, et al., 2016). Additionally, wildfire smoke events can alter the particulate composition, and the resulting $PM_{2.5}$ may be up to ten times more harmful than non-smoke $PM_{2.5}$, especially for respiratory health (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Liu & Peng, 2019). In comparison to dust particles of a similar size, wildfire-emitted particulate matter is more similar in composition to particulate matter emitted from diesel or fuels, making it more hazardous at lower exposure thresholds (Adetona et al., 2016; Naeher et al., 2007). The increased incidence of hospital admissions for asthma, bronchitis, chronic obstructive pulmonary disease (COPD), and respiratory infections during wildfire smoke events highlights this relationship between $PM_{2.5}$ exposure and negative respiratory health outcomes (Aguilera et al., 2021; Alman et al., 2016; DeFlorio-Barker et al., 2019.; Leibel et al., 2019; Reid & Maestas, 2019).

Exposure to wildfire smoke has been documented in multiple ways. There are studies that examine exposure on a continuous level, often looking at associated impacts for 10 $\mu g/m^3$ increases in $PM_{2.5}$ (Henry et al., 2021; Yao et al., 2016). However, some studies choose a binary exposure variable often title “smoke day” or “smoke wave”. These studies take into account the observed $PM_{2.5}$ values by local monitors, whether those measured values exceed $PM_{2.5}$ values that are designated by a specific threshold, and the $PM_{2.5}$ values of preceding or consecutive days (Doubleday et al., 2020; Liu et al., 2017; Liu & Peng, 2019). When selecting a threshold for a wildfire “smoke day”, some studies choose to set those through relative comparison to other daily, county-level $PM_{2.5}$ estimates; for example, the smoke wave day designation is given if the $PM_{2.5}$ values are in the top 2% of these county-level estimates (Liu et al., 2017; Liu & Peng, 2019). However, other studies choose a certain $PM_{2.5}$ level (e.g., 20.7 $\mu g/m^3$) based on observed levels in the chosen study area, and smoke day determination is given based on that threshold (Doubleday et al., 2020).

Association Between Wildfire Smoke and Health Outcomes

For people residing in areas near wildfires or for occupations involving fire response, health effects can occur due to direct contact to flames and radiant heat such as burns, trauma-related injuries, mental health impacts, and mortality (Finlay et al., 2012). In California, research from as early as 2003 depicts associations between healthcare outcomes and wildfire with increasing intensity resulting in increased incidence in hospital admission (Alman et al., 2016; Delfino et al., 2009; Flannigan et al., 2013; Gan et al., 2017; Leibel et al., 2019). A study investigating mortality from wildfire occurrence in Athens, Greece, documented increased mortality due to respiratory causes, and it also revealed a dose-response relationship that depicted increased mortality with increasing forest area burned (Analitis et al., 2012). Additionally, fires burning above 1,000,000 m^2 were associated with increases in respiratory mortality (Analitis et al., 2012).

Respiratory

On a cellular level, wildfire-smoke emitted PM affects the lungs by inducing inflammation, toxicity, and oxidative stress (Adetona et al., 2013; Holm et al., 2021). The induced oxidative stress can be attributed to the relatively elevated levels of PAHs and organic compounds emitted during wildfire events when compared to background pollution signals (Black et al., 2017). Lung function can be especially decreased for those having pre-existing respiratory conditions, such as asthma or bronchial hyperactivity (Reid et al., 2016). Multiple studies have documented the relationship between air pollution and the resulting acute health impacts after exposure, but one of the first studies examining chronic wildfire smoke impact on humans determined the presence of chronic impacts in populations after exposure to Indonesian forest fires in 1997 (Y. Kim et al., 2017). The adult men exposed to smoke during those wildfires had decreased lung function after ten years (despite accounting for temporal factors and changes), and though the exposed children appeared to have recovered their lung function after the ten years, there is still not as much documentation of pediatric health outcomes specifically attributed to wildfire smoke (Reid & Maestas, 2019).

Along with decreases in lung function, wildfire smoke events are also associated with the exacerbation of respiratory conditions ranging from asthma to upper respiratory infections. Average increases of 70 $\mu g/m^3$ $PM_{2.5}$ during heavy smoke conditions compared with $PM_{2.5}$ in the pre-wildfire period were associated with 34% increases in asthma admissions (Delfino et al., 2009). Within different age demographics, the strongest asthma associations were seen in people 65– 99 years old, 0–4 years old, and 20–64 years old, respectively (Delfino et al., 2009). A study modeling continuous PM (particulate matter) exposure reported that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ resulted in a 6% increase in physician visits for asthma exacerbations as well as increased association with COPD and lower respiratory infection visits (Yao et al., 2016). Other admissions include increases in chronic obstructive pulmonary disease (COPD) admissions for ages 20–64 years and increases in pneumonia admissions for ages 5–18 years (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Delfino et al., 2009; Reid & Maestas, 2019).

Acute bronchitis and pneumonia admissions also increased after the fires (DeFlorio-Barker et al., 2019; Delfino et al., 2009; Gan et al., 2017). In California, Delfino et al. (2009) reported that for acute bronchitis, hospital admissions across all ages increased by 9.6% in regard to increases in $PM_{2.5}$ by 10 ug/m^3, and these admissions increased after the fires as well. Overall, it can be difficult to assess the association between the broad category of "respiratory infections" and wildfire smoke as many studies vary in terms of how they group together respiratory conditions (Reid & Maestas, 2019). For example, Johnston et al. (2014) found null significance between wildfire smoke and hospitalizations for pneumonia and acute bronchitis did not align with prior papers which reported an association (Gan et al., 2017; Reid et al., 2016; Reid & Maestas, 2019). However, it was found that prior studies grouped together pneumonia and bronchitis as respiratory infections, so these discordant health outcome categorizations can lead to different conclusions when comparing overall respiratory infections associations across different studies (Reid et al., 2016; Reid & Maestas, 2019). Additionally, the intensity of the wildfire plays a role in respiratory-related health outcomes and resulting healthcare burden, as Liu et al. (2017) demonstrated that days with higher smoke wave intensity were associated with a 7.2% increase in respiratory admissions to California hospitals. Smoke wave intensity was determined by three different thresholds of 23 ug/m^3, 28 ug/m^3, and 37 ug/m^3 (Liu et al., 2017).

Multiple studies in peer-reviewed literature demonstrate a positive association between COPD and wildfire smoke exposure (Black et al., 2017; Reid et al., 2016). These studies often examine COPD exacerbation through ED visits, and at least four studies within the past decade have reported significance when examining health impacts in the western United States (Dohrenwend et al., 2013; Reid et al., 2016). In analyses of wildfire seasons in Colorado, Washington, and California there were significant associations reported between hospitalizations and/or ED visits for COPD as well as atmospheric model-derived data for $PM_{2.5}$ (Alman et al., 2016; Dohrenwend et al., 2013; Reid et al., 2016). However, due to the chronic nature of this disease, it often is not as applicable to children as evidenced by the exclusion criteria for COPD studies that often focus on adults that are well above 18 (Sutherland et al., 2005).

Dermal Conditions

Accumulating evidence also links excess PM exposure to multiple dermal conditions. PAHs and VOCs, environmental pollutants and documented carcinogens that are often emitted from industrial processes, can be coated on the surface of PM emitted during wildfire smoke events (Balmes, 2018). Due to the lipophilic properties of those carcinogenic compounds, they can surpass the skin's lipid barrier, resulting in numerous dermal effects (W. Roberts, 2021).

Along with this method, other mechanisms through which wildfire pollutants can influence dermal health are through direct accumulation on the cutaneous surface, absorption through hair follicles, and pollutants present in plasma that can diffuse into deeper levels of skin tissue (W. Roberts, 2021). The pollutants can also be inhaled, or once they have deposited on surfaces, they can be transferred to food and ingested (W. Roberts, 2021).

Frequent and chronic exposure to high levels of PM can have negative, deleterious effects on skin cells, and this can manifest through photo damage, melasma, and premature aging symptoms such as wrinkles or pigment spots (Okada et al., 2013; W. Roberts, 2021). A host of other conditions are exacerbated by outdoor air pollution, such as allergic sensitization, eczema, urticaria (hives), skin cancer, psoriasis, and acne (Bonamonte et al., 2019; Krutmann et al., 2014; Oudin et al., 2016; W. Roberts, 2021). In particular, atopic dermatitis, characterized by itchy, red skin, often starts at early stages of life (infancy or childhood) and can be further irritated by higher levels of pollutants such as VOCs, toluene, and PM (Bonamonte et al., 2019; J. Kim et al., 2019). Literature also links inflammatory acne symptoms with elevated levels of air pollution, and one documented mechanism for this pathway is through increased sebum, an oily substance secreted from sebaceous glands that contains antibacterial properties (Dréno et al., 2018; Eudier et al., 2019; Krutmann et al., 2014; Li et al., 2021). When looking at populations of acne patients, a study conducted in Mexico City determined that high levels of air pollution resulted in reduced vitamin E and squalene, and these two measures of reduced skin quality also serve as potential indicators of acne precursors (Eudier et al., 2019; Lefebvre et al., 2015). As wildfire smoke events elevate levels of ambient air pollution along with often contributing additional toxic chemicals, these dermal conditions can be exacerbated beyond the associations recorded for day-to-day pollutants (Liu & Peng, 2019; W. Roberts, 2021).

Mental Health

The act of experiencing a wildfire itself can lend to significant mental health effects as people undergo trauma, loss of property and loved ones, and displacement, and this can manifest in mental illnesses such as post-traumatic stress disorder, insomnia, and depression (Belleville et al., 2019). Children are particularly susceptible to psychological impacts, and a twenty-year follow-up study demonstrated this point by reporting that an exposure to wildfires in youth resulted in an increased likelihood of mental illness in adulthood (Brown et al., 2019; McFarlane & Van Hooff, 2009). Furthermore, wildfire events have been associated with a subsequent decrease in academic performance in children (Gibbs et al., 2019).

On a physiological level, air pollution has multiple symptoms and mechanisms through which it can affect mental health, especially through inflammation which serves as a common contributing factor to many psychiatric conditions. Post-mortem examinations have concluded that different pollutants, mostly fine and ultrafine PM, can cross the blood-brain barrier and reach the brain (Calderón-Garcidueñas et al., 2008). Additionally, air pollution has been proven to cause stress hormone increases and to alter metabolic behavior (Thomson, 2019). The most significant associations when examining mental health disorders are reported from exposure to fine PM, including $PM_{2.5}$ (Szyszkowicz et al., 2020).

While few studies have specifically examined mental health impacts of wildfire smoke, a significant body of literature from geographically diverse areas exists for traffic-related air pollutants (TRAP). A Swedish study demonstrated that children living in areas with higher air pollution levels were more likely to have medication given for a psychiatric condition during their follow-up appointments (Oudin et al., 2016). Another Toronto study reported that $PM_{2.5}$

exposure was associated with increased anxiety symptoms, and temporality was a factor as more recent exposure had a higher likelihood of relevance in displayed symptoms when compared to previously experienced exposures (Power et al., 2015). Additionally, there have been investigations regarding whether $PM_{2.5}$ can interact with chronic mental stressors to perpetuate compounded negative health impacts (Zhou et al., 2013). A study conducted in South California involving students in elementary school reported significantly higher asthma risk attributed to TRAP for children with higher parental stress when compared to their peers that had lower stress levels (Shankardass et al., 2009). Another study examining Canadian children aged 9 - 18 found an interaction between chronic stress and TRAP-attributed asthma outcomes as well (E. Chen et al., 2008).

Other Health Conditions

The evidence for an association between cardiovascular hospitalizations and wildfire $PM_{2.5}$ is mixed. A national study of over 600 counties in the United States found a statistically significant 0.61% increase in cardiovascular hospitalizations per each 10 $\mu g/m^3$ increase in $PM_{2.5}$ during wildfire season (DeFlorio-Barker et al., 2019), and another California-specific smoke study found an increase in out-of-hospital cardiac arrest risk with days that had heavier smoke (Jones Caitlin G. et al., 2020). These positive associations were seen in adult populations (ages 35 and older). However, other studies during different time periods show no association between emergency room and hospital admissions for cardiovascular health outcomes and smoke days (H. Chen et al., 2021; Reid et al., 2016). Current methodological limitations in determining wildfire smoke exposure can result in inconsistencies, potentially explaining why there are incongruous findings in reported cardiovascular associations (Yao et al., 2016). Overall, due to the inconclusive results, more granular research must be conducted in determining the impacts of cardiovascular health impacts such as heart failure, stroke, myocardial infarction, and other life-threatening outcomes (Cascio, 2018).

Apart from potential cardiovascular impacts, survivors face direct health impacts from being near to wildfire itself such as burns, mortality, mental health impacts (like post-traumatic stress disorder), or trauma (Xu et al., 2020) In addition to these direct impacts, Domitrovich & Sharkey (2010) report that individuals who spend time in areas with elevated temperature and smoke (like during wildfire events) are prone to common heat illnesses such as heat cramps, exhaustion, or heat stroke. The underlying mechanism for adverse heat-related health outcomes is thought to be due to dehydration coupled with an electrolyte imbalance, and both of these can lead to inadequate circulation in the cardiovascular system, leading to the aforementioned illnesses (Domitrovich & Sharkey, 2010).

Outcomes Observed in Washington

A few Washington State-focused studies have explored the impacts of wildfire smoke events on health outcomes. Gan et al (2017) found that during the 2012 wildfire season a 10 $\mu g/m^3$ increase in geographically weighted smoke $PM_{2.5}$ was associated with an 8% increased risk in asthma-related and a 8.4% increase in COPD-related hospital admissions (Gan et al., 2017). Specifically looking at pediatric populations, Gan et al. (2017) found a 7% increased association between smoke and all respiratory admission. (Gan et al., 2017) Findings from Harborview Medical Center have reported increased hospital emergency room visits (3 people per day with respiratory conditions), while Snohomish County reported an increase in numbers of emergency medical service call volume (DeFlorio-Barker Stephanie et al., 2019).

Equity Considerations

Wildfire events can present different impacts with varying severity based on community characteristics, and exposure assessment protocols can also pose challenges when viewing monitoring from an equity lens. When these communities are disproportionately impacted, the children in these situations may face a higher risk of adverse health outcomes from exposure. In particular, undocumented Latin and Indigenous migrants face a disproportionate exposure burden along with lower accessibility to resources during these extreme events (Méndez, 2022). During the 2017 Thomas Fire in Ventura County, California, wildfire event warning, resources, and advisories were not translated in a culturally sensitive manner (such as including mistranslations for the word "wildfire" in Spanish), and many Indigenous languages were not even included in material as certain Indigenous communities from Mexico are grouped with Hispanic communities despite speaking only their native language (Méndez, 2022). This posed a significant barrier in accessing resources, especially considering that undocumented migrant farmworkers are often at higher risk from wildfire smoke due to increased susceptibility from underlying chronic respiratory conditions, such as asthma, and increased vulnerability from working long hours outdoors. The lack of language accommodation and the political power inequity due to their undocumented status makes it difficult for the migrant workers to advocate for personal protections or time off from their occupation, drastically increasing the longevity and severity of their exposure. Additionally, undocumented status prevents migrant worker families from attaining health insurance to treat any resulting symptoms or receiving unemployment for losing their job due to the fire. Extreme events such as wildfires tend to exacerbate already present inequalities, and the Thomas Fire case study in California is just one example (Méndez, 2022).These inequalities also affect children in migrant worker families that may not be educated or have the medical resources to treat any symptoms resulting from wildfire smoke exposure.

Inequalities may also be present when collecting data regarding air quality during wildfires, especially in rural areas. Current air quality exposure methods utilize $PM_{2.5}$ monitor networks that are distributed in a spatially sparse manner when compared to the large geographic spread of smoke in rural areas (Yao et al., 2016). Urban areas often have more densely clustered monitor networks, and this can even be seen in Washington State's regulatory air quality network monitors (Washington Department of Ecology). Additionally, when measuring health outcomes, the smaller size of rural populations often cannot provide the necessary statistical power to detect significant increases in adverse health outcomes (Yao et al., 2016). Though newer exposure methods utilize air pollution models and satellite data to address these limitations, more comparative epidemiological studies must be conducted to determine the accuracy of these alternate, novel methods (Yao et al., 2016). Validation of these exposure assessment methods would allow us to better characterize population-specific risk, to understand the variability in exposure for rural populations that can often be heavily affected by smoke events, and to determine statistically significant increases in adverse health outcomes for these smaller groups.

Lack of Literature Regarding Pediatric Populations

Around the United States, there are approximately 7.4 million children annually impacted by wildfire smoke, and the Pacific Northwest contains a significant amount of that population. Children and adolescents comprise some of the most vulnerable populations to $PM_{2.5}$ and the neurological impacts of air pollution as their brains are still in development (Balmes, 2018; S.

Roberts et al., 2018). For children living in certain parts of the United States, up to 20% of the PM that they are exposed to come from wildfires (Holm et al., 2021). Their vulnerability can be attributed to increased exposure to PM from lifestyle factors (often spending more time outdoors in comparison to adults), inhaling more air relative to their body size, higher proportions of particles penetrating deeply into their lungs (due to less nasal deposition), and the ongoing developing of their respiratory and immune systems (Bennett et al., 2007; Holm et al., 2021; Vanos, 2015).

There is documented evidence of the association between $PM_{2.5}$ and adverse respiratory outcomes in children, particularly asthma exacerbations (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Leibel et al., 2019; Oliveira et al., 2019). Similar to the findings in adults, children that are closer (or directly downwind) from the fire are at highest risk, and their respiratory symptoms can manifest as upper respiratory irritations in the eyes, nose, and throat, increased visits to the hospital, and increases in infections (such as pneumonia) (Holm et al., 2021; Leibel et al., 2019). Limited pediatric research demonstrates that respiratory admissions can increase with 10-unit increases in $PM_{2.5}$ (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021) and that children from 0 – 5 years old are likely to be the most at risk as seen by their age group having the highest incidence of hospital admissions during smoke events (Leibel et al., 2019).

Aside from respiratory symptoms, a study conducted in California demonstrated that exposure to PM during the 2003 wildfire season during pregnancy can reduce newborn birthweight (Holstius et al., 2012). A retrospective cohort study conducted in China conducted by Guo et al. (2020) demonstrated that an increase in interquartile range (IQR) $PM_{2.5}$ exposure from ambient air pollution for the entire pregnancy can reduce birthweight by three grams. This is of particular concern as $PM_{2.5}$ levels can often exceed three times the regulatory limits set by the EPA during wildfire smoke season (Liu et al., 2015), which constitutes at least three months (about one third of a pregnancy period).

Though there are recent studies that have been conducted on the impact of wildfire smoke on pediatric populations (primarily in California) (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Leibel et al., 2019), more research must be conducted to characterize this association further. Though we know negative health outcomes resulting from specific air pollutants (particularly TRAP)), we would likely observe a heightened severity of adverse health effects during wildfire smoke effects due to the increase in induced oxidative stress, toxicity of pollutant types emitted during wildfire events, and the volume of $PM_{2.5}$ that is emitted. The published effect estimates on adult health would likely be very similar to the effect estimates found in children, but more research must be done to determine which specific pediatric populations are most at risk. Studies indicate that increased exposure to PM results in poor respiratory outcomes in children that can contribute to adverse, lifelong chronic conditions, but when exposure to pollutants is decreased, growth in lung function improves (Gauderman et al., 2015; Miller & Marty, 2010). This finding suggests that determining which populations are most at risk and developing interventions to decrease their exposure to PM can have an impact on child health that would improve their overall quality of life. Characterizing the risk between pediatric health and wildfire smoke events would supply us with the information we need to not only protect children during smoke events but to maintain respiratory health into adulthood.

STUDY AIMS

As Pacific Northwest residents, we recognize that wildfire incidence and severity has been increasing over the past three decades, and climate change has been proven to exacerbate wildfire events (Abatzoglou & Williams, 2016). Previous studies have confirmed a correlation between wildfire smoke events and emergency department admissions, asthma admissions, accurate bronchitis, and COPD admissions (Alman et al., 2016; Henry et al., 2021; Reid & Maestas, 2019). Despite all this knowledge on health effects, there is limited literature on pediatric health outcomes resulting from wildfire exposure, and our long-term goal is to elucidate this relationship.

The main objective of this study is to determine what health outcomes Seattle Children's Hospital reports during wildfire smoke influenced days when compared to non-wildfire smoke influenced days. Since wildfire smoke events are correlated with negative respiratory health outcomes in adults, and we know exposure responses tend to be exacerbated in children (Aguilera, Corringham, Gershunov, Leibel, et al., 2021), we predict that we will see an increase in utilization of pediatric hospital services during wildfire smoke influenced days. To evaluate the relationship between pediatric health and wildfire smoke, we will utilize three specific aims.

Specific Aim 1: *Characterize the association between wildfire smoke exposure and all same-day pediatric admission, including lag effects.*

- **Aim 1.1**: Estimate the odds ratio of overall pediatric emergency department visits and inpatient/observational hospital admissions on wildfire smoke days when compared to referent days using a case crossover study.

- **Aim 1.2:** Estimate lag effect for days 0 - 4 from exposure to $PM_{2.5}$ on wildfire smoke days for pediatric hospital services.

Specific Aim 2: *Characterize the association between wildfire smoke and pediatric health outcomes through stratification of results by individual-level traits, such as race, age group, sex, socioeconomic status, and cause of admission.*

Specific Aim 3: *Communicate findings to various audiences through oral and visual presentations.*

Since the literature base regarding pediatric outcomes is scarce, this study could significantly illuminate the acute effects of wildfire smoke influenced days on pediatric health, especially as current trends reflect an increase in wildfire severity and exposure. Determining health outcomes during these effects could allow us to shape better interventions and prophylactic activities during wildfire smoke influenced days to protect pediatric health.

AIMS 1 & 2: PEDIATRIC OUTCOMES ASSOCIATION ANALYSES

Pediatric Hospital Services Utilization Associated with Wildfire Smoke Exposure in Seattle Children's Hospital, 2006 – 2020

Daaniya Iyaz[1], Alexandra Perkins[2], Jon Cogen[2], Coralynn Sack[1], Annie Doubleday[1], María Valdez[3], Tania Busch Isaksen[1]

[1]Department of Environmental and Occupational Health Science, University of Washington
[2]Pulmonary Department, Seattle Children's Hospital
[3]Department of Biostatistics, University of Washington

Abstract

Background: Pacific Northwest wildfire smoke events have been increasing in prevalence and severity over the past three decades, resulting in documented negative health outcomes in adults. However, when examining the effect of wildfire smoke on pediatric health, the literature is scarce.

Objectives: We evaluated the association between wildfire smoke exposure and Seattle Children's Hospital service utilization, as well as impacts due to previous day exposure (lag effect).

Methods: We utilized a case crossover study to determine the odds of pediatric hospital admission at Seattle Children's Hospital on smoke days versus non-smoke days during wildfire season (June to September) in 2006 to 2020. Our analysis used a conditional logistic regression model that controls for temperature and humidity to evaluate the association between daily average $PM_{2.5}$ concentrations (based on patient zip code) and pediatric admissions for all causes and specific conditions. We also evaluated the impact of individual-level characteristics (such as age and sex) and lag effects for days 0 - 4 to further characterize the association.

Results: The results indicate a 0.0% (95% CI: -3.0%, 3.0%) change in odds for all overall same-day emergency department (ED) visits on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex, but we observed a 7.0% (95% CI: 3.0% - 12.0%) increase in odds for overall inpatient/observational hospital admissions. When stratifying by health outcomes, we found a 9.0% (95% CI:1.0% - 17.0%) and a 11.0% (95% CI:1.0% - 21.0%) increase in odds for being admitted to the ER for a respiratory and respiratory infection-related concern, respectively. We also observed statistically significant increases for inpatient/observational admissions odds, ranging from 5.0% - 8.0%, for lagged exposure on days 1 – 4 when using lag models that adjusted for Humidex. When evaluating inpatient/observational cases, we found a 44.0% (95% CI: 3.0% – 102.0%) increase in odds for trauma-related cases.

Discussion: Our results mostly aligned with other research findings on wildfire smoke impacts on health. This study reinforced the impacts of wildfire smoke on pediatric respiratory health, and it will help educate practitioners, patients, and patient families on the risks they can expect during wildfire smoke seasons in Washington state.

INTRODUCTION

In the Pacific Northwest, changes in extreme heat, precipitation, and wind patterns have led to an increase in wildfire smoke occurrence (Xu et al., 2020). Globally, factors such as forest management practices, deforestation, and increases in the interface between human residence and densely forested areas have contributed to wildfire occurrence (Adetona et al., 2016). However, multiple studies have demonstrated that climate change is a key factor in increasing wildfire occurrence and severity (Abatzoglou & Williams, 2016; Turco et al., 2018; Xu et al., 2020). Due to climate change, resulting from human behaviour, forest fires consumed an estimated 4.2 million hectares of area during the years of 1984 – 2015, and this trend is expected to persevere in the future (Abatzoglou & Williams, 2016). Large fires are almost five times more frequent in the western US on an annual level when compared to fifty years ago, and these longer-lasting fires expose people to higher levels of smoke than historically experienced (Balmes, 2018). As fire seasons become more severe (Flannigan et al., 2013), investigating the health outcomes resulting from wildfire smoke events has become a priority, especially since exposure is estimated to cost $11 – 20 billion in health care costs per year in the United States (Reid & Maestas, 2019).

Wildfire smoke contains multiple pollutants in the form of gaseous emissions, including criteria air pollutants such as carbon monoxide (caused by incomplete combustion from smoldering fuel), nitrogen dioxide, small to ultrafine particulate matter, ozone (through secondary photochemical reactions), and sulfur dioxide as well as Hazardous Air Pollutants (HAPs), such as benzene and mercury (Balmes, 2018; Evans et al., 1977; Finlay et al., 2012; Naeher et al., 2007). The smoke composition also includes carcinogenic compounds such as aldehydes, Polycyclic Aromatic Hydrocarbons (PAHs), and Volatile Organic Compounds (VOCs), and the latter two can also be found on the surface of the particulate matter due to adsorption (Balmes, 2018; Naeher et al., 2007; W. Roberts, 2021). With increasing wildfires, air quality is expected to worsen during wildfire seasons. The primary pollutant of concern within wildfire smoke is a particulate less than 2.5 microns in diameter ($PM_{2.5}$) that can penetrate deep into the lungs, cross epithelial borders and enter systemic circulation for transportation throughout the body due to its size (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Reid & Maestas, 2019). Wildfire smoke events can increase the amount of ambient $PM_{2.5}$ in incident areas considerably, often exceeding the 35 μg/m^3 National Ambient Air Quality Standard (NAAQS) set for daily average $PM_{2.5}$ (Liu et al., 2015). A study observing smoke events in the Western US from 2004 to 2009 noted that over 70% of total $PM_{2.5}$ emissions on days exceeding the regulatory standards can be attributed to wildfires, indicating the significant pollution contribution by these events (Liu, Mickley, Sulprizio, Dominici, et al., 2016). Additionally, wildfire smoke events can alter the particulate composition, and the resulting $PM_{2.5}$ may be up to ten times more harmful than non-smoke $PM_{2.5}$, especially for respiratory health (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Liu & Peng, 2019). The increased incidence of hospital admissions for asthma, bronchitis, chronic obstructive pulmonary disease (COPD), and respiratory infections during wildfire smoke events highlights this relationship between $PM_{2.5}$ exposure and negative respiratory health outcomes (Aguilera et al., 2021; Alman et al., 2016; DeFlorio-Barker et al., 2019.; Leibel et al., 2019; Reid & Maestas, 2019).

Exposure to wildfire smoke has been documented in multiple ways. There are studies that examine exposure on a continuous level, often looking at associated impacts for 10 μg/m3 increases in PM2.5 (Henry et al., 2021; Yao et al., 2016). However, some studies choose a binary exposure variable often title "smoke day" or "smoke wave". These studies take into account the observed PM2.5 values by local monitors, whether those measured values exceed PM2.5 values that are designated by a specific threshold, and the PM2.5 values of preceding or consecutive days (Doubleday et al., 2020; Liu et al., 2017; Liu & Peng, 2019). When selecting a threshold for a wildfire "smoke day", some studies choose to set those through relative comparison to other daily, county-level PM2.5 estimates; for example, the smoke wave day designation is given if the PM2.5 values are in the top 2% of these county-level estimates (Liu et al., 2017; Liu & Peng, 2019). However, other studies choose a certain PM2.5 level (e.g., 20.7 μg/m3) based on observed levels in the chosen study area, and smoke day determination is given based on that threshold (Doubleday et al., 2020).

In California, research from as early as 2003 depicts associations between healthcare outcomes and wildfire with increasing intensity resulting in increased incidence in hospital admission (Alman et al., 2016; Delfino et al., 2009; Flannigan et al., 2013; Gan et al., 2017; Leibel et al., 2019). A study investigating mortality from wildfire occurrence in Athens, Greece, documented increased mortality due to respiratory causes, and it also revealed a dose-response relationship that depicted increased mortality with increasing forest area burned (Analitis et al., 2012). Additionally, fires burning above 1,000,000 m^2 were associated with increases in respiratory mortality (Analitis et al., 2012).

Respiratory

On a cellular level, wildfire-smoke emitted PM affects the lungs by inducing inflammation, toxicity, and oxidative stress (Adetona et al., 2013; Holm et al., 2021). Lung function can be especially decreased for those having pre-existing respiratory conditions, such as asthma or bronchial hyperactivity (Reid et al., 2016). Multiple studies have documented the relationship between air pollution and the resulting acute health impacts after exposure, but one of the first studies examining chronic wildfire smoke impact on humans determined the presence of chronic impacts in populations after exposure to Indonesian forest fires in 1997 (Y. Kim et al., 2017). The adult men exposed to smoke during those wildfires had decreased lung function after ten years (despite accounting for temporal factors and changes), and though the exposed children appeared to have recovered their lung function after the ten years, there is still not as much documentation of pediatric health outcomes specifically attributed to wildfire smoke (Reid & Maestas, 2019).

Along with decreases in lung function, wildfire smoke events are also associated with the exacerbation of respiratory conditions ranging from asthma to upper respiratory infections. Average increases of 70 $\mu g/m^3$ $PM_{2.5}$ during heavy smoke conditions compared with $PM_{2.5}$ in the pre-wildfire period were associated with 34% increases in asthma admissions (Delfino et al., 2009). Within different age demographics, the strongest asthma associations were seen in people 65– 99 years old, 0–4 years old, and 20–64 years old, respectively (Delfino et al., 2009). A study modeling continuous PM (particulate matter) exposure reported that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ resulted in a 6% increase in physician visits for asthma exacerbations as well as increased association with COPD and lower respiratory infection visits (Yao et al., 2016). Acute bronchitis and pneumonia admissions also increased after the fires (DeFlorio-Barker Stephanie et al., 2019; Delfino et al., 2009; Gan et al., 2017). In California, Delfino et al. (2009) reported that for acute bronchitis, hospital admissions across all ages increased by 9.6% in regard to increases in $PM_{2.5}$ by 10 ug/m^3, and these admissions increased after the fires as well. Additionally, the intensity of the wildfire plays a role in respiratory-related health outcomes and resulting healthcare burden, as Liu et al. (2017) demonstrated that days with higher smoke wave intensity were associated with a 7.2% increase in respiratory admissions to California hospitals. Smoke wave intensity was determined by three different thresholds of 23 ug/m^3, 28 ug/m^3, and 37 ug/m^3 (Liu et al., 2017).

Dermal Conditions

Accumulating evidence also links excess PM exposure to multiple dermal conditions. Frequent and chronic exposure to high levels of PM can have negative, deleterious effects on skin cells, and this can manifest through photo damage, melasma, and premature aging symptoms such as wrinkles or pigment spots (Okada et al., 2013; W. Roberts, 2021). A host of other conditions are exacerbated by outdoor air pollution, such as allergic sensitization, eczema, urticaria (hives), skin cancer, psoriasis, and acne (Bonamonte et al., 2019; Krutmann et al., 2014; Oudin et al., 2016; W. Roberts, 2021). In particular, atopic dermatitis, characterized by itchy, red skin, often starts at early stages of life (infancy or childhood) and can be further irritated by higher levels of pollutants such as VOCs, toluene, and PM (Bonamonte et al., 2019; J. Kim et al., 2019). Literature also links inflammatory acne symptoms with elevated levels of air pollution, and one documented mechanism for this pathway is through increased sebum, an oily substance secreted from sebaceous glands that contains antibacterial properties (Dréno et al., 2018; Eudier et al., 2019; Krutmann et al., 2014; Li et al., 2021). As wildfire smoke events elevate levels of ambient air pollution along with often contributing additional toxic chemicals, these dermal conditions can be exacerbated beyond the associations recorded for day-to-day pollutants (Liu & Peng, 2019; W. Roberts, 2021).

Mental Health

Children are particularly susceptible to psychological impacts, and a twenty-year follow-up study demonstrated this point by reporting that an exposure to wildfires in youth resulted in an increased likelihood of mental illness in adulthood (Brown et al., 2019; McFarlane & Van Hooff, 2009). Furthermore, wildfire events have been associated with a subsequent decrease in academic performance in children (Gibbs et al., 2019). While few studies have specifically examined mental health impacts of wildfire smoke, a significant body of literature from geographically diverse areas exists for traffic-related air pollutants (TRAP). A Swedish study demonstrated that children living in areas with higher air pollution levels were more likely to have medication given for a psychiatric condition during their follow-up appointments (Oudin et al., 2016). Another Toronto study reported that $PM_{2.5}$ exposure was associated with increased anxiety symptoms, and temporality was a factor as more recent exposure had a higher likelihood of relevance in displayed symptoms when compared to previously experienced exposures (Power et al., 2015).

Outcomes Observed in Washington

A few Washington State-focused studies have explored the impacts of wildfire smoke events on health outcomes. Gan et al (2017) found that during the 2012 wildfire season a 10 $\mu g/m^3$ increase in geographically weighted smoke $PM_{2.5}$ was associated with an 8% increased risk in asthma-related and a 8.4% increase in COPD-related hospital admissions (Gan et al., 2017). Specifically looking at pediatric populations, Gan et al. (2017) found a 7% increased association between smoke and all respiratory admission. (Gan et al., 2017) Findings from Harborview Medical Center have reported increased hospital emergency room visits (3 people per day with respiratory conditions), while Snohomish County reported an increase in numbers of emergency medical service call volume (DeFlorio-Barker et al., 2019).

Pediatric-Specific Outcomes

Around the United States, there are approximately 7.4 million children annually impacted by wildfire smoke, and the Pacific Northwest contains a significant amount of that population. Children and adolescents comprise some of the most vulnerable populations to $PM_{2.5}$ and the neurological impacts of air pollution as their brains are still in development (Balmes, 2018; S. Roberts et al., 2018). For children living in certain parts of the United States, up to 20% of the PM that they are exposed to come from wildfires (Holm et al., 2021). Their vulnerability can be attributed to increased exposure to PM from lifestyle factors (often spending more time outdoors in comparison to adults), inhaling more air relative to their body size, higher proportions of particles penetrating deeply into their lungs (due to less nasal deposition), and the ongoing developing of their respiratory and immune systems (Bennett et al., 2007; Holm et al., 2021; Vanos, 2015).

There is documented evidence of the association between $PM_{2.5}$ and adverse respiratory outcomes in children, particularly asthma exacerbations (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Leibel et al., 2019; Oliveira et al., 2019). Similar to the findings in adults, children that are closer (or directly downwind) from the fire are at highest risk, and their respiratory symptoms can manifest as upper respiratory irritations in the eyes, nose, and throat, increased visits to the hospital, and increases in infections (such as pneumonia) (Holm et al., 2021; Leibel et al., 2019). Limited pediatric research demonstrates that respiratory admissions can increase with 10-unit increases in $PM_{2.5}$ (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021) and that children from 0 – 5 years old are likely to be the most at risk as seen by their age group having the highest incidence of hospital admissions during smoke events (Leibel et al., 2019).

Though there are recent studies that have been conducted on the impact of wildfire smoke on pediatric populations (primarily in California) (Aguilera, Corringham, Gershunov, & Benmarhnia, 2021; Leibel et al., 2019), more research must be conducted to characterize this association further. The published effect estimates on adult health would likely be very similar to the effect estimates found in children, but more research must be done to determine which specific pediatric populations are most at risk. Studies indicate that increased exposure to PM results in poor respiratory outcomes in children that can contribute to adverse, lifelong chronic conditions, but when exposure to pollutants is decreased, growth in lung function improves (Gauderman et al., 2015; Miller & Marty, 2010). This finding suggests that determining which populations are most at risk and developing interventions to decrease their exposure to PM can have an impact on child health that would improve their overall quality of life. Characterizing the risk between pediatric health and wildfire smoke events would supply us with the information we need to not only protect children during smoke events but to maintain respiratory health into adulthood.

METHODS

Health Outcome Data

The University of Washington's Institutional Review Board (IRB) reviewed this study and determined it exempt from further review on June 7, 2021. The Seattle Children's IRB board approved this study on June 3, 2021. The health outcomes in this study were sourced from Pediatric Health Information System (PHIS) dataset (https://www.childrenshospitals.org/phis), and provided to the study team by Seattle Children's Hospital. The PHIS dataset is comprised of deidentified patient data from 51 children's hospitals in the United States regarding diagnoses, length of stay, demographic information (such as age, sex, and race), and billing (https://www.childrenshospitals.org/phis). This study obtained PHIS data for all-cause ED visits and inpatient/observational admissions from only Seattle Children's main campus hospital. Each patient was provided with a unique ID code to ensure de-identification, and each clinical encounter was also given its own unique ID code. Our study period is confined to June 1st to September 30st from 2006 to 2020. Data uses . The utilized variables included date of the medical visit, type of service provided (ED, inpatient, or observational), and patient diagnosis (reported as an All Patient Refined Diagnostic Related Group). The PHIS dataset reports discharge APR-DRG codes which are more accurate than admission APR-DRGs because the principal diagnosis of the patient has been determined, secondary diagnoses have been established, and any procedures or tests that occur during the patient's admission are documented (Appendix A, Table A1). Individual-level variables consisted of age, sex, race (either assigned by hospital staff or reported by the patient/patient guardian during the clinical encounter), and insurance type. Other included variables that were not utilized in this analysis were length of stay, mechanical ventilation flag, intensive care unit (ICU) flag, cost of the visit, and International Classification of Disease (ICD)-9 and ICD-10 codes. The ICD codes were not used to determine cause of the hospital visit as the study period stretched over the switch from ICD-9 to ICD-10 which occurred in 2015; therefore, using the APR-DRG code allowed us to view all clinical encounter diagnoses in a uniform manner across the study period, allowing for comparison between all years. Cases were included if $\geq$ 19 years old with a Washington state residential zip code. We excluded any patients who had a COVID-19 cause of admission because at the time of methods development, research was unclear regarding the impacts of wildfire smoke on COVID-19, and we were uncertain if patients visiting the ED had positive COVID-19 tests (or if they had merely come for testing purposes). We also omitted patients admitted on the fourth or fifth of July due to confounding levels of air pollution resulting from fireworks.

Our age exclusion criteria was based on other pediatric wildfire studies (Aguilera, Corringham, Gershunov, Leibel, et al., 2021; Henry et al., 2021) and with input from pediatricians and pulmonologists. Additionally, due to the differences in hospital services and case severity, we chose to split our health outcome data into two groups for analysis: ED visits and inpatient/observational admissions. For each service type, we stratified the data by age categories (0 – 5, 6 – 12, 13 – 19) derived from literature (Aguilera, Corringham, Gershunov, Leibel, et al., 2021), sex (male or female), race group as reported in PHIS ("Non-Hispanic White", "Non-Hispanic Black", "Hispanic", "Asian", "Other"), insurance type as reported in PHIS ("Government", "Private", or "Other"), and clinical encounter cause categories (respiratory, respiratory infections, dermal conditions, trauma, and mental health). We defined the clinical encounter cause categories based on established literature regarding the impacts of wildfire smoke on all aged populations (as cited in the background section) and from physiologically plausible links suggested by practicing pediatricians and pulmonologists. The full-list of APR-DRG conditions included in each clinical encounter cause category can be found in Appendix A (Table A1).

Exposure Classification

We assessed exposure for cases using an exposure grid as described previously (Doubleday et al., 2020). $PM_{2.5}$ and meteorological variables (temperature and relative humidity) were collected from 68 monitors across the state (Doubleday et al., 2020). These federal and state-run monitors in Washington State track various pollutant concentrations ($PM_{2.5}$,

PM_{10}, ozone, NO_2, SO_2, and carbon monoxide) and meteorology (daily relative humidity and daily maximum temperature) averaged over 24 hours (Washington State Department of Ecology, 2022). These two values were used to calculate Humidex, a measure of perceived temperature. Data collected from the monitors were combined with the AIRPACT (Air Indicator Report for Public Awareness and Community Tracking) model to create a 4x4 kilometer exposure grid of 24-hour mean $PM_{2.5}$ via methods modeled in Doubleday et al., 2020. In order to derive the population-weighted average, the 4x4km grid was overlaid with zip code-level population data to determine what percentage of the population resided in each grid cell per zip code. To calculate full exposure to $PM_{2.5}$ per zip code, each grid cell's population percentage was multiplied by daily average $PM_{2.5}$ exposure values and summed across all grid cells in that zip code. Areas where more people lived were weighted more to prevent over-estimating population exposure in areas with less residents. For grid cells missing exposure data, we utilized the nearest zip code neighbor values as an estimate through methods reported in Doubleday et al., 2020. We joined the exposure dataset with the PHIS dataset using R 4.0.2 (R Core Development Team, 2020) and assigned each clinical encounter a corresponding 24-hour average $PM_{2.5}$ concentration and Humidex value.

Wildfire "Smoke Day" Classification

This study uses a binary "smoke day" classification based on met threshold conditions from Doubleday et al. (2020). They defined a wildfire smoke day as "[zip code] days with a 24-hour average $PM_{2.5}$ concentration greater than 20.4 $\mu g/m^3$". In areas where wildfire smoke events may have lower $PM_{2.5}$ concentrations, they applied an additional set of criteria for smoke day classifications. The criteria are as follows: "For the [zip code] days between 9 and 20.4 $\mu g/m^3$…:

1. The day must be part of an event in which 2 of 3 consecutive days are greater than 9 $\mu g/m^3$
2. One of the days in the event window must be greater than 15 $\mu g/m^3$
3. For urban areas (Seattle, Tacoma, Spokane), at least 50% of the air monitors in those areas must be greater than 9 $\mu g/m^3$" (Doubleday et al., 2020).

We used this smoke day classification criteria to determine if the 24-hour average $PM_{2.5}$ concentration in our exposure dataset constituted a smoke day. The smoke day threshold is based on how relative the recorded 24-hour average $PM_{2.5}$ levels were to background anthropogenic PM levels across the study period and if the PM concentration would likely cause adverse health effects (based on background literature findings). This classification method was chosen because it reduces smoke-related exposure misclassification in urban areas that report higher levels of background PM concentration and in rural areas which typically have lower background PM concentrations. The resulting exposure assessment is a more accurate binary classification of wildfire smoke affected day (Doubleday et al., 2020). We utilized R 4.0.2 (R Core Development Team, 2020) to assess exposure for each patient encounter.

Statistical Analysis

This study utilized a case-crossover design to study the impacts of acute $PM_{2.5}$ exposure and effect modification on the subsequent risk of a pediatric hospital encounters. In case-cross over design, each case serves as their own control by comparing the "at risk" exposure period with the level of exposure during a referent period (i.e., a period during which the patient was exposed but did not result in a hospital visit) (Levy et al., 2001). These "referent windows" create within-subject comparisons that control for time-invariant confounders (Janes et al., 2005). To adequately control for time-dependent confounders (like long-term time trends, seasonality (variations due to season), and day-of-week impacts), we utilized a bidirectional, time-stratified design to match the referent days by the same day of the week as the "case" day within the same month and year (Hutchinson et al., 2018; Janes et al., 2005; Levy et al., 2001). This matching method yields multiple referent days within the same month. We removed any referent days that overlapped as a case day if a patient was admitted multiple times within the same month and year. We also removed any referent days on the fourth or fifth of July due to confounding levels of air pollution resulting from fireworks. Our only covariate was Humidex, and we controlled for its impact on our effect estimates using a spline with three degrees of freedom as documented in Doubleday et al. (2020) . We report our results as an odds ratio, or a percent difference in odds of a pediatric hospital encounter at Seattle Children's on wildfire smoke days versus non-smoke days, after adjusting for Humidex. To obtain the odds ratios, we used conditional logistic regression with 95% confidence intervals with the *clogit* function within the *survival* R package (R Core Development Team, 2020). We define statistical significance as any result that has a p-value (p-val) of less than or equal to 0.05, referring to a 95% confidence interval (95% CI).

To further characterize the association between wildfire smoke and pediatric health, we stratified the results by age group, sex, race category, insurance type, and cause of hospital encounter. We conducted a lag analysis by comparing cases to smoke day exposure from days 0 – 4. For same-day exposures, we reported our findings by age group and sex within the cause of admission categories (respiratory, respiratory infections, dermal conditions, trauma, and mental health).

RESULTS

Population

Between June to September from 2006 to 2020, there were a total of 194,887 pediatric hospital encounters. 132,408 (67.94%) of them were visits to the ED, and 62,479 (32.06%) of them were inpatient or observational admissions to the hospital (Table 1). Both ED and inpatient/observational patients were more likely to be male (~53% versus ~46% female for both categories). In terms of age categories, most patients were in the youngest category (0 – 5). For ED visits, the breakdown was 56.9% for ages 0 – 5, 27.2% for ages 6 – 12, and 15.9% for ages 13 – 19 (Table 1). In terms of inpatients/observational visits, 46.4% were ages 0 – 5, 25.8% were ages 6 – 12, and 27.7% were ages 13 – 19 (Table 1). Regarding races seen in the PHIS dataset, the most common race was Non-Hispanic White, followed by Other and Hispanic (Table 1). For ED visits, the next most common race was Non-Hispanic Black (10.5%) followed by Asian (7.03%). For inpatient/observation admissions, Asian was the next most common race (5.91%) followed by Non-Hispanic Black (5.69%). The most common insurance type for ED visits was government (49.3%) and for inpatients/observational admissions, private (48.8%). When looking at health outcome categories, ED visits had trauma as the most common outcome (19.22%), followed by overall respiratory conditions (15.60%) and respiratory infections (10.35%). Inpatient/observational admissions had respiratory conditions as the highest outcome (10.09%), followed by mental health (6.55%) and respiratory infections (3.45%) (Table 1). In terms of location, for ED visits, most patients were from King County (75.3%), Snohomish County 19.4%), and Pierce County (1.5%) as seen by the case density breakdown in Figure 1. For a closer look at Seattle (King County) zip code density, please see Figure A1 in Appendix A. For inpatient/observational admissions, cases were a little more distributed, but most patients were still from King County (47.2%), Snohomish County (19.4%), and Pierce County (6.2%), followed by the counties of Yakima (4.7%), Whatcom (3.1%), Skagit (2.5%), Thurston (2.5%), and Kitsap (2.4%) as seen in Figure 2. For a zoomed in view of King and Snohomish County zip codes, please see Figure A2 in Appendix A.

Table 1. Case characteristics separated by ED visits (left column) and Inpatient/Observational (abbreviated as Inpt./Obvs.) admissions (right column).

		ED	Inpt./Obvs.
Characteristic		**N of Total Cases (%)**	**N of Total Cases (%)**
All Cases		132,408 (67.94)	62,479 (32.06)
Age group (years)		**n (%)**	**n (%)**
	0 – 5	75,328 (56.9)	29,005 (46.4)
	6 – 12	36,079 (27.2)	16,145 (25.8)
	13 – 19	21,001 (15.9)	17,329 (27.7)
Sex			
	Male	71,441 (53.96)	33,661 (53.88)
	Female	60,964 (46.04)	28,807 (46.11)
Race			
	Non-Hispanic White	48,317 (36.5)	27,735 (44.4)
	Non-Hispanic Black	13,879 (10.5)	3,552 (5.69)
	Hispanic	28,388 (21.4)	11,322 (18.1)
	Asian	9,312 (7.03)	3,694 (5.91)
	Other	32,512 (24.6)	16,176 (25.9)
Insurance Type			
	Government	65,252 (49.3)	28,912 (46.3)
	Private	62,325 (47.1)	30,483 (48.8)
	Other	4,831 (3.65)	3,084 (4.94)
Health Outcomes			
	All Cases	132,408 (100)	62,479 (100)
	Respiratory	20,653 (15.60)	6,304 (10.09)
	Resp. Infections	13,710 (10.35)	2,147 (3.45)
	Dermal Conditions	5,404 (4.08)	346 (0.55)
	Trauma	25,451 (19.22)	1,020 (1.63)
	Mental Health	4,203 (3.17)	4,093 (6.55)

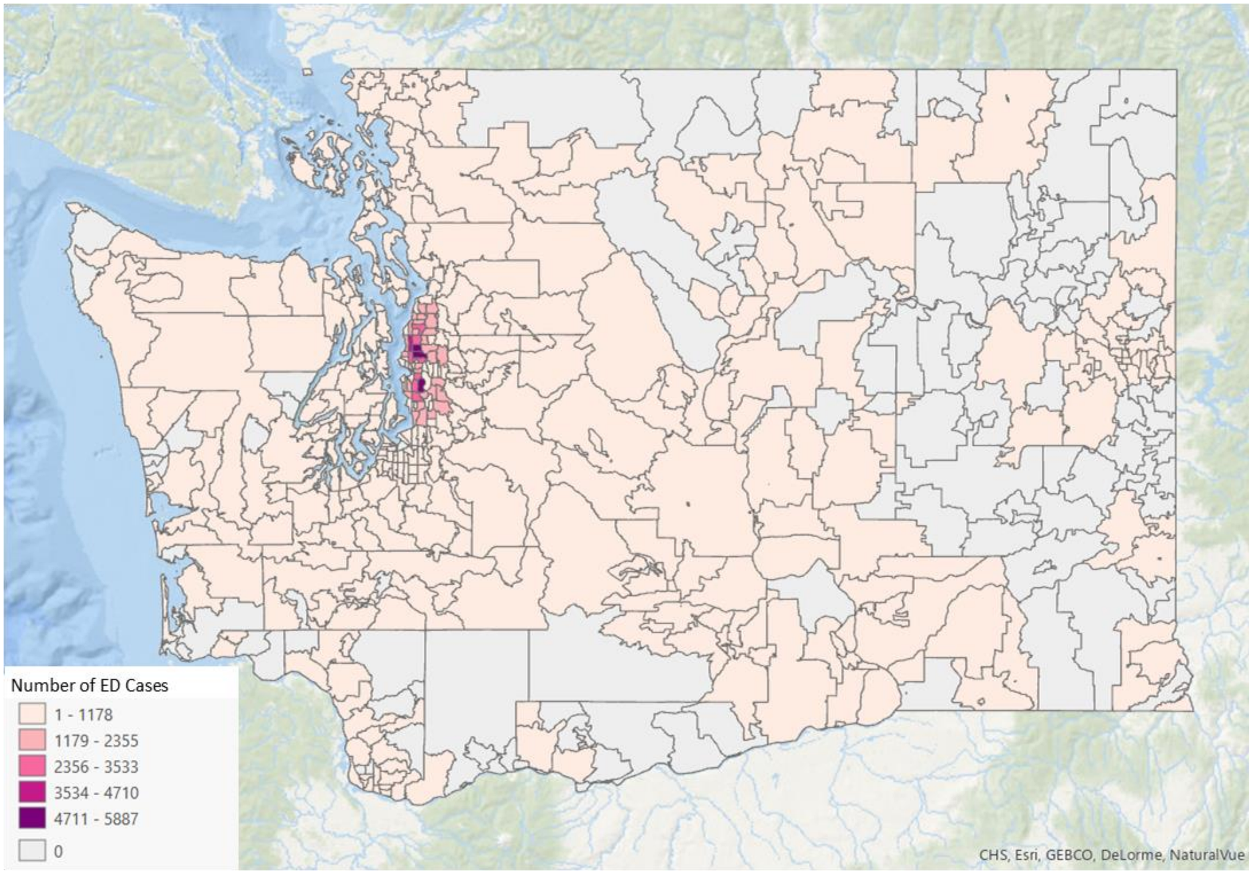

Figure 1. Case distribution for all ED cases (132,408) in the study dataset.

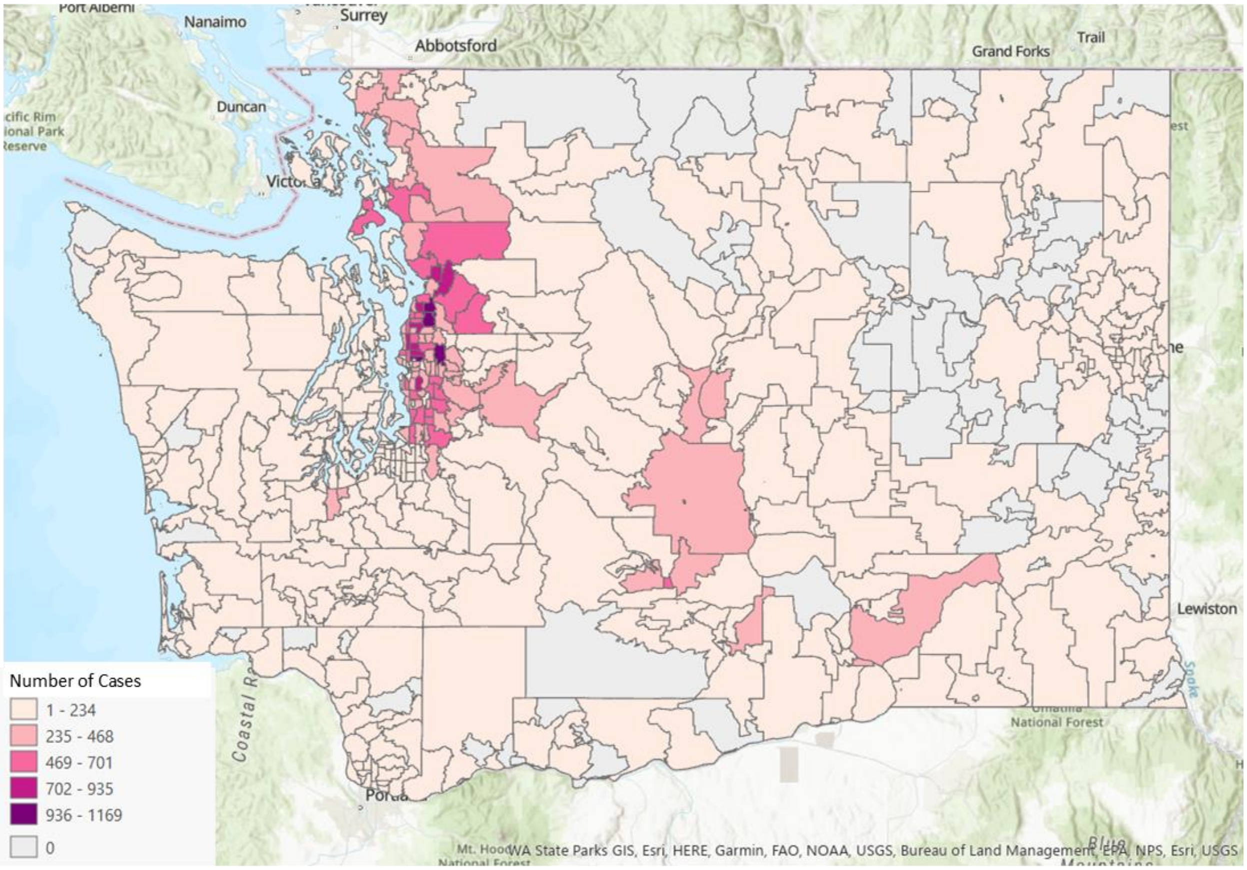

Figure 2. Case distribution for all inpatient/observational cases (62,479) in the study dataset.

Exposure

Table 2 presents a summary of the exposure variables ($PM_{2.5}$ and Humidex) for both ED and inpatient/observational hospital visits. For ED visits, the average $PM_{2.5}$ concentration on case days was 6.67 μg/m^3 (SD: 9.50) while the average $PM_{2.5}$ concentration on referent days was also 6.67 μg/m^3 (SD: 9.59). When comparing the distribution of daily average $PM_{2.5}$ using quantiles, the 75% quantile was 0.07 μg/m^3 higher for case days versus referent days. For inpatient/observational visits, the average $PM_{2.5}$ concentration on case days was 6.73 μg/m^3 (SD: 12.36) while the average $PM_{2.5}$ concentration on referent days was 6.58 μg/m^3 (SD: 10.63). When comparing the distribution of daily average $PM_{2.5}$ using quantiles, the 75% quantile was 0.15 μg/m^3 higher for case days versus referent days. The maximum concentration of $PM_{2.5}$ was always higher for case days regardless of ED or inpatient/observational status. In the entire dataset (all case and referent days from 2006 – 2020 that met the exclusion criteria), there were 32,950 wildfire smoke days. For wildfire smoke days in the ED dataset, the average $PM_{2.5}$ value was 29.78 µg/m^3 (SD: 30.58) and the average Humidex value was 28.18 ˚C (SD: 5.18). Non-wildfire smoke days reported an average $PM_{2.5}$ value of 5.27 µg/m^3 (SD: 2.49) and an average Humidex value of 23.31 ˚C (SD: 4.91). In the inpatient/observation dataset, wildfire smoke days had an average $PM_{2.5}$ value of 31.30 µg/m^3 (SD: 35.89) and an average Humidex value of 29.34 ˚C (SD: 5.31). Non-wildfire smoke days reported an average $PM_{2.5}$ value of 5.03 µg/m^3 (SD: 2.46) and an average Humidex value of 23.88 ˚C (SD: 5.24).

Table 2. Exposure characteristics (daily average $PM_{2.5}$ and Humidex) values for case and referent days reported in frequency, quantile, mean, and standard deviation (SD). Separated by ED visits and Inpatient/Observational (abbreviated as Inpt./Obvs.) admissions.

	ED			**Inpt./Obvs.**		
Characteristic	**Number**			**Number**		
Total Event Days	132,408			62,479		
Total Referent Days	444,590			208,898		
Exposure Values	**PM2.5 (µg/m^3) Mean (SD)**	**Quantile (µg/m^3) 25%, 75%**	**Maximum Concentra -tion**	**PM2.5 (µg/m^3) Mean (SD)**	**Quantile (µg/m^3) 25%, 75%**	**Maximum Concentra -tion**
Case Days	6.67 (9.50)	3.57, 7.17	324.49	6.73 (12.36)	3.36, 6.90	1094.15
Referent Days	6.67 (9.59)	3.55, 7.10	254.94	6.58 (10.63)	3.35, 6.85	523.69
Exposure Metrics	**Number**	**PM2.5 (µg/m^3) Mean (SD)**	**Humidex (C˚) Mean (SD)**	**Number**	**PM2.5 (µg/m^3) Mean (SD)**	**Humidex (C˚) Mean (SD)**
Wildfire Smoke Days	32,950	29.78 (30.58)	28.18 (5.18)	16,332	31.30 (35.89)	29.34 (5.31)
Non-Wildfire Smoke Days	544,048	5.27 (2.49)	23.31 (4.91)	255,045	5.03 (2.46)	23.88 (5.24)

All-cause and Lag

Table 3 depicts the results of the statistical analysis for both ED and inpatient/observational datasets. The results indicate a 0.0% (95% CI: -3.0% - 3.0%) change in odds for all overall same-day ED visits on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex, but we observed a 7.0% (95% CI: 3.0% - 12.0%) increase in odds for overall inpatient/observational hospital admissions under the same conditions. This finding can also be visualized in Figure A3 in Appendix A. When modelling lag of daily average $PM_{2.5}$ exposure levels one to four days before a case encounter, we did not observe any statistically significant changes in odds for ED visits (as seen in Figure 1). However, when evaluating the inpatient/observational admissions, we observed statistically significant increases in odds in every day of lag exposure (days 1 - 4) when adjusting for Humidex (Figure 3 and Table 4). The highest increase in odds we observed was an 8.0% (95% CI: 4.0% - 12.0%) elevation in odds from the null for inpatient/observational patients that were exposed one day prior (1 day lag) to wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. The statistically significant increase in odds continued throughout the lag days we tested for, and the results are summarized in Figure 3 and Table 4.

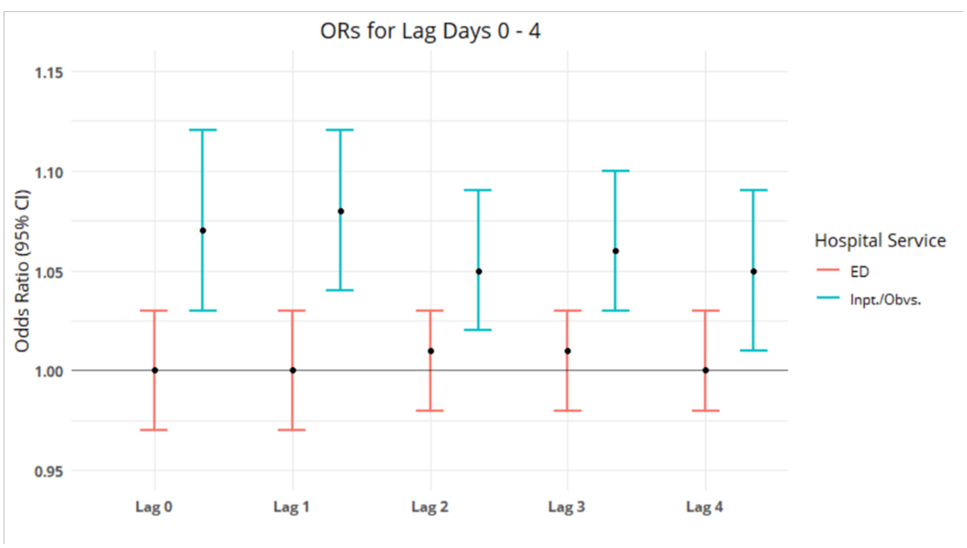

Figure 3. ORs from the lag model for days 0 - 4.

Individual-level Characteristics

We conducted additional same-day analyses by stratifying further by individual-level characteristics, specifically age, sex, and insurance type. These results are summarized in Table 3. In terms of age groups, the only statistically significant results were seen for the inpatient/observational admissions category. Ages 0 – 5 and 6 – 12 saw a 6.0% (95% CI: 0.0% - 13.0%) and 8.0% (95% CI: 0.1% - 12.0%) increase in odds of admission on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex, respectively. There was also an 8.0% (95% CI: 0.0% - 16.0%) increase in odds under the same conditions for ages 13 – 19, but the finding was not statistically significant. For ED visits, only ages 6 – 12 had an odds ratio that deviated from the null indicating a 2.0% (95% CI: -4.0% - 7.0%) non-statistically significant increase in odds for ED visits on wildfire smoke days versus non-wildfire smoke days, when controlling for Humidex. In terms of sex, only inpatient/observational admissions deviated from the null, and female cases were found to have a 10.0% (95% CI: 4.0% - 17.0%) increase in odds of admissions on wildfire smoke days versus non-wildfire smoke days, when controlling for Humidex. Male cases were found to have a 5.0% (95% CI: -1.0% - 11.0%) non-statistically significant increase in odds under the same conditions. Similar to sex, only the inpatient/observational admissions had statistically significant results regarding race. The non-Hispanic White population reported a 7.0% (95% CI: 4.0% - 18.0%) increase in odds for inpatient/observational admissions on wildfire smoke days compared to non-smoke days, while the non-Hispanic Black population observed a 30.0% (95% CI: 11.0% - 53.0%) increase in odds. For ED visits, non-Hispanic Black cases observed a 2.0% (95% CI: -6.0% - 11.0%) non-statistically significant increase in odds wildfire smoke days versus non-wildfire smoke days. Hispanic cases also observed a 2.0% (95% CI: -4.0% - 9.0%) non-statistically significant increase in odds for ED visits and a 7.0% (95% CI: -2.0% - 16.0%) increase in odds for inpatient/observational admissions under the same conditions. Asian populations were associated with a non-statistically significant decrease in odds for both hospital services. For insurance type, the only statistically significant results were for inpatient/observation cases which had a 7.0% increase in odds for both government (95% CI: 1.0% - 13.0%) or private insurance (95% CI: 0.3% - 13.0%) on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. These findings can also be visualized in figures A4 – A7 in Appendix A.

We also stratified the results by cause of admission (from APR-DRG codes). The only statistically significant results for ED visits were for overall respiratory cause of admission that estimated a 9.0% (95% CI: 1.0% - 17.0%) increase in odds for presenting to the ER for a respiratory-related concern on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. Similarly, the respiratory sub-category of respiratory infections reported a 11.0% (95% CI: 1.0% - 21.0%) increase in odds under the same conditions. When evaluating inpatient/observational cases, the only statistically significant finding was a 44.0% (95% CI: 3.0% - 102.0%) increase in odds for trauma-related cases on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. These findings can also be visualized in Figure A8 in Appendix A.

Table 3. ORs separated by ED visits and Inpatient/Observational admissions. Bolded values are statistically significant findings.

	ED			Inpt./Obvs.		
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
All Cases (no Lag)	1.00 (0.97, 1.03)	0.974	132,408 (67.94)	**1.07 (1.03, 1.12)**	**0.001**	62,479 (32.06)
Age group (years)						
0 – 5	1.00 (0.96, 1.04)	0.915	75,328 (56.9)	**1.06 (1.00, 1.13)**	**0.046**	29,005 (46.4)
6 – 12	1.02 (0.96, 1.07)	0.613	36,079 (27.2)	**1.08 (1.00, 1.12)**	**0.047**	16,145 (25.8)
13 – 19	1.00 (0.93, 1.07)	0.932	21,001 (15.9)	1.08 (1.00, 1.16)	0.065	17,329 (27.7)
Sex						
Male	1.00 (0.96, 1.04)	0.956	71,441 (53.96)	1.05 (0.99, 1.11)	0.112	33,661 (53.88)
Female	1.00 (0.96, 1.04)	0.996	60,964 (46.04)	**1.10 (1.04, 1.17)**	**0.002**	28,807 (46.11)
Race						
Non-Hispanic White	1.00 (0.96, 1.05)	0.957	48,317 (36.5)	**1.11 (1.04, 1.18)**	**0.001**	27,735 (44.4)
Non-Hispanic Black	1.02 (0.94, 1.11)	0.615	13,879 (10.5)	**1.30 (1.11, 1.53)**	**0.001**	3,552 (5.69)
Hispanic	1.02 (0.96, 1.09)	0.441	28,388 (21.4)	1.07 (0.98, 1.16)	0.148	11,322 (18.1)
Asian	0.98 (0.88, 1.09)	0.741	9,312 (7.03)	0.86 (0.71, 1.03)	0.097	3,694 (5.91)
Other	0.97 (0.91, 1.04)	0.380	2,512 (24.6)	1.00 (0.91, 1.09)	0.932	16,176 (25.9)
Insurance Type						
Government	1.01 (0.97, 1.05)	0.819	65,252 (49.3)	**1.07 (1.01, 1.13)**	**0.031**	28,912 (46.3)
Private	0.99 (0.95, 1.04)	0.765	62,325 (47.1)	**1.07 (1.00, 1.13)**	**0.039**	30,483 (48.8)
Other	1.05 (0.90, 1.23)	0.529	4,831 (3.65)	1.23 (0.99, 1.51)	0.058	3,084 (4.94)
Health Outcomes						
Respiratory	**1.09 (1.01, 1.17)**	**0.029**	20,653 (15.60)	0.97 (0.84, 1.11)	0.645	6,304 (10.09)
Resp. Infections	**1.11 (1.01, 1.21)**	**0.033**	13,710 (10.35)	1.14 (0.89, 1.45)	0.306	2,147 (3.45)
Dermal Conditions	1.10 (0.95, 1.27)	0.199	5,404 (4.08)	1.61 (0.88, 2.94)	0.126	346 (0.55)
Trauma	0.99 (0.92, 1.05)	0.652	25,451 (19.22)	**1.44 (1.03, 2.02)**	**0.034**	1,020 (1.63)
Mental Health	0.94 (0.80, 1.10)	0.433	4,203 (3.17)	1.15 (0.98, 1.35)	0.081	4,093 (6.55)
All Cases	1.00 (0.97, 1.03)	0.974	132,408 (100)	**1.07 (1.03, 1.12)**	**0.001**	62,479 (100)

Table 4. ORs for lag analysis separated by ED visits and Inpatient/Observational admissions. Bolded values are statistically significant findings.

	ED			**Inpt./Obvs.**		
Lag	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
1 Day	1.00 (0.97, 1.03)	0.928	130,128 (98.28)	**1.08 (1.04, 1.12)**	**0.000**	61,426 (98.32)
2 Day	1.01 (0.98, 1.03)	0.672	127,798 (96.52)	**1.05 (1.02, 1.09)**	**0.005**	60,472 (96.79)
3 Day	1.01 (0.98, 1.03)	0.558	125,665 (94.91)	**1.06 (1.03, 1.10)**	**0.001**	59,552 (95.32)
4 Day	1.00 (0.98, 1.03)	0.841	123,601 (93.35)	**1.05 (1.01, 1.09)**	**0.005**	58,588 (93.77)

We also conducted additional secondary analyses where we further stratified by age group and sex within each health outcome category, and the results can be seen in Figure 2 and 3 below (respectively). These sub-analyses did not reveal many statistically significant deviations away from the null. For same-day ED visits, the only significant age-related result we observed was a 12.0% (95% CI: 0.4% - 24.0%) increase in odds for ages 0 – 5 that visited the ED for respiratory infections on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex (Figure 3). Regarding same-day inpatient/observational cases, there were three documented statistically significant increases in odds for age categories. For 13 – 19-year-old children, there was a 136.0% (95% CI: 27.0% - 339.0%) increase in odds when admitted for a respiratory condition and a 113.0% (95% CI: 13.0% - 239.0%) increase in odds when admitted for a trauma-related condition on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex (Figure 3). For the 0 – 5 age group, odds of trauma-related admissions, increased by 67.0% (95% CI: 2.0% - 175.0%) (Figure 2). With respect to patient sex, only female patients visiting the ER for two specific health outcomes resulted in statistically significant increases in odds. We documented a 13.0% (95% CI: 0.4% - 27.0%) increase in odds for female respiratory-related ED visits and a 17.0% (95% CI: 2.0% - 35.0%) increase in odds for female respiratory infection-related ED visits on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex (Figure 3). We also looked into lag analyses for each of the health outcome strata, but we did not find any statistically significant results except for a decreased odds for lag days 1 – 4 when looking at mental health-related visits to the ED. However, the same-day mental health-related ED visits association was not statistically significant (Table 3). The full results can be seen in Table A2 in Appendix A.

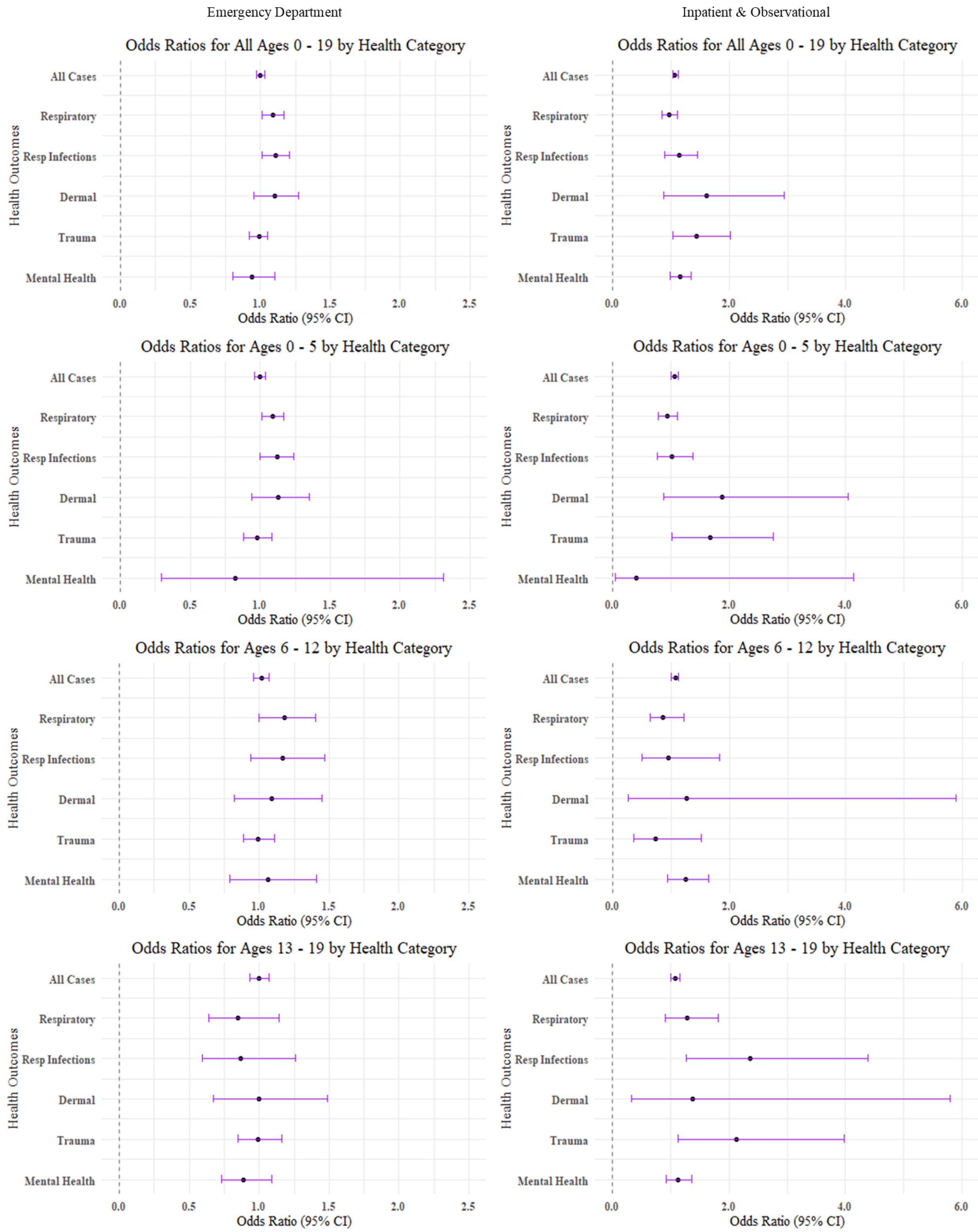

Figure 4. ORs for each age category stratified by health outcome and separated by ED cases (left column) and inpatient/observational cases (right column).

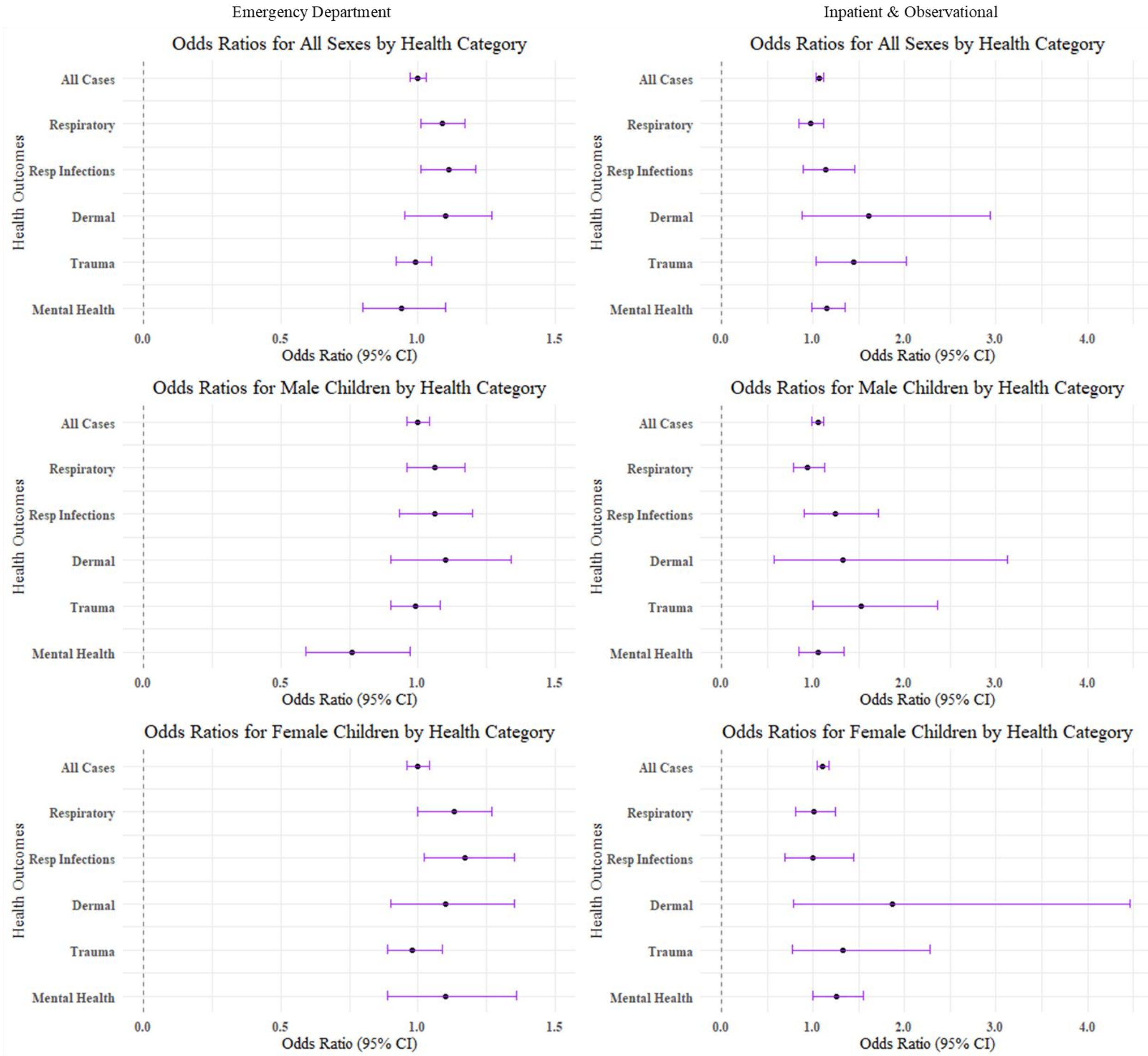

Figure 5. ORs for sex stratified by health outcome, by ED cases (left column) and inpatient/observational cases (right column).

.

DISCUSSION

In this case-crossover study, we analyzed data consisting of 132,408 ED visits and 62,479 inpatient/observational hospital admissions to Seattle Children's Hospital, a single center, tertiary pediatric hospital in Washington state, from 2006 – 2020. We found that wildfire smoke exposure was associated with an increased odds of respiratory-related ED visits and all-cause hospital admissions. The risk remained elevated with exposure to wildfire smoke up to 4 days prior to inpatient and observational hospital admission. Though other studies have investigated the impact between wildfire smoke and overall health, relatively few studies have specifically focused on children (Leibel et al., 2019). Furthermore, there are even fewer studies that investigate health impacts in the state of Washington which is an area of pronounced wildfire occurrence and smoke impacts (Gan et al., 2017). To our knowledge, this is the first study to focus on the risks of wildfire smoke exposure in a pediatric population in Washington state. .

Though our OR estimate for overall ED visits did not deviate from the null hypothesis, when we stratified by respiratory-related health outcomes, we found a 9.0% (95% CI: 1.0% - 17.0%) increase in odds for presenting to the ED for a respiratory-related concern on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. When focusing specifically on respiratory infections, we found an 11.0% (95% CI: 1.0% - 21.0%) increase in odds under the same conditions. Both statistically significant findings align with other studies, both adult and pediatric-specific, which report associations between wildfire smoke and negative respiratory health outcomes. For example, our results are similar to a study conducted in California with Medicare enrollees (aged 65 and above) which reported a 7.2% (95% CI: 0.25%

– 15.0%) increase in emergency respiratory admissions on wildfire smoke-wave-days (wildfire-specific $PM_{2.5}$ >37 μg/m^3) versus matched non-smoke-wave days (Liu et al., 2017).

Since our APR-DRG respiratory-related outcome groups included asthma, upper respiratory tract infections, and pneumonia (Table A1), we compared our findings to studies that looked at asthma, bronchitis, and pneumonia. A study examining respiratory admissions for the 2003 California wildfire season found that an average increase of 70 μg/m^3 $PM_{2.5}$ during wildfire smoke events was associated with a 34% increase in asthma admissions for all-ages when compared to pre-wildfire smoke periods (Delfino et al., 2009). When looking particularly the pediatric age group for 0 – 4, they found an 8.3% (95% CI: 2.2% - 14.9%) increase in asthma associations per 10 μg/m^3 $PM_{2.5.}$ This is similar to the 9.0% (95% CI: 1.0% - 17.0%) increase in odds that we observed for ages 0 – 19 admitted to the ED for a respiratory-related concern on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. This finding was echoed in other papers that looked at respiratory admissions across all ages (DeFlorio-Barker Stephanie et al., 2019; Reid & Maestas, 2019). Specifically looking at our study area, our results aligned with a study examining the health impacts of the 2012 wildfire smoke season in Washington state which also reported a 8.0% (95% CI: 2.0% - 14.0%) increase in risk for all-age asthma-related hospital admissions for a 10 μg/m^3 increase in wildfire smoke (Gan et al., 2017).

When looking specifically at studies that only considered pediatric populations, our results also mirrored the general trends in findings. A study that focused on the 2017 Lilac Fire (a wind-driven wildfire) in San Diego County and its association with pediatric patient visits to ED and urgent care clinics found that the fire occurrence was associated with an average of 16 additional respiratory visits per day (95% CI: 11.2 – 20.9) across all pediatric age groups, with ages 6 – 12 reporting the highest relative increase in cases of 3.4 (95% CI: 2.3 – 4.6) excess visits per day (Leibel et al., 2019). However, children aged 0 – 5 years reported 7.3 (96% CI: 3.0 – 11.7) excess respiratory visits per day, which was the highest in terms of absolute visits (not comparing relative increase) (Leibel et al., 2019). In comparison, when considering same-day ED visits, the only significant age-related result we observed was a 12.0% (95% CI: 1.4% - 24.0%) increase in odds for ages 0 – 5 that visited the ED for respiratory infections on wildfire smoke days versus non-wildfire smoke days when controlling for Humidex (Figure 2). This finding of ours is similar to multiple other studies in the California region. A study looking at pediatric ED and urgent care admissions for San Diego County during 2011 – 2017 also observed that children ages 0 – 5 were the most affected by $PM_{2.5}$ specifically from wildfire smoke where a 10 $\mu g/m^3$ increase in wildfire smoke-related $PM_{2.5}$ resulted in a 34.5% (95% CI: 14.8% - 54.3%) increase in pediatric respiratory visits, and this was the largest effect estimate reported for the pediatric age groups (Aguilera, Corringham, Gershunov, Leibel, et al., 2021). For another paper examining the health impacts associated with the 2007 San Diego wildfires on Medi-Cal pediatric patients, children ages 0 – 4 reported a 136% increase in ED visits associated with asthma, and within this subgroup, the youngest children (aged 0 – 1) reported a 243% increase (Hutchinson et al., 2018).

To our knowledge, this is the first study that evaluates inpatient/observational pediatric hospitalization outcomes in Washington state associated with wildfire smoke. Other recent case crossover pediatric studies in Washington (Bernstein et al., 2022.) or other wildfire smoke-related pediatric studies (Aguilera, Corringham, Gershunov, Leibel, et al., 2021; Henry et al., 2021; Hutchinson et al., 2018; Leibel et al., 2019) focus on ED and urgent care clinic visits. In our inpatient/observational dataset, the patients are not necessarily admitted from the ED, and it was within this dataset that we observed statistically significant increases in odds of hospitalization for all lagged exposure days from 1 – 4. The highest increase in odds we observed was an 8.0% (95% CI: 4.0% - 12.0%) elevation in odds for patients that were exposed one day prior to wildfire smoke days versus non-wildfire smoke days when controlling for Humidex. The lowest increase in lag odds was a 5.0% (95% CI: 2.0% - 9.0%) increase for two-day lag and a 5.0% (95% CI: 1.0% - 9.0%) increase for four-day lag. This finding aligns with a study on ED admissions in Sydney, Australia during 1996 – 2007 where the positive associations between ED visits for all non-trauma conditions and respiratory outcomes (including asthma and COPD) in all ages continued for one to three days after the wildfire smoke event (Johnston et al., 2014). In a study examining United States (U.S.) counties within 200 kilometers of 123 large wildfires from 2008 – 2010, there was a generally positive association between lags of 0 – 6 days and modeled wildfire-associated $PM_{2.5}$ values for people 65 and over (DeFlorio-Barker et al., 2019). The highest lag effect observed by DeFlorio-Barker et al. (2019) was one day lagged exposure where they reported a 5.78% (95% CI: 2.85% – 8.71%) increase in asthma, bronchitis, and wheezing hospitalizations per a 10 $\mu g/m^3$ increase in wildfire smoke-related $PM_{2.5}$. For Washington-specific wildfire smoke lag trends, a study examining the health impacts in all ages of the 2012 wildfire smoke season in Washington state reported that an increase in blended model $PM_{2.5}$ estimates was associated with an increase in “all respiratory” admissions for all six lag days (ranging from a 3.0% [95% CI: 1.4% - 6.0%] increase in odds for lag day 5 to a 5.0% [95% CI: 3.0% - 9.0%] increase in odds for lag day 3) (Gan et al., 2017). Finally, a study examining mortality associated with wildfire smoke exposure in Washington state from 2006 - 2017 reported a 1.3% (95% CI: 0.0% - 2.0%) increase in the odds of next day non-traumatic mortality in all ages for previous day wildfire smoke exposure when compared to non-wildfire smoke exposure after controlling for exposure four days prior to death and Humidex (Doubleday et al., 2020).

Limitations and Future Study Recommendations

When assessing wildfire smoke exposure, this analysis was limited by the assumption that individual exposure for the entire zip code is represented by the recorded $PM_{2.5}$ level at the nearest monitor. This assumption relies on monitor performance as well as the idea that all exposure occurred at the residential address. If a patient spent most of their time outside of the residential address, that could result in potential exposure misclassification. As mentioned in (Yao et al., 2016), this concern is particularly applicable for rural communities in which there are fewer monitors per geographic area. Additionally, the classification of wildfire smoke day (using the 20.4 $\mu g/m^3$ threshold) can be challenging due to the lack of available methodology to distinguish between anthropogenic and wildfire related $PM_{2.5}$. However, the tailored methodology reduced misclassification from anthropogenic $PM_{2.5}$ sources such as ambient PM pollution or event-based from fireworks. Finally, when merging the exposure dataset with the health outcomes dataset, there were 35 cases that we had to drop since there was no available exposure data for those data/zip code combinations.

Regarding health outcomes, this paper was limited in terms of granularity due to the use of APR-DRGs. We utilized this broad grouping strategy because it allowed us to easily compare cases before and after the switch from ICD-9 to ICD-10, and it increased our study power by allowing more cases in our groupings that may be influenced by $PM_{2.5}$-related effects. If we wanted to look at the evaluate the impact on certain individual conditions (such as only bronchitis or only a certain type of trauma), then we could utilize the ICD codes that were included in the PHIS dataset. Additionally, we dropped 32 cases with a COVID-19 diagnosis from the analysis in 2020 since the symptoms from COVID-19 could artificially inflate the respiratory associations, and there was not enough power to conduct an interaction analysis. Considering the dataset used, our findings may not be as generalizable to children as a whole, because we were not able to account for underlying conditions and because our data was from a single site study at a tertiary care hospital. As a result, Seattle Children's tends to serve children that may have more complex medica conditions compared to other pediatric patients, so a severity analysis would need to be completed to make any statements about generalizability.

There is a plethora of directions future research can take to address these limitations and further explore the associations relayed in this study. First, exposure could be modeled as a continuous variable (such the cited papers which model $PM_{2.5}$ exposure in increases of 10 $\mu g/m^3$) as this would address some of the concerns associated with a binary smoke day definition. Second, as mentioned in the limitations, underlying conditions within the patient population can affect associations between pollutants admitted from wildfires and adverse health outcomes (Bernstein et al., 2022; Méndez, 2022). Conducting an analysis that examines the interaction with certain pre-existing conditions (such as diabetes) and wildfire smoke exposure would shed additional insight on vulnerability in sub-populations. Third, we would strongly recommend conducting lag analyses farther out than four days. As observed in our own results, there was still a statistically significant increase in odds of inpatient/observational pediatric admission on wildfire smoke days versus non-wildfire smoke days after adjusting for Humidex on the fourth lag day, and conversations with the doctors from Seattle Children's Hospital have yielded a recommendation of exploring lag up to 10 – 14 days, especially since families may wait for longer than 1 – 4 days to bring their children to the hospital. We hypothesize this is because wildfire-related symptoms may start out less severe as their severity develops over multiple days, parents may wait longer than 4 days before inpatient/observational admission.

CONCLUSIONS

The findings from this study mostly aligned with similar studies conducted in California, and to our knowledge, it is the first study to examine the associations between pediatric inpatient/observational hospital admissions and wildfire smoke exposure in Washington state. Our findings add to the current understanding by observing increased odds of respiratory-related ED visits and all-cause hospital admissions for Seattle Children's patients on smoke days versus non-smoke days when adjusting to Humidex. The risk remained elevated with exposure to wildfire smoke up to 4 days prior to inpatient and observational hospital admission. The results can be used to inform pediatric practitioners about the impacts that patients and their families can expect during smoke seasons. Further research should be conducted to determine how underlying conditions, such as diabetes or immune disorders, affect vulnerability to and effect impact from wildfire smoke.

AIM 3. COMMUNICATION OF FINDINGS

For Aim 3, our goal was to disseminate our findings and methods to various audiences that would find the information relevant and to convey our results and recommendations. To accomplish this aim via oral presentations, we chose three main audiences: Seattle Children's practitioners, my peers and the general public, and environmental health-specific conferences (academics).

We presented the background, methods, and initial findings to a group of doctors from Seattle Children's in April of 2022 to receive feedback on the underlying mechanisms for the trends that we viewed. Next, we presented the results to a group of peers and general audience members at the Program on Climate Change Research Symposium on May 19th. The presentation was a shortened version of the full talk, and this version focused on the relevant results and potential applications of the research. This framing allowed us to connect more with audience members to relay the significance of my findings on their daily lives. Finally, we will be presenting the findings in an elongated 55-minute talk at the National Environmental Health Association (NEHA) Annual Education Conference (AEC) conference in Spokane, Washington from June 28, 2022 – July 1, 2022. The intended audience here will be comprised of mostly academics and practitioners, so the information will be methods and detail-heavy to demonstrate credibility and thoroughness in the findings.

REFERENCES

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, *113*(42), 11770–11775. https://doi.org/10.1073/pnas.1607171113

About the air quality monitoring network—Washington State Department of Ecology. (n.d.). Retrieved April 26, 2022, from https://ecology.wa.gov/Regulations-Permits/Guidance-technical-assistance/Air-quality-monitoring-network

Adetona, O., Reinhardt, T. E., Domitrovich, J., Broyles, G., Adetona, A. M., Kleinman, M. T., Ottmar, R. D., & Naeher, L. P. (2016). Review of the health effects of wildland fire smoke on wildland firefighters and the public. *Inhalation Toxicology*, *28*(3), 95–139. https://doi.org/10.3109/08958378.2016.1145771

Adetona, O., Zhang, J. (Junfeng), Hall, D. B., Wang, J.-S., Vena, J. E., & Naeher, L. P. (2013). Occupational exposure to woodsmoke and oxidative stress in wildland firefighters. *Science of The Total Environment*, *449*, 269–275. https://doi.org/10.1016/j.scitotenv.2013.01.075

Aguilera, R., Corringham, T., Gershunov, A., & Benmarhnia, T. (2021). Wildfire smoke impacts respiratory health more than fine particles from other sources: Observational evidence from Southern California. *Nature Communications*, *12*(1), 1493. https://doi.org/10.1038/s41467-021-21708-0

Aguilera, R., Corringham, T., Gershunov, A., Leibel, S., & Benmarhnia, T. (2021). Fine Particles in Wildfire Smoke and Pediatric Respiratory Health in California. *Pediatrics*, *147*(4), e2020027128. https://doi.org/10.1542/peds.2020-027128

Alman, B. L., Pfister, G., Hao, H., Stowell, J., Hu, X., Liu, Y., & Strickland, M. J. (2016). The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: A case crossover study. *Environmental Health*, *15*(1), 64. https://doi.org/10.1186/s12940-016-0146-8

Analitis, A., Georgiadis, I., & Katsouyanni, K. (2012). Forest fires are associated with elevated mortality in a dense urban setting. *Occupational and Environmental Medicine*, *69*(3), 158–162. https://doi.org/10.1136/oem.2010.064238

Balmes, J. R. (2018). Where There's Wildfire, There's Smoke. *New England Journal of Medicine*, *378*(10), 881–883. https://doi.org/10.1056/NEJMp1716846

Belleville, G., Ouellet, M.-C., & Morin, C. M. (2019). Post-Traumatic Stress among Evacuees from the 2016 Fort McMurray Wildfires: Exploration of Psychological and Sleep Symptoms Three Months after the Evacuation. *International Journal of Environmental Research and Public Health*, *16*(9), 1604. https://doi.org/10.3390/ijerph16091604

Bennett, W. D., Zeman, K. L., & Jarabek, A. M. (2007). Nasal Contribution to Breathing and Fine Particle Deposition in Children Versus Adults. *Journal of Toxicology and Environmental Health, Part A*, *71*(3), 227–237. https://doi.org/10.1080/15287390701598200

Bernstein, A. S., Sun, S., Weinberger, K. R., Spangler, K. R., Sheffield, P. E., & Wellenius, G. A. (n.d.). Warm Season and Emergency Department Visits to U.S. Children's Hospitals. *Environmental Health Perspectives*, *130*(1), 017001. https://doi.org/10.1289/EHP8083

Black, C., Tesfaigzi, Y., Bassein, J. A., & Miller, L. A. (2017). Wildfire smoke exposure and human health: Significant gaps in research for a growing public health issue.

Environmental Toxicology and Pharmacology, *55*, 186–195. https://doi.org/10.1016/j.etap.2017.08.022

Bonamonte, D., Filoni, A., Vestita, M., Romita, P., Foti, C., & Angelini, G. (2019). The Role of the Environmental Risk Factors in the Pathogenesis and Clinical Outcome of Atopic Dermatitis. *BioMed Research International*, *2019*, 1–11. https://doi.org/10.1155/2019/2450605

Bowman, D. M. J. S., Kolden, C. A., Abatzoglou, J. T., Johnston, F. H., van der Werf, G. R., & Flannigan, M. (2020). Vegetation fires in the Anthropocene. *Nature Reviews Earth & Environment*, *1*(10), 500–515. https://doi.org/10.1038/s43017-020-0085-3

Brown, M. R. G., Agyapong, V., Greenshaw, A. J., Cribben, I., Brett-MacLean, P., Drolet, J., McDonald-Harker, C., Omeje, J., Mankowsi, M., Noble, S., Kitching, D., & Silverstone, P. H. (2019). After the Fort McMurray wildfire there are significant increases in mental health symptoms in grade 7–12 students compared to controls. *BMC Psychiatry*, *19*, 18. https://doi.org/10.1186/s12888-018-2007-1

Calderón-Garcidueñas, L., Solt, A. C., Henríquez-Roldán, C., Torres-Jardón, R., Nuse, B., Herritt, L., Villarreal-Calderón, R., Osnaya, N., Stone, I., García, R., Brooks, D. M., González-Maciel, A., Reynoso-Robles, R., Delgado-Chávez, R., & Reed, W. (2008). Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid beta-42 and alpha-synuclein in children and young adults. *Toxicologic Pathology*, *36*(2), 289–310. https://doi.org/10.1177/0192623307313011

Cascio, W. E. (2018). Wildland Fire Smoke and Human Health. *The Science of the Total Environment*, *624*, 586–595. https://doi.org/10.1016/j.scitotenv.2017.12.086

Chen, E., Schreier, H. M. C., Strunk, R. C., & Brauer, M. (2008). Chronic Traffic-Related Air Pollution and Stress Interact to Predict Biologic and Clinical Outcomes in Asthma. *Environmental Health Perspectives*, *116*(7), 970–975. https://doi.org/10.1289/ehp.11076

Chen, H., Samet, J. M., Bromberg, P. A., & Tong, H. (2021). Cardiovascular health impacts of wildfire smoke exposure. *Particle and Fibre Toxicology*, *18*. https://doi.org/10.1186/s12989-020-00394-8

DeFlorio-Barker Stephanie, Crooks James, Reyes Jeanette, & Rappold Ana G. (2019). Cardiopulmonary Effects of Fine Particulate Matter Exposure among Older Adults, during Wildfire and Non-Wildfire Periods, in the United States 2008–2010. *Environmental Health Perspectives*, *127*(3), 037006. https://doi.org/10.1289/EHP3860

Delfino, R. J., Brummel, S., Wu, J., Stern, H., Ostro, B., Lipsett, M., Winer, A., Street, D. H., Zhang, L., Tjoa, T., & Gillen, D. L. (2009). The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. *Occupational and Environmental Medicine*, *66*(3), 189–197. https://doi.org/10.1136/oem.2008.041376

Dohrenwend, P. B., Le, M. V., Bush, J. A., & Thomas, C. F. (2013). The Impact on Emergency Department Visits for Respiratory Illness During the Southern California Wildfires. *Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health*, *14*(2). https://doi.org/10.5811/westjem.2012.10.6917

Domitrovich, J., & Sharkey, B. (n.d.). *Heat Illness Basics for Wildland Firefighters*. 8.

Doubleday, A., Schulte, J., Sheppard, L., Kadlec, M., Dhammapala, R., Fox, J., & Busch Isaksen, T. (2020). Mortality associated with wildfire smoke exposure in Washington state, 2006–2017: A case-crossover study. *Environmental Health*, *19*(1), 4. https://doi.org/10.1186/s12940-020-0559-2

Dréno, B., Bettoli, V., Araviiskaia, E., Sanchez Viera, M., & Bouloc, A. (2018). The influence of exposome on acne. *Journal of the European Academy of Dermatology and Venereology*, *32*(5), 812–819. https://doi.org/10.1111/jdv.14820

Eudier, F., Hucher, N., Picard, C., Savary, G., & Grisel, M. (2019). Squalene Oxidation Induced by Urban Pollutants: Impact on Skin Surface Physico-Chemistry. *Chemical Research in Toxicology*, *32*(2), 285–293. https://doi.org/10.1021/acs.chemrestox.8b00311

Evans, L. F., Weeks, I. A., Eccleston, A. J., & Packham, D. R. (1977). Photochemical ozone in smoke from prescribed burning of forests. *Environmental Science & Technology*, *11*(9), 896–900. https://doi.org/10.1021/es60132a006

Finlay, S. E., Moffat, A., Gazzard, R., Baker, D., & Murray, V. (2012). Health Impacts of Wildfires. *PLoS Currents*, *4*, e4f959951cce2c. https://doi.org/10.1371/4f959951cce2c

Flannigan, M., Cantin, A. S., de Groot, W. J., Wotton, M., Newbery, A., & Gowman, L. M. (2013). Global wildland fire season severity in the 21st century. *Forest Ecology and Management*, *294*, 54–61. https://doi.org/10.1016/j.foreco.2012.10.022

Gan, R. W., Ford, B., Lassman, W., Pfister, G., Vaidyanathan, A., Fischer, E., Volckens, J., Pierce, J. R., & Magzamen, S. (2017). Comparison of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. *GeoHealth*, *1*(3), 122–136. https://doi.org/10.1002/2017GH000073

Gauderman, W. J., Urman, R., Avol, E., Berhane, K., McConnell, R., Rappaport, E., Chang, R., Lurmann, F., & Gilliland, F. (2015). Association of Improved Air Quality with Lung

Development in Children. *New England Journal of Medicine*, *372*(10), 905–913. https://doi.org/10.1056/NEJMoa1414123

Gibbs, L., Nursey, J., Cook, J., Ireton, G., Alkemade, N., Roberts, M., Gallagher, H. C., Bryant, R., Block, K., Molyneaux, R., & Forbes, D. (2019). Delayed Disaster Impacts on Academic Performance of Primary School Children. *Child Development*, *90*(4), 1402–1412. https://doi.org/10.1111/cdev.13200

Henry, S., Ospina, M. B., Dennett, L., & Hicks, A. (2021). Assessing the Risk of Respiratory-Related Healthcare Visits Associated with Wildfire Smoke Exposure in Children 0–18 Years Old: A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(16), 8799. https://doi.org/10.3390/ijerph18168799

Holm, S. M., Miller, M. D., & Balmes, J. R. (2021). Health effects of wildfire smoke in children and public health tools: A narrative review. *Journal of Exposure Science & Environmental Epidemiology*, *31*(1), 1–20. https://doi.org/10.1038/s41370-020-00267-4

Holstius, D. M., Reid, C. E., Jesdale, B. M., & Morello-Frosch, R. (2012). Birth Weight following Pregnancy during the 2003 Southern California Wildfires. *Environmental Health Perspectives*, *120*(9), 1340–1345. https://doi.org/10.1289/ehp.1104515

Hutchinson, J. A., Vargo, J., Milet, M., French, N. H. F., Billmire, M., Johnson, J., & Hoshiko, S. (2018). The San Diego 2007 wildfires and Medi-Cal emergency department presentations, inpatient hospitalizations, and outpatient visits: An observational study of smoke exposure periods and a bidirectional case-crossover analysis. *PLOS Medicine*, *15*(7), e1002601. https://doi.org/10.1371/journal.pmed.1002601

Janes, H., Sheppard, L., & Lumley, T. (2005). Case–Crossover Analyses of Air Pollution Exposure Data: Referent Selection Strategies and Their Implications for Bias. *Epidemiology*, *16*(6), 717–726. https://doi.org/10.1097/01.ede.0000181315.18836.9d

Johnston, F. H., Purdie, S., Jalaludin, B., Martin, K. L., Henderson, S. B., & Morgan, G. G. (2014). Air pollution events from forest fires and emergency department attendances in Sydney, Australia 1996–2007: A case-crossover analysis. *Environmental Health*, *13*(1), 105. https://doi.org/10.1186/1476-069X-13-105

Jones Caitlin G., Rappold Ana G., Vargo Jason, Cascio Wayne E., Kharrazi Martin, McNally Bryan, Hoshiko Sumi, & null null. (2020). Out-of-Hospital Cardiac Arrests and Wildfire-Related Particulate Matter During 2015–2017 California Wildfires. *Journal of the American Heart Association*, *9*(8), e014125. https://doi.org/10.1161/JAHA.119.014125

Karnauskas, K. B., Lundquist, J. K., & Zhang, L. (2018). Southward shift of the global wind energy resource under high carbon dioxide emissions. *Nature Geoscience*, *11*(1), 38–43. https://doi.org/10.1038/s41561-017-0029-9

Kim, J., Kim, B. E., & Leung, D. Y. M. (2019). Pathophysiology of atopic dermatitis: Clinical implications. *Allergy and Asthma Proceedings*, *40*(2), 84–92. https://doi.org/10.2500/aap.2019.40.4202

Kim, Y., Knowles, S., Manley, J., & Radoias, V. (2017). Long-run health consequences of air pollution: Evidence from Indonesia's forest fires of 1997. *Economics & Human Biology*, *26*, 186–198. https://doi.org/10.1016/j.ehb.2017.03.006

Krutmann, J., Liu, W., Li, L., Pan, X., Crawford, M., Sore, G., & Seite, S. (2014). Pollution and skin: From epidemiological and mechanistic studies to clinical implications. *Journal of Dermatological Science*, *76*(3), 163–168. https://doi.org/10.1016/j.jdermsci.2014.08.008

Lefebvre, M.-A., Pham, D.-M., Boussouira, B., Bernard, D., Camus, C., & Nguyen, Q.-L. (2015). Evaluation of the impact of urban pollution on the quality of skin: A multicentre study in Mexico. *International Journal of Cosmetic Science*, *37*(3), 329–338. https://doi.org/10.1111/ics.12203

Leibel, S., Nguyen, M., Brick, W., Parker, J., Ilango, S., Aguilera, R., Gershunov, A., & Benmarhnia, T. (2019). Increase in Pediatric Respiratory Visits Associated with Santa Ana Wind–Driven Wildfire Smoke and PM2.5 Levels in San Diego County. *Annals of the American Thoracic Society*, *17*(3), 313–320. https://doi.org/10.1513/AnnalsATS.201902-150OC

Levy, D., Sheppard, L., Checkoway, H., Kaufman, J., Lumley, T., Koenig, J., & Siscovick, D. (2001). A Case-Crossover Analysis of Particulate Matter Air Pollution and Out-of-Hospital Primary Cardiac Arrest. *Epidemiology*, *12*(2), 193–199.

Li, X., An, S.-J., Liu, X.-L., Ji, A.-L., Cao, Y., Xiang, Y., Ma, X.-Y., Hu, Q., Yuan, Z.-Q., Li, Y.-F., Lu, Y.-G., & Cai, T.-J. (2021). The Association Between Short-Term Air Pollution Exposure and Post-Adolescent Acne: The Evidence from a Time Series Analysis in Xi’an, China. *Clinical, Cosmetic and Investigational Dermatology*, *14*, 723–731. https://doi.org/10.2147/CCID.S320248

Liu, J. C., Mickley, L. J., Sulprizio, M. P., Dominici, F., Yue, X., Ebisu, K., Anderson, G. B., Khan, R. F. A., Bravo, M. A., & Bell, M. L. (2016). Particulate air pollution from wildfires in the Western US under climate change. *Climatic Change*, *138*(3), 655–666. https://doi.org/10.1007/s10584-016-1762-6

Liu, J. C., Mickley, L. J., Sulprizio, M. P., Yue, X., Peng, R. D., Dominici, F., & Bell, M. L. (2016). Future respiratory hospital admissions from wildfire smoke under climate change in the Western US. *Environmental Research Letters*, *11*(12), 124018. https://doi.org/10.1088/1748-9326/11/12/124018

Liu, J. C., & Peng, R. D. (2019). The impact of wildfire smoke on compositions of fine particulate matter by ecoregion in the Western US. *Journal of Exposure Science & Environmental Epidemiology*, *29*(6), 765–776. https://doi.org/10.1038/s41370-018-0064-7

Liu, J. C., Pereira, G., Uhl, S. A., Bravo, M. A., & Bell, M. L. (2015). A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. *Environmental Research*, *136*, 120–132. https://doi.org/10.1016/j.envres.2014.10.015

Liu, J. C., Wilson, A., Mickley, L. J., Dominici, F., Ebisu, K., Wang, Y., Sulprizio, M. P., Peng, R. D., Yue, X., Son, J.-Y., Anderson, G. B., & Bell, M. L. (2017). Wildfire-specific Fine Particulate Matter and Risk of Hospital Admissions in Urban and Rural Counties. *Epidemiology*, *28*(1), 77–85. https://doi.org/10.1097/EDE.0000000000000556

McFarlane, A. C., & Van Hooff, M. (2009). Impact of childhood exposure to a natural disaster on adult mental health: 20-year longitudinal follow-up study. *The British Journal of Psychiatry: The Journal of Mental Science*, *195*(2), 142–148. https://doi.org/10.1192/bjp.bp.108.054270

Méndez, M. (2022, January 12). Behind the Bougainvillea Curtain: Wildfires and Inequality. *Issues in Science and Technology*. https://issues.org/wildfires-inequality-indigenous-undocumented-workers-mendez/

Miller, M. D., & Marty, M. A. (2010). Impact of Environmental Chemicals on Lung Development. *Environmental Health Perspectives*, *118*(8), 1155–1164. https://doi.org/10.1289/ehp.0901856

Naeher, L. P., Brauer, M., Lipsett, M., Zelikoff, J. T., Simpson, C. D., Koenig, J. Q., & Smith, K. R. (2007). Woodsmoke Health Effects: A Review. *Inhalation Toxicology*, *19*(1), 67–106. https://doi.org/10.1080/08958370600985875

Okada, H. C., Alleyne, B., Varghai, K., Kinder, K., & Guyuron, B. (2013). Facial changes caused by smoking: A comparison between smoking and nonsmoking identical twins. *Plastic and Reconstructive Surgery*, *132*(5), 1085–1092. https://doi.org/10.1097/PRS.0b013e3182a4c20a

Oliveira, M., Slezakova, K., Delerue-Matos, C., Pereira, M. C., & Morais, S. (2019). Children environmental exposure to particulate matter and polycyclic aromatic hydrocarbons and biomonitoring in school environments: A review on indoor and outdoor exposure levels, major sources and health impacts. *Environment International*, *124*, 180–204. https://doi.org/10.1016/j.envint.2018.12.052

Oudin, A., Bråbäck, L., Åström, D. O., Strömgren, M., & Forsberg, B. (2016). Association between neighbourhood air pollution concentrations and dispensed medication for psychiatric disorders in a large longitudinal cohort of Swedish children and adolescents. *BMJ Open*, *6*(6), e010004. https://doi.org/10.1136/bmjopen-2015-010004

Power, M. C., Kioumourtzoglou, M.-A., Hart, J. E., Okereke, O. I., Laden, F., & Weisskopf, M. G. (2015). The relation between past exposure to fine particulate air pollution and prevalent anxiety: Observational cohort study. *BMJ*, *350*, h1111. https://doi.org/10.1136/bmj.h1111

Reid, C. E., Brauer, M., Johnston, F. H., Jerrett, M., Balmes, J. R., & Elliott, C. T. (2016). Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives*, *124*(9), 1334–1343. https://doi.org/10.1289/ehp.1409277

Reid, C. E., & Maestas, M. M. (2019). Wildfire smoke exposure under climate change: Impact on respiratory health of affected communities. *Current Opinion in Pulmonary Medicine*, *25*(2), 179–187. https://doi.org/10.1097/MCP.0000000000000552

Roberts, S., Arseneault, L., & Barratt, B. (2018). Exploration of NO2 and PM2.5 air pollution and mental health problems using high-resolution data in London-based children from a UK longitudinal cohort study. *Psychiatry Res.*, *272*, 8–17. https://doi.org/10.1016/j.psychres.2018.12.050

Roberts, W. (2021). Air pollution and skin disorders. *International Journal of Women's Dermatology*, *7*(1), 91–97. https://doi.org/10.1016/j.ijwd.2020.11.001

Shankardass, K., McConnell, R., Jerrett, M., Milam, J., Richardson, J., & Berhane, K. (2009). Parental stress increases the effect of traffic-related air pollution on childhood asthma incidence. *Proceedings of the National Academy of Sciences*, *106*(30), 12406–12411. https://doi.org/10.1073/pnas.0812910106

Sun, Q., Miao, C., Hanel, M., Borthwick, A. G. L., Duan, Q., Ji, D., & Li, H. (2019). Global heat stress on health, wildfires, and agricultural crops under different levels of climate warming. *Environment International*, *128*, 125–136. https://doi.org/10.1016/j.envint.2019.04.025

Sutherland, E. R., Make, B. J., Vedal, S., Zhang, L., Dutton, S. J., Murphy, J. R., & Silkoff, P. E. (2005). Wildfire smoke and respiratory symptoms in patients with chronic obstructive pulmonary disease. *Journal of Allergy and Clinical Immunology*, *115*(2), 420–422. https://doi.org/10.1016/j.jaci.2004.11.030

Szyszkowicz, M., Zemek, R., Colman, I., Gardner, W., Kousha, T., & Smith-Doiron, M. (2020). Air Pollution and Emergency Department Visits for Mental Disorders among Youth.

International Journal of Environmental Research and Public Health, *17*(12), 4190. https://doi.org/10.3390/ijerph17124190

Thomson, E. M. (2019). Air Pollution, Stress, and Allostatic Load: Linking Systemic and Central Nervous System Impacts. *Journal of Alzheimer's Disease: JAD*, *69*(3), 597–614. https://doi.org/10.3233/JAD-190015

Turco, M., Rosa-Cánovas, J. J., Bedia, J., Jerez, S., Montávez, J. P., Llasat, M. C., & Provenzale, A. (2018). Exacerbated fires in Mediterranean Europe due to anthropogenic warming projected with non-stationary climate-fire models. *Nature Communications*, *9*(1), 3821. https://doi.org/10.1038/s41467-018-06358-z

Vanos, J. K. (2015). Children's health and vulnerability in outdoor microclimates: A comprehensive review. *Environment International*, *76*, 1–15. https://doi.org/10.1016/j.envint.2014.11.016

Xu, R., Yu, P., Abramson, M. J., Johnston, F. H., Samet, J. M., Bell, M. L., Haines, A., Ebi, K. L., Li, S., & Guo, Y. (2020). Wildfires, global climate change, and human health. *The New England Journal of Medicine*, *383*(22), 2173–2181. https://doi.org/10.1056/NEJMsr2028985

Yao, J., Eyamie, J., & Henderson, S. B. (2016). Evaluation of a spatially resolved forest fire smoke model for population-based epidemiologic exposure assessment. *Journal of Exposure Science & Environmental Epidemiology*, *26*(3), 233–240. https://doi.org/10.1038/jes.2014.67

Zhou, C., Baïz, N., Banerjee, S., Charpin, D. A., Caillaud, D., de Blay, F., Raherison, C., Lavaud, F., & Annesi-Maesano, I. (2013). The relationships between ambient air pollutants and childhood asthma and eczema are modified by emotion and conduct

problems. *Annals of Epidemiology*, *23*(12), 778-783.e3.
https://doi.org/10.1016/j.annepidem.2013.09.004

APPENDIX A

Table A1. List of APR-DRGs that fall under each health outcome sub-category.

Subcategory	APR-DRG Conditions within Category (Code #)
Respiratory	
	Infections of upper respiratory tract (113)
	Respiratory system diagnosis with ventilator support of 96+ hours (130)
	Cystic fibrosis: pulmonary disease (131)
	BPD & other chronic respiratory dis-arising in perinatal period (132)
	Pulmonary edema & respiratory failure (133)
	Pulmonary embolism (134)
	Major respiratory infections & inflammations (137)
	Bronchiolitis & RSV pneumonia (138)
	Pneumonia (not elsewhere specified) (139)
	Chronic obstructive pulmonary disease (140)
	Asthma (141)
	Interstitial & alveolar lung diseases (142)
	Respiratory diagnoses not elsewhere specified except signs, symptoms & minor diagnoses (143)
	Respiratory signs, symptoms & minor diagnoses (144)
Respiratory Infections	
	Infections of upper respiratory tract (113)
	Major respiratory infections & inflammations (137)
	Bronchiolitis & RSV pneumonia (138)

	Pneumonia (not elsewhere specified) (139)
Dermal Conditions	
	Other skin, subcutaneous tissue & related procedures (364) Major skin disorders (381) Other skin, subcutaneous tissue & breast disorders (385)
Trauma	
	Head trauma with coma >1 hour or hemorrhage (55) Brain contusion/laceration & comp skull fracture, coma < 1 hour or no coma (56) Concussion, closed skull fracture, uncomp intracranial injury, coma <1 hour or no coma (57) Major chest & respiratory trauma (135) Fracture of femur (340) Fracture of pelvis or dislocation of hip (341) Fractures & dislocations except femur, pelvis & back (342) Other back & neck disorders, fractures & injuries (347) Other musculoskeletal system & connective tissue diagnoses (351) Contusion, open wound & other trauma to skin & subcutaneous tissue (384)
Mental Health	
	Schizophrenia (750) Major depressive disorders & other/unspecified psychoses (751) Disorders of personality & impulse control (752) Bipolar disorders (753) Depression except major depressive disorder (754) Adjustment disorders & neuroses except depressive diagnosis (755) Acute anxiety & delirium states (756)

	Organic mental health disturbances (757) Childhood behavioral disorders (758) Eating disorders (759) Other mental health disorders (760)

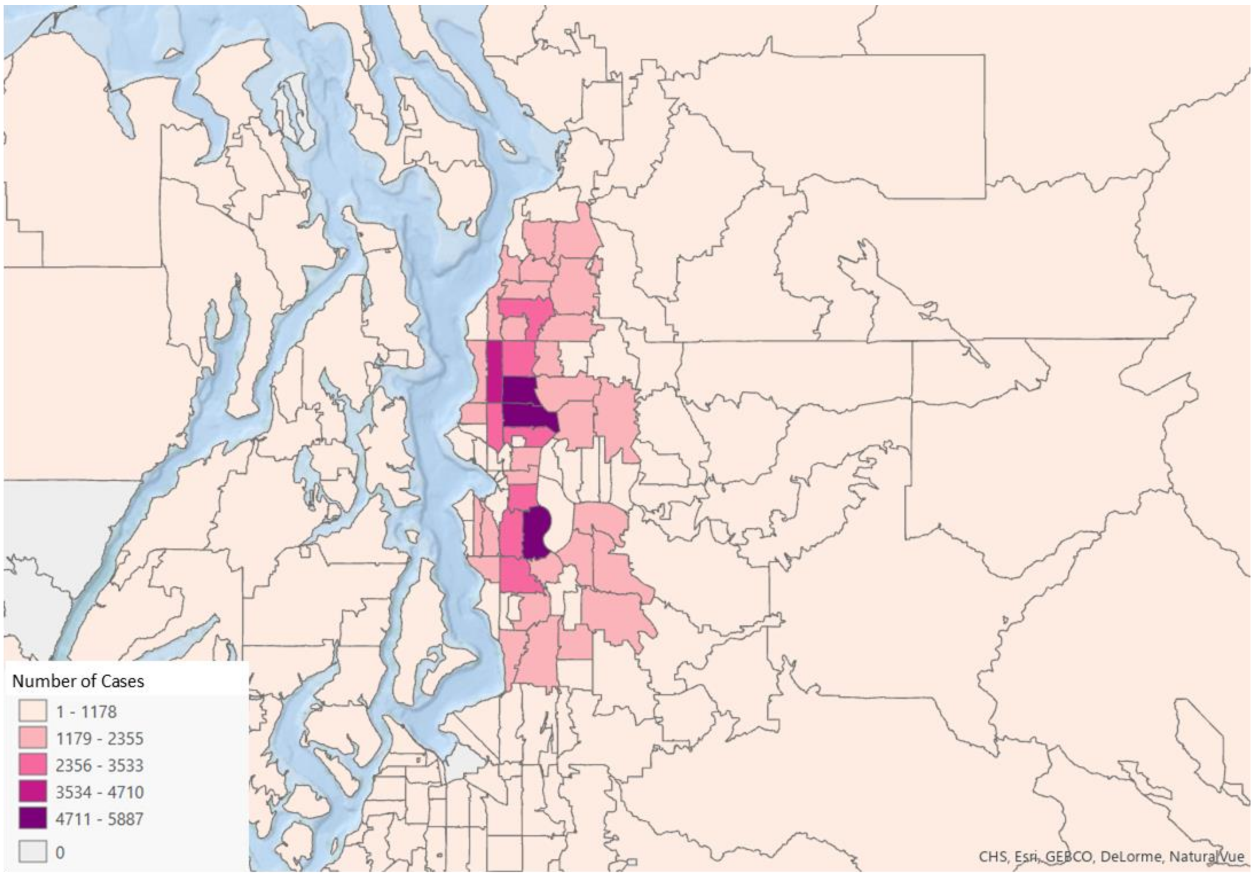

Figure A1. Zoomed in view of Seattle area with the highest density of ED cases.

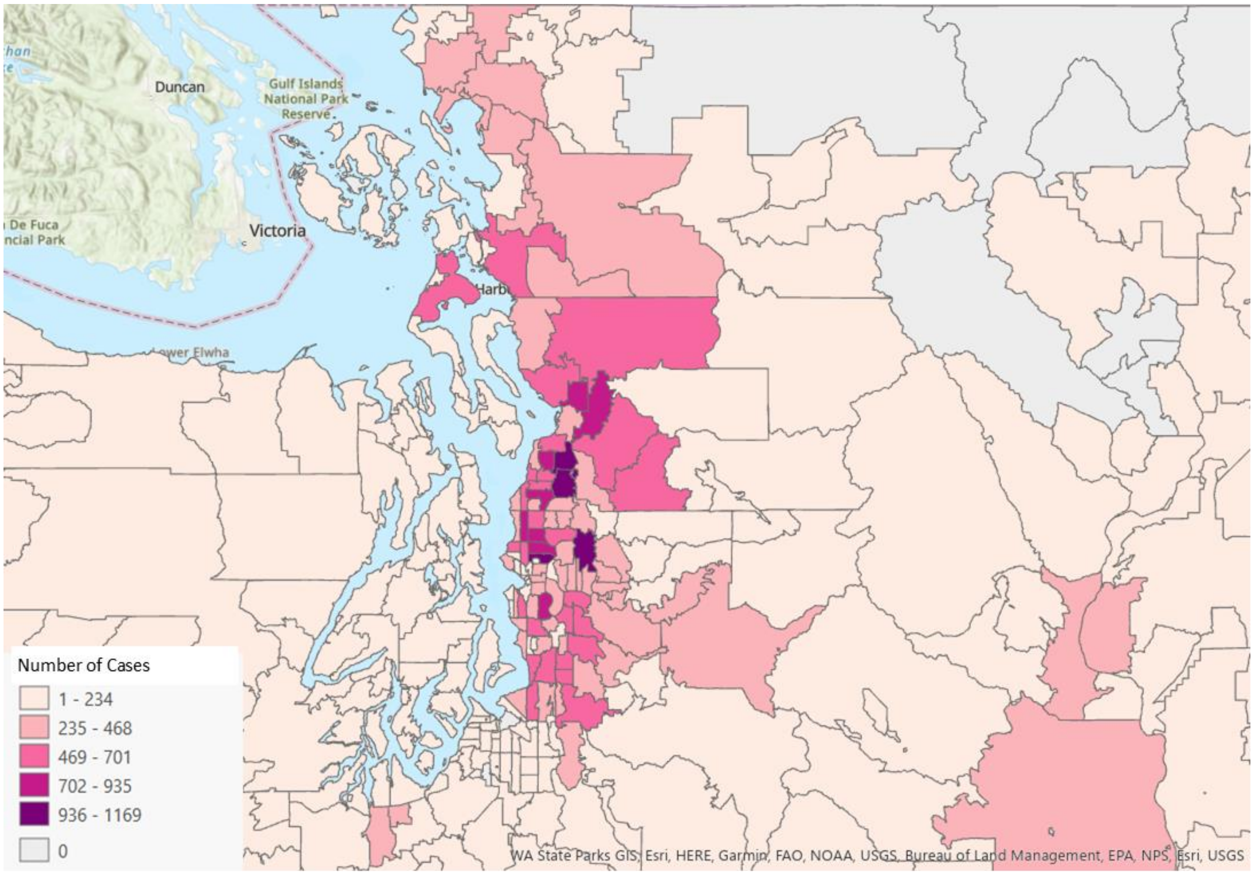

Figure A2. Zoomed in view of King County and Snohomish County areas with the highest density of inpatient/observational cases.

Table A2. ORs for secondary analyses stratified by age, sex, and lag within each health outcome.

	ED			Inpt./Obvs.		
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
Respiratory	**1.09 (1.01, 1.17)**	**0.029**	20,653 (15.60)	0.97 (0.84, 1.11)	0.645	6,304 (10.09)
Age group (years)						
0 – 5	1.09 (1.00, 1.19)	0.055	15,163 (73.42)	0.93 (0.78, 1.11)	0.413	4,178 (66.28)
6 – 12	1.18 (1.00, 1.40)	0.052	4,047 (19.60)	0.86 (0.64, 1.22)	0.449	1,271 (20.16)
13 – 19	0.85 (0.64, 1.14)	0.289	1,443 (6.99)	1.28 (0.91, 1.81)	0.154	855 (13.56)
Sex						
Male	1.06 (0.96, 1.17)	0.253	11,920 (57.72)	0.94 (0.78, 1.13)	0.495	3,739 (59.31)
Female	**1.13 (1.00, 1.27)**	**0.042**	8,732 (42.28)	1.01 (0.81, 1.24)	0.958	2,565 (40.69)
Lag						
0 Day	**1.09 (1.01, 1.17)**	**0.029**	20,653 (15.60)	0.97 (0.84, 1.11)	0.645	6,304 (10.09)
1 Day	1.05 (0.98, 1.13)	0.147	20,265 (98.12)	1.00 (0.88, 1.13)	0.942	6,188 (98.16)
2 Day	1.04 (0.98, 1.11)	0.200	19,867 (96.19)	0.99 (0.87, 1.11)	0.818	6,091 (96.62)
3 Day	1.02 (0.96, 1.09)	0.501	19,540 (94.61)	0.98 (0.88, 1.11)	0.787	5,996 (95.11)
4 Day	1.01 (0.95, 1.08)	0.684	19,189 (92.91)	0.97 (0.87, 1.08)	0.566	5,899 (93.58)
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
Respiratory Infections	**1.11 (1.01, 1.21)**	**0.033**	13,710 (10.35)	1.14 (0.89, 1.45)	0.306	2,147 (3.45)
	OR (95% CI)	**p-val**	**N (%) of ED Cases**	**OR (95% CI)**	**p-val**	**N (%) of Inpt./Obvs Cases**
Age group (years)						
0 – 5	**1.12 (1.00, 1.24)**	**0.042**	10,423 (76.03)	1.02 (0.76, 1.37)	0.890	1,584 (73.44)
6 – 12	1.17 (0.94, 1.47)	0.167	2,416 (17.62)	0.96 (0.50, 1.83)	0.904	341 (15.81)
13 – 19	0.87 (0.59, 1.26)	0.454	871 (6.35)	**2.36 (1.27, 4.39)**	**0.006**	232 (10.76)
Sex						
Male	1.06 (0.93, 1.20)	0.391	7,720 (56.31)	1.24 (0.90, 1.71)	0.183	1,282 (59.43)
Female	**1.17 (1.02, 1.35)**	**0.023**	5,989 (43.68)	1.00 (0.69, 1.44)	0.981	875 (40.57)
Lag						
0 Day	**1.11 (1.01, 1.21)**	**0.033**	13,710 (10.35)	1.14 (0.89, 1.45)	0.306	2,147 (3.45)

	ED			Inpt./Obvs.		
1 Day	1.06 (0.97, 1.15)	0.203	13,435 (97.99)	1.12 (0.89, 1.40)	0.339	2,111 (97.87)
2 Day	1.05 (0.97, 1.14)	0.232	13,149 (95.91)	1.09 (0.88, 1.35)	0.429	2,076 (96.25)
3 Day	1.03 (0.95, 1.11)	0.525	12,918 (94.22)	1.10 (0.90, 1.35)	0.350	2,040 (94.58)
4 Day	1.02 (0.95, 1.10)	0.639	12,659 (92.33)	1.02 (0.84, 1.24)	0.818	2,008 (93.09)
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
Dermal	1.10 (0.95, 1.27)	0.199	5,404 (4.08)	1.61 (0.88, 2.94)	0.126	346 (0.55)
	OR (95% CI)	**p-val**	**N (%) of ED Cases**	**OR (95% CI)**	**p-val**	**N (%) of Inpt./Obvs Cases**
Age group (years)						
0 – 5	1.13 (0.94, 1.35)	0.196	3,269 (60.49)	1.88 (0.87, 4.04)	0.107	219 (63.30)
6 – 12	1.09 (0.82, 1.45)	0.573	1,467 (27.15)	1.27 (0.27, 5.89)	0.760	63 (18.21)
13 – 19	1.00 (0.67, 1.49)	0.998	668 (12.36)	1.37 (0.33, 5.79)	0.665	64 (18.50)
Sex						
Male	1.10 (0.90, 1.34)	0.380	2,787 (51.57)	1.33 (0.57, 3.13)	0.507	172 (49.71)
Female	1.10 (0.90, 1.35)	0.354	2,617 (48.43)	1.87 (0.78, 4.46)	0.159	174 (50.29)
Lag						
0 Day	1.10 (0.95, 1.27)	0.199	5,404 (4.08)	1.61 (0.88, 2.94)	0.126	346 (0.55)
1 Day	1.09 (0.95, 1.24)	0.207	5,302 (98.11)	1.26 (0.71, 2.22)	0.436	345 (99.71)
2 Day	1.04 (0.91, 1.17)	0.598	5,200 (96.23)	1.28 (0.75, 2.17)	0.364	337 (97.40)
3 Day	1.05 (0.93, 1.18)	0.465	5,122 (94.78)	1.18 (0.72, 1.92)	0.522	327 (94.51)
4 Day	1.03 (0.92, 1.16)	0.629	5,030 (93.08)	1.10 (0.67, 1.80)	0.706	325 (93.93)
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
Trauma	0.98 (0.92, 1.05)	0.652	25,451 (19.22)	**1.44 (1.03, 2.02)**	**0.034**	1,020 (1.63)
	OR (95% CI)	**p-val**	**N (%) of ED Cases**	**OR (95% CI)**	**p-val**	**N (%) of Inpt./Obvs Cases**
Age group (years)						
0 – 5	0.98 (0.88, 1.08)	0.637	10,953 (43.04)	**1.67 (1.02, 2.75)**	**0.043**	519 (50.88)
6 – 12	0.99 (0.89, 1.11)	0.915	9,540 (37.48)	0.73 (0.35, 1.51)	0.394	269 (26.37)

	ED			Inpt./Obvs.		
13 – 19	0.99 (0.85, 1.16)	0.910	4,958 (19.48)	**2.13 (1.13, 3.98)**	**0.019**	232 (22.75)
Sex						
Male	0.99 (0.90, 1.08)	0.750	14,541 (57.13)	1.53(1.00, 2.36)	0.053	599 (58.73)
Female	0.98 (0.89, 1.09)	0.728	10,910 (42.87)	1.33 (0.77, 2.28)	0.306	421 (41.28)
Lag						
0 Day	0.98 (0.92, 1.05)	0.652	25,451 (19.22)	**1.44 (1.03, 2.02)**	**0.034**	1,020 (1.63)
1 Day	1.00 (0.94, 1.06)	0.931	25,076 (98.53)	1.35 (0.98, 1.85)	0.064	1,004 (98.43)
2 Day	1.01 (0.95, 1.07)	0.796	24,649 (96.85)	1.22 (0.90, 1.66)	0.206	993 (97.35)
3 Day	1.00 (0.95, 1.06)	0.881	24,239 (95.24)	1.26 (0.94, 1.70)	0.122	975 (95.59)
4 Day	0.99 (0.94, 1.05)	0.759	23,854 (93.73)	1.19 (0.90, 1.59)	0.230	963 (94.41)
Category	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**	**OR (95% CI)**	**p-val**	**N of Total Cases (%)**
Mental Health	0.94 (0.80, 1.10)	0.433	4,203 (3.17)	1.15 (0.98, 1.35)	0.081	4,093 (6.55)
	OR (95% CI)	**p-val**	**N (%) of ED Cases**	**OR (95% CI)**	**p-val**	**N (%) of Inpt./Obvs Cases**
Age group (years)						
0 – 5	0.82 (0.29, 2.31)	0.705	136 (3.24)	0.41 (0.04, 4.14)	0.450	61 (1.49)
6 – 12	1.06 (0.79, 1.41)	0.702	1,385 (32.95)	1.25 (0.94, 1.65)	0.122	1,283 (31.35)
13 – 19	0.89 (0.73, 1.09)	0.267	2,682 (63.81)	1.12 (0.92, 1.36)	0.249	2,749 (67.16)
Sex						
Male	0.76 (0.59, 0.97)	0.029	1,943 (46.23)	1.06 (0.84, 1.34)	0.621	1,892 (46.23)
Female	1.10 (0.89, 1.36)	0.375	2,260 (53.77)	1.25 (1.00, 1.55)	0.047	2,201 (53.78)
Lag						
0 Day	0.94 (0.80, 1.10)	0.433	4,203 (3.17)	1.15 (0.98, 1.35)	0.081	4,093 (6.55)
1 Day	**0.86 (0.74, 1.00)**	**0.044**	4,137 (98.43)	1.12 (0.97, 1.30)	0.126	4,013 (98.05)
2 Day	**0.86 (0.75, 1.00)**	**0.047**	4,052 (96.41)	1.13 (0.98, 1.30)	0.087	3,945 (96.38)
3 Day	**0.86 (0.75, 0.98)**	**0.029**	3,971 (94.48)	1.09 (0.95, 1.25)	0.230	3,902 (95.33)
4 Day	**0.85 (0.74, 0.97)**	**0.017**	3,907 (92.96)	1.11 (0.97, 1.27)	0.128	3,832 (93.62)

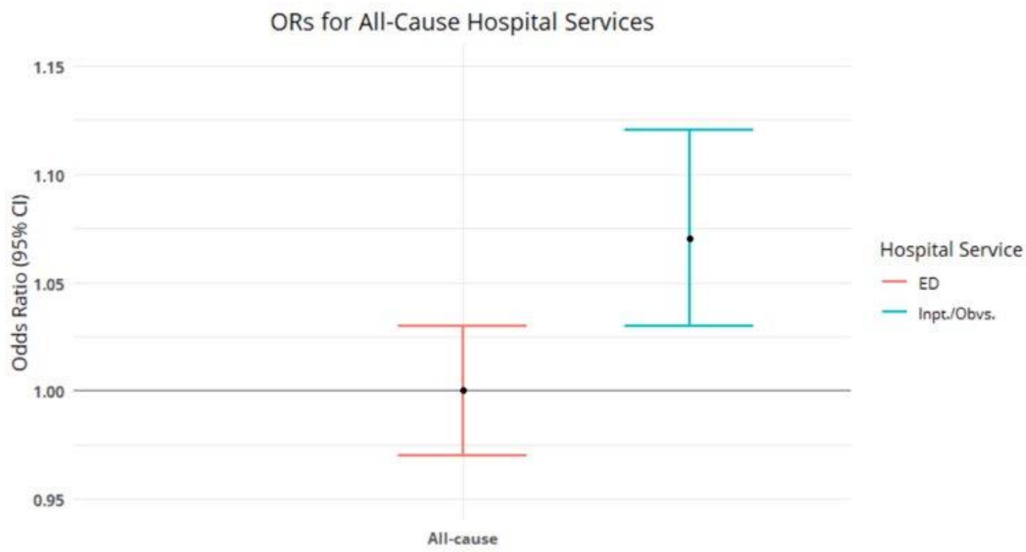

Figure A3. Plot of ORs for all-cause hospital encounters by ED visits and inpatient/observational admissions.

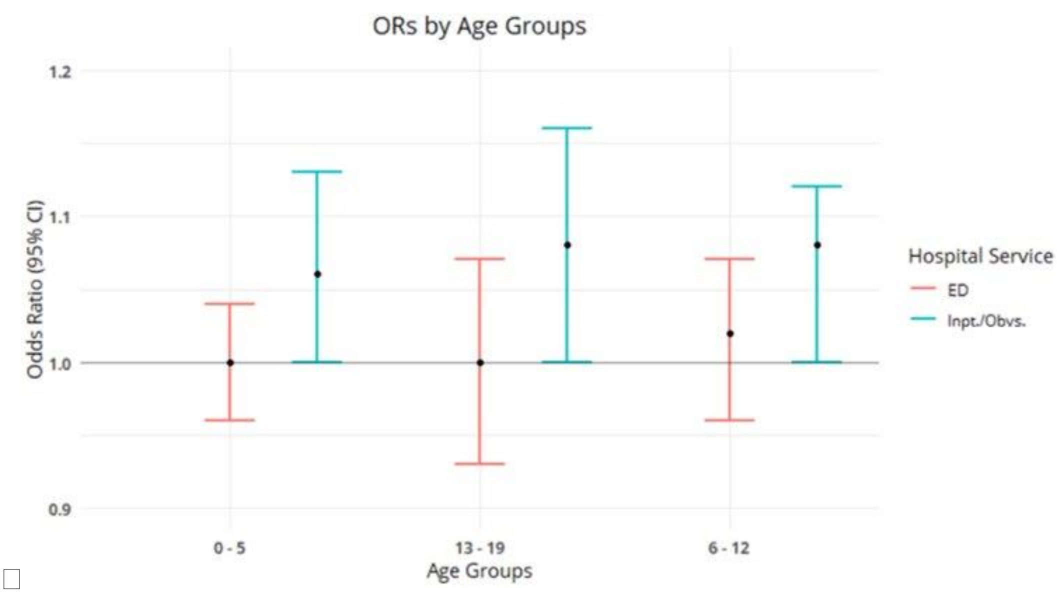

Figure A4. Plot of ORs per each age category by ED visits and inpatient/observational admissions.

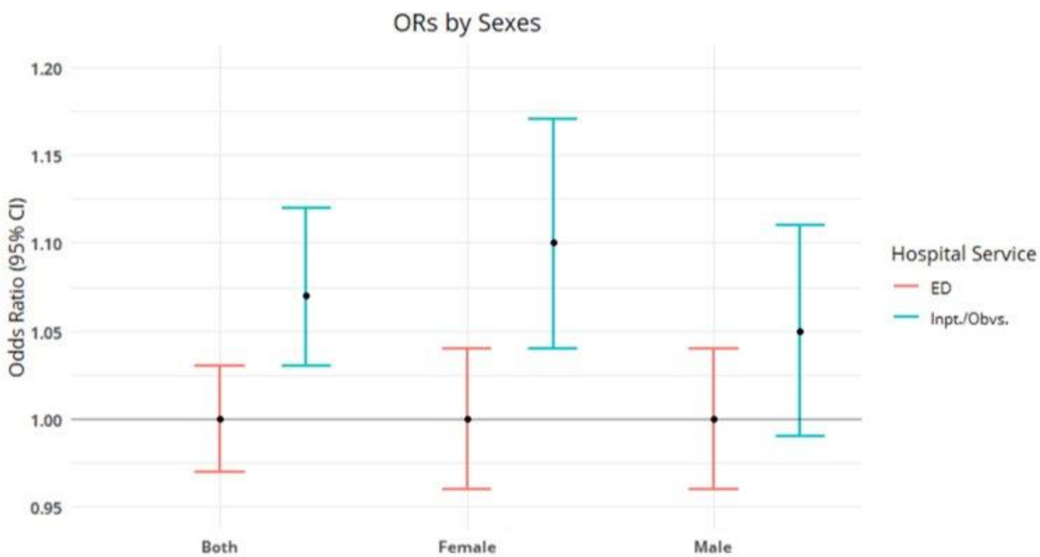

Figure A5. Plot of ORs for sex by ED visits and inpatient/observational admissions.

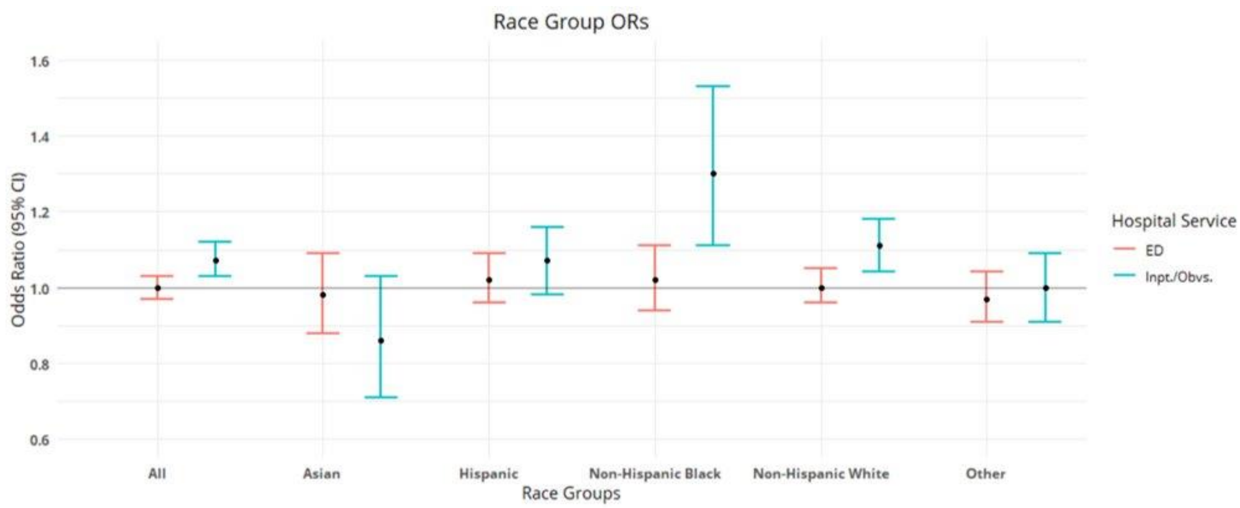

Figure A6. Plot of ORs per race/ethnicity category by ED visits and inpatient/observational admissions.

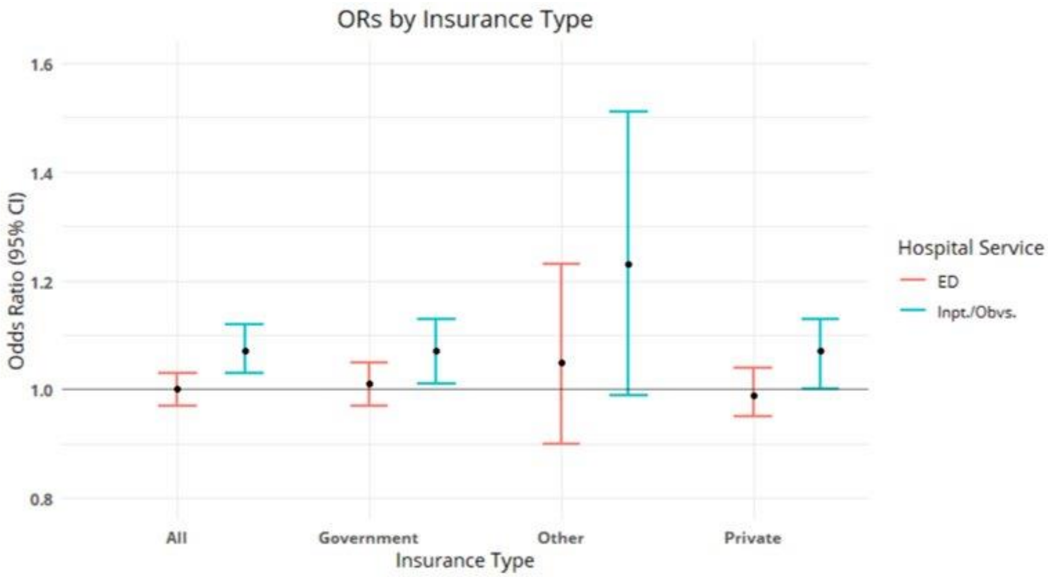

Figure A7. Plot of ORs for insurance types by ED visits and inpatient/observational admissions.

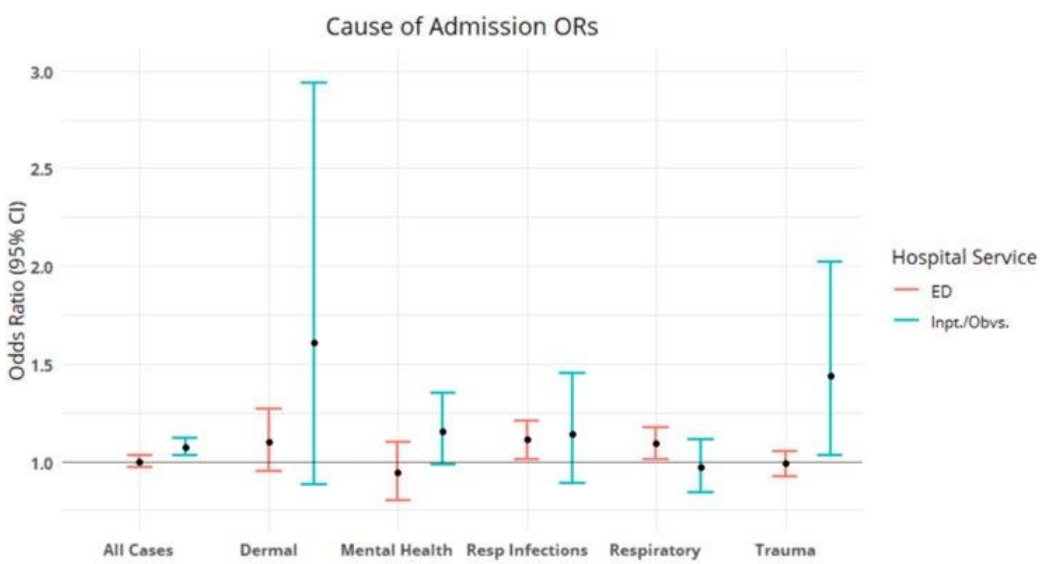

Figure A8. Plot of ORs per cause of admission category by ED visits and inpatient/observational admissions.

APPENDIX B

```
## Code for preparing and cleaning dataset, running models, and creating visualizations
## Loading necessary packages and setting directory
library(lattice)
library(psych)
library(lmtest)
library(dplyr)
library(tidyr)
library(lubridate)
library(readr)
library(readxl)
library(survival) #clogit function

## ---------------------------------------------------
## Set working directory
projectWD <- "C:/Users/daani/Downloads/phis/datasets"
wd <- getwd()
setwd(projectWD)

## ---------------------------------------------------
##Purpose: Merge PHIS and humidex, remove all non-WA residents, remove above age 19
##Load PHIS datasets
phis_load <- read_excel("phis_pm2.5_master_decrypt.xlsx")
humidexwa <- read_csv("03_phis_humidex.csv")
totalhumidex <- read_csv("master_humidex.csv")

summary(phis_load)
summary(humidexwa)
summary(totalhumidex)

View(phis_load)
## ---------------------------------------------------
##Zipcode correction

#delete zip column
phis = subset(phis_load, select = -zip)

summary(phis)

#change name of variable in phis dataset
phis <- phis %>%
  rename(zip = zipd)
```

```
#change name of variable in humidex dataset
humidexwa <- humidexwa %>%
  rename(zip = zipcode)

#change both zipcodes to character to match
phis$zip <- as.character(phis$zip)
humidexwa$zip <- as.character(humidexwa$zip)

class(phis$zip)
class(humidexwa$zip)

##Date correction
#change both dates to same format to match
phis$date <- as.POSIXct(phis$date)
humidexwa$date <- as.POSIXct(humidexwa$date)

class(phis$date)
class(humidexwa$date)

summary(phis$date)
summary(humidexwa$date)

## ---------------------------------------------------
##Checking zipcodes in the phis dataset
#Loading in dataset
wazips <- read_excel("wazips_counties_only.xlsx")
summary(wazips)

#Checking class of zipcodes
class(wazips$zip)

#converting to char
wazips$zip <- as.character(wazips$zip)

## ---------------------------------------------------
##Attempting merge only on WA zip codes
phis_wazips = merge(x = phis, y = wazips, by = "zip", all.x= TRUE)
summary(phis_wazips)
View(phis_wazips)

sum(is.na(phis_wazips$state_name)) #1798
sum(is.na(phis_wazips$pm25_zipwt)) #4515
sum(is.na(phis_wazips$hcia_pat_id)) #0

#Removing all NAs for WA state
phis_onlywazips <- phis_wazips %>% drop_na(state_name)
```

```
summary(phis_onlywazips)
View(phis_onlywazips)

sum(is.na(phis_onlywazips$state_name)) #0
sum(is.na(phis_onlywazips$pm25_zipwt)) #2901
sum(is.na(phis_onlywazips$hcia_pat_id)) #0

## --------------------------------------------------
##Merging phis with ONLY WA zips with humidex dataset

#Checking for NA in humidex dataset
sum(is.na(humidexwa$Humidex)) #no NA --> YEE HAW!!!!

#Checking classes for merge
class(phis_onlywazips$zip)
class(humidexwa$zip)

class(phis_onlywazips$date)
class(humidexwa$date)

#Merge
waphis_humidex = merge(x = phis_onlywazips, y = humidexwa, by = c("date", "zip"), all.x
= TRUE)

summary(waphis_humidex)
View(waphis_humidex)

sum(is.na(waphis_humidex$pm25_zipwt)) #2906
sum(is.na(waphis_humidex$Humidex)) #2896
sum(is.na(waphis_humidex$hcia_pat_id)) #0

## --------------------------------------------------
##Removing all ages above 19
waphis_upto19 <- subset(waphis_humidex, age <= 19)

summary(waphis_upto19)
sum(is.na(waphis_upto19$pm25_zipwt)) #1618
sum(is.na(waphis_upto19$Humidex)) #1294

## --------------------------------------------------
##Descriptive Statistics for all binary/categorical variables
#Descriptive Statistics for all WA PHIS
table(waphis_humidex$p_type)
table(waphis_humidex$sex_group)
table(waphis_humidex$race_group)
```

```
table(waphis_humidex$pay_group)
table(waphis_humidex$can_v32)

#Descriptive Statistics for all WA PHIS upto age 19
table(waphis_upto19$p_type)
table(waphis_upto19$sex_group)
table(waphis_upto19$race_group)
table(waphis_upto19$pay_group)
table(waphis_upto19$can_v32)
View(waphis_upto19$pm25_zipwt)

## ---------------------------------------------------
## SAVE final merged file with ages up to 19

waphis <- waphis_upto19

## Change working directory to save in datasets folder
projectWD <- "C:/Users/daani/Downloads/phis/datasets"
wd <- getwd()
setwd(projectWD)
save(waphis, file = "waphis11.rda")

## ---------------------------------------------------
## Finding out where most PM2.5 NAs are

load(file = "waphis2.rda")
summary(waphis)

## ---------------------------------------------------
## Finding out where most PM2.5 NAs are

waphis_pm2.5NA <- waphis[is.na(waphis$pm25_zipwt), ]
View(waphis_pm2.5NA$zip)
table(waphis_pm2.5NA$date)#they are mostly July 4th or 5th

## ---------------------------------------------------
## Removing July 4th and 5th

waphis.nofireworks2 <- waphis[!as.Date(waphis$date) %in% as.Date(c("2006-07-04",
"2006-07-05",
                                            "2007-07-04", "2007-07-05",
                                            "2008-07-04", "2008-07-05",
                                            "2009-07-04", "2009-07-05",
                                            "2010-07-04", "2010-07-05",
                                            "2011-07-04", "2011-07-05",
                                            "2012-07-04", "2012-07-05",
```

```
                                                "2013-07-04", "2013-07-05",
                                                "2014-07-04", "2014-07-05",
                                                "2015-07-04", "2015-07-05",
                                                "2016-07-04", "2016-07-05",
                                                "2017-07-04", "2017-07-05",
                                                "2018-07-04", "2018-07-05",
                                                "2019-07-04", "2019-07-05",
                                                "2020-07-04", "2020-07-05")), ]

summary(waphis.nofireworks2$pm25_zipwt) #40 only YAY

##Findings NAs for new dataset

waphis_nofireworks_pm2.5NA <-
waphis.nofireworks2[is.na(waphis.nofireworks2$pm25_zipwt), ]
View(waphis_pm2.5NA$zip)
table(waphis_nofireworks_pm2.5NA$date)
table(waphis_nofireworks_pm2.5NA$zip)

##Isolating the date and zip of the PM2.5 NA values
library(dplyr)

day_zip_pm2.5NA <- waphis_nofireworks_pm2.5NA %>% select(date, zip)

write.csv(day_zip_pm2.5NA, "day_zip_pm2.5NA.csv")

##SAVE dataset
waphis_rm070405 <- waphis.nofireworks2
save(waphis_rm070405, file = "waphis_rm070405.rda")
summary(waphis_rm070405)

##Descriptive stats for new dataset
table(waphis_rm070405$p_type)
table(waphis_rm070405$sex_group)
table(waphis_rm070405$race_group)
table(waphis_rm070405$pay_group)
table(waphis_rm070405$can_v32)
View(waphis_rm070405)

sum(is.na(waphis_rm070405_v2$Humidex))
day_zip_pm2.5NA2 <- waphis_nofireworks_pm2.5NA %>% select(date, zip, Humidex,
RH_pct, Temp_C )
View(day_zip_pm2.5NA2)

## ---------------------------------------------------
## Finding out where most Humidex NAs are
```

```
load(file = "waphis_rm070405_v2.rda")
summary(waphis_rm070405_v2)
View(waphis_rm070405_v2)

waphis_HumidexNA <- waphis_rm070405_v2[is.na(waphis_rm070405_v2$Humidex), ]
View(waphis_HumidexNA$Humidex)
table(waphis_HumidexNA$date)

##Isolating the date and zip of the Humidex NA values
library(dplyr)

day_zip_HumidexNA <- waphis_HumidexNA %>% select(date, zip, RH_pct, Temp_C)

View(day_zip_HumidexNA)

write.csv(day_zip_HumidexNA, "day_zip_HumidexNA.csv")

## -----------------------------------------------------
##Thoughts on converting PM2.5 to binary smoke-day
#converting continuous PM2.5 to binary smoke_day

load(file = "waphis_rm070405_v2.rda")
pm25_zipwt_new <- ifelse(waphis_rm070405_v2$pm25_zipwt > 20.4, 1, 0)
table(pm25_zipwt_new, waphis_rm070405_v2$pm25_zipwt, useNA = "ifany")
table(pm25_zipwt_new)
##Make a new variable; assign a new variable name and assign it to if/else statement
summary(pm25_zipwt_new == 1)
summary(waphis_rm070405_v2$pm25_zipwt > 20.4)

summary(waphis_rm070405_v2)

#confirming if conversion was done correctly
#ensure NAs are coded as missing
tapply(waphis_rm070405_v2$pm25_zipwt, pm25_zipwt_new, summary)

load(file = "waphis_rm070405.rda")
summary(waphis_rm070405$pm25_zipwt)

View(waphis_rm070405_v2)

##-----------------------------------------------------
##Reading in datasets

##Prep the dataset from part 01 (where July 4th and 5th were removed)
waphis_noNA <- waphis_rm070405
```

```
##Read in the datasets to fill in
pm25_fill <- read.csv("pm2.5_fill.csv")
humidex_fill <- read.csv("humidex_fill.csv")

##---------------------------------------------------
#PM2.5 Merge
##---------------------------------------------------

##---------------------------------------------------
##Add a column that has row counts for ID merge
waphis_noNA$ID <- as.numeric(rownames(waphis_noNA))
class(waphis_noNA$ID) #make sure class is numeric

##Add the missing row numbers from the PM2.5 NA dataset
pm25_fill$ID <- c(24132, 24133, 24346, 24560, 24561, 24883, 25120, 26005,
            26006, 26115, 26617, 28299, 28301, 28774, 28875, 29339,
            42190, 56227, 105463, 124161, 125158, 125159, 125202,
            125203, 125216, 132828, 153650, 167122, 167123, 168814,
            168815, 168938, 168939, 177704, 177705, 177706, 177707,
            192273, 194996, 201585)

class(pm25_fill$ID) #make sure this is numeric as well

##Remove extraneous variables from pm fill dataset
pm_fill <- subset(pm25_fill, select = -c(X, zip, date))

##---------------------------------------------------
##Merge
waphis_pm25merge <- merge(x = waphis_noNA, y = pm_fill, by = "ID", all.x = TRUE)

##Make a new variable with the coalesce -- this works with no NA
pm25_zipwt_new <- coalesce(waphis_mergepls$pm25_zipwt.x,
waphis_mergepls$pm25_zipwt.y)

#Add this new variable to the waphis dataset
waphis_pm25merge$pm25_zipwt <- pm25_zipwt_new
#Check if there are any NAs
sum(is.na(waphis_pm25merge$pm25_zipwt)) #0

#Drop the useless columns now
waphis_NAhum <- subset(waphis_pm25merge, select = -c(pm25_zipwt.x, pm25_zipwt.y))

##---------------------------------------------------
#Removing NA for humidex
##---------------------------------------------------
```

```
##----------------------------------------------------

#Load fill datasets
humidex_fill$ID <- c(24132, 24133, 24346, 24560, 24561, 24883, 25120, 26005,
                  26006, 26115, 26617, 28299, 28301, 28774, 28875, 29339,
                  42190, 56227, 105463, 124161, 125158, 125159, 125202,
                  125203, 125216, 132828, 153650,192273, 194996, 201585)

class(hum_fill2$ID) #make sure numeric

#Remove extraneous variables from humidex fill dataset
humi_fill <- subset(humidex_fill, select = -c(zip, date))

##----------------------------------------------------
##Merge
waphis_humimerge <- merge(x = waphis_NAhum, y = humi_fill, by = "ID", all.x = TRUE)

#Making new variables with coalesce
humi_new <- coalesce(waphis_humimerge$Humidex.x, waphis_humimerge$Humidex.y)
RH_new <- coalesce(waphis_humimerge$RH_pct.x, waphis_humimerge$RH_pct.y)
C_new <- coalesce(waphis_humimerge$Temp_C.x, waphis_humimerge$Temp_C.y)

waphis_humimerge$Humidex <- humi_new
waphis_humimerge$RH_pct <- RH_new
waphis_humimerge$Temp_C <- C_new

#Checking for NA
sum(is.na(waphis_humimerge$Humidex)) #0
sum(is.na(waphis_humimerge$RH_pct)) #0
sum(is.na(waphis_humimerge$Temp_C)) #0

#Drop the unused columns now
waphis_complete <- subset(waphis_humimerge, select = -c(Humidex.x, RH_pct.x,
Temp_C.x,
                              Humidex.y, RH_pct.y, Temp_C.y, ID))
summary(waphis_complete)

##----------------------------------------------------
#Saving final file with NO NAs
##----------------------------------------------------
save(waphis_complete, file = "waphis_complete_jan22.rda")

## ----------------------------------------------------
## Running replicate analysis on patient ID
## ----------------------------------------------------
```

```
#Finding duplicates
sum(duplicated(waphis_complete$hcia_pat_id)) #5 repeat patients
waphis_complete$hcia_pat_id[duplicated(waphis_complete$hcia_pat_id)]

#Vector of unique patient IDs
waphis_unique <- unique(waphis_complete$hcia_pat_id) #197,628 unique patients

#Making a dataset of only unique patients (no repeat)
waphis_uni <- waphis_complete %>% distinct(hcia_pat_id, .keep_all = TRUE)

## ---------------------------------------------------
## Looking at trends of first 4 days of June
## ---------------------------------------------------

##Attempt 2 --> can only get 1 day at a time (06-01, 06-02, etc.)

##Making datasets of each lag day in June
waphis_june01 <- waphis_uni[grep("06-01", as.character(waphis_uni$date)), ]
waphis_june02 <- waphis_uni[grep("06-02", as.character(waphis_uni$date)), ]
waphis_june03 <- waphis_uni[grep("06-03", as.character(waphis_uni$date)), ]
waphis_june04 <- waphis_uni[grep("06-04", as.character(waphis_uni$date)), ]

##Combining them into one dataset
waphis_juneall <- rbind(waphis_june01, waphis_june02, waphis_june03, waphis_june04)
#There are 7329 observations total
summary(waphis_juneall$pm25_zipwt)
summary(waphis_uni$pm25_zipwt)

#Removing June 01 - 04 from dataset; 190,299 observations total
waphis_0605 <- anti_join(waphis_uni, waphis_juneall) #dplyr
View(waphis_0605)

## ---------------------------------------------------
## BONUS!!!! Adding smoke day column from Maria's code
## ---------------------------------------------------
#Need to rename the column in smoke day dataset as zip first
Zipcode_merge = rename(Zipcode_final, zip = zipcode)
View(Zipcode_merge)
#Save the file
save(Zipcode_merge, file = "Zipcode_merge.rda")

##checking classes
class(waphis_juneall$date) #originally POSIXct
class(Zipcode_merge$date) #originally char
waphis_juneall$date <- as.Date(waphis_juneall$date)
```

```
Zipcode_merge$date <- as.Date(Zipcode_merge$date)

class(waphis_juneall$zip) #originally char
class(Zipcode_merge$zip) #int
Zipcode_merge$zip <- as.character(Zipcode_merge$zip)

#dropping extraneous vars
sd_merge <- subset(Zipcode_merge, select = -c(pm25_zipwt, city, date_num))

#Merging on the removed june dates as a test
waphis_juneallsd = merge(x = waphis_juneall, y = sd_merge,
                 by = c("date", "zip"), all.x = TRUE)
View(waphis_juneallsd)
table(waphis_juneallsd$Smoke.day) #160 smoke days

## ---------------------------------------------------
##Repeat merge procedure for dataset with removed June dates

#fixing class
class(waphis_0605$date) #originally POSIXct
waphis_0605$date <- as.Date(waphis_0605$date)

#merge
waphis_sd = merge(x = waphis_0605, y = sd_merge,
             by = c("date", "zip"), all.x = TRUE)
View(waphis_sd)
table(waphis_sd$Smoke.day) #11,420 sd; 178,844 non-sd
waphis_0605_sd <- waphis_sd

#Saving final dataset without first 4 days of June
save(waphis_0605_sd, file = "waphis_0605_sd.rda")
write.csv(waphis_0605_sd, "waphis_0605_sd.csv")

load(file = "waphis_0605_sd.rda")

## ---------------------------------------------------
## Coding smoke days
#---------- Reading Dataset -----------------------

Zipcode_final <- read.csv("D:/Washington_Files/Autumn2021/BIOST590/Daaniya - Final Project/DataSets/Zipcode_wt_pm25_final.csv",header = TRUE)
#Zipcode_final <- Zipcode_final[,-1] #I always get and extra column when reading data with read.csv, so this might not be necessary for you

#--------- List of zipcodes that are considered Urban areas --------------------
```

#Lists found on google

Tacoma_zipcodes <- c(98402, 98403, 98404, 98405, 98406, 98407, 98408, 98409, 98416, 98418, 98421, 98422, 98424, 98444, 98445, 98465, 98466, 98467, 98499)

Seattle_zipcodes <- c(98101, 98102, 98103, 98104, 98105, 98106, 98107, 98108, 98109, 98112, 98115, 98116, 98117, 98118, 98119, 98121, 98122, 98125, 98126, 98133, 98134, 98136, 98144, 98146, 98154, 98164, 98174, 98177, 98178, 98195, 98199)

Spokane_zipcodes <- c(99026, 99201, 99202, 99203, 99204, 99205, 99207, 99223)

#----------Codifying smoke.day---

#We add the variable city, which indicates which urban area is the zipcode or if it is simply not urban
Zipcode_final$city <- ifelse(Zipcode_final$zipcode %in% Tacoma_zipcodes, "Tacoma",
ifelse(Zipcode_final$zipcode %in% Seattle_zipcodes, "Seattle",
ifelse(Zipcode_final$zipcode %in% Spokane_zipcodes, "Spokane", "Not-urban")))

Zipcode_final$date_num <- as.Date(Zipcode_final$date, "%Y-%m-%d")#Change the string variable date to an actual date variable

Zipcode_final <- Zipcode_final[order(Zipcode_final$date_num),]#We want the data to be grouped by zipcode and ordered by date
Zipcode_final <- Zipcode_final[order(Zipcode_final$zipcode),]

##

#We are creating a list of smoke.day status for the data set. We checked manually if the first two days]
#or the last two days were smoke days (they were not). The rest of days will be checked for smoke.day status
#depending on the PM2.5 values on that day, the two days prior and the two days after.
smoke.day <- c(0,0,sapply((3:(length(Zipcode_final$date)-2)), function(i){
if (Zipcode_final$pm25_zipwt[i]>20.4){
return(1) #It is a smoke day if it is above 20.4
} else if(Zipcode_final$pm25_zipwt[i]<9){
return (0) # It is a non-smoke day if it is below 9
} else {
#Checking the value of PM2.5 two days before
d_m2 <- ifelse((Zipcode_final$zipcode[i-2]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i-2]==Zipcode_final$date_num[i]-2),
Zipcode_final$pm25_zipwt[i-2],

```
                    ifelse((Zipcode_final$zipcode[i-
1]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i-1]==Zipcode_final$date_num[i]-
2),
                         Zipcode_final$pm25_zipwt[i-1],NA))
        #Checking the value of PM2.5 one day prior
        d_m1 <- ifelse((Zipcode_final$zipcode[i-
1]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i-1]==Zipcode_final$date_num[i]-
1),
                    Zipcode_final$pm25_zipwt[i-1],NA)
        #Checking the value of PM2.5 that same day
        d <- Zipcode_final$pm25_zipwt[i]
        #Checking the value of PM2.5 two days after
        d_p2 <-
ifelse((Zipcode_final$zipcode[i+2]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i+
2]==Zipcode_final$date_num[i]+2),
                    Zipcode_final$pm25_zipwt[i+2],

ifelse((Zipcode_final$zipcode[i+1]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i+
1]==Zipcode_final$date_num[i]+2),
                         Zipcode_final$pm25_zipwt[i+1],NA))
        #Checking the value of PM2.5 one day after
        d_p1 <-
ifelse((Zipcode_final$zipcode[i+1]==Zipcode_final$zipcode[i])&&(Zipcode_final$date_num[i+
1]==Zipcode_final$date_num[i]+1),
                    Zipcode_final$pm25_zipwt[i+1],NA)

        if(!is.na(d_m2) && !is.na(d_m1) && ((d_m2>=9) || (d_m1>=9)) && ((d_m2>15) ||
(d>15) || (d_m1>15))){
          #Checked that the days: two days prior + current day, hold the characteristics needed.
          return(1)
        }else if (!is.na(d_m1) && !is.na(d_p1) && ((d_m1>=9) || (d_p1>=9)) && ((d_m1>15) ||
(d>15) || (d_p1>15))){
          #Checked that the days: one day prior to one day after hold the characteristics needed.
          return(1)
        }else if(!is.na(d_p2) && !is.na(d_p1) && ((d_p2>=9) || (d_p1>=9)) && ((d_p2>15) ||
(d>15) || (d_p1>15))){
          #Checked that the days: two days after + current day, hold the characteristics needed.
          return(1)
        } else {
          #None of the potential groups of three days held the characteristics needed to be a smoke
day.
          return(0)
        }
      }
    }),0,0)
```

```
Zipcode_final$Smoke.day <- smoke.day #Adding the smoke.day status to the table.

#We have to check the third condition for potential smoke days in urban areas
#Creating data sets for each urban area
Tacoma_dataset <-
Zipcode_final[Zipcode_final$city=="Tacoma",c("date_num","pm25_zipwt")]
Seattle_dataset <-
Zipcode_final[Zipcode_final$city=="Seattle",c("date_num","pm25_zipwt")]
Spokane_dataset <-
Zipcode_final[Zipcode_final$city=="Spokane",c("date_num","pm25_zipwt")]

#Assess which zipcodes-dates report values of PM2.5 above 9 in each data set.
Tacoma_dataset$Over_9 <- 1*(Tacoma_dataset$pm25_zipwt>=9)
Seattle_dataset$Over_9 <- 1*(Seattle_dataset$pm25_zipwt>=9)
Spokane_dataset$Over_9 <- 1*(Spokane_dataset$pm25_zipwt>=9)

#Compute the percentage of zipcodes that report above 9 per date.
Tacoma_summary <- aggregate(.~ date_num, data =
Tacoma_dataset[,c("date_num","Over_9")], FUN = mean)
Seattle_summary <- aggregate(.~ date_num, data =
Seattle_dataset[,c("date_num","Over_9")], FUN = mean)
Spokane_summary <- aggregate(.~ date_num, data =
Spokane_dataset[,c("date_num","Over_9")], FUN = mean)

#Creating list of dates that do not hold the condition of at least 50% of values reported being
over 9 for each urban area
Tacoma_list_dates <- Tacoma_summary[Tacoma_summary$Over_9<0.5,]$date_num
Seattle_list_dates <- Seattle_summary[Seattle_summary$Over_9<0.5,]$date_num
Spokane_list_dates <- Spokane_summary[Spokane_summary$Over_9<0.5,]$date_num

#For each date without the third condition, each zipcode in the Tacoma area, and PM2.5 not
as high, we change smoke.day status to 0
for (dn in Tacoma_list_dates){
  Zipcode_final[(Zipcode_final$date_num==dn) & (Zipcode_final$city=="Tacoma") &
(Zipcode_final$pm25_zipwt<=20.4),]$Smoke.day <- 0
}

#For each date without the third condition, each zipcode in the Seattle area, and PM2.5 not as
high, we change smoke.day status to 0
for (dn in Seattle_list_dates){
  Zipcode_final[(Zipcode_final$date_num==dn) & (Zipcode_final$city=="Seattle") &
(Zipcode_final$pm25_zipwt<=20.4),]$Smoke.day <- 0
}

#For each date without the third condition, each zipcode in the Spokane area, and PM2.5 not
as high, we change smoke.day status to 0
```

```
for (dn in Spokane_list_dates){
  Zipcode_final[(Zipcode_final$date_num==dn) & (Zipcode_final$city=="Spokane") & (Zipcode_final$pm25_zipwt<=20.4),]$Smoke.day <- 0
}

############### Coding Exposure_k for lag analysis ########################

#Making sure the order is the correct order

Zipcode_final <- Zipcode_final[order(Zipcode_final$date_num),] #The data.frame should be ordered by date after ordering per zipcode
Zipcode_final <- Zipcode_final[order(Zipcode_final$zipcode),] #The data.frame should be ordered by zipcode

length(unique(Zipcode_final[,c("date_num","zipcode")])$date_num) #Making sure there are not pairs date-zipcode repeated

Exposure_1 <- sapply((1:length(Zipcode_final$X)), function(i){
  zc <- Zipcode_final$zipcode[i] #We save the current Zipcode
  date_current <- Zipcode_final$date_num[i] # and the current date
  list_sd_temp <- c(Zipcode_final$Smoke.day[i]) #Create a list of days we will check for any smoke day

  if (i > 1){ #We can only check the previous entry if i>1
    if ((Zipcode_final$zipcode[i-1]==zc) && ((date_current-1) <= Zipcode_final$date_num[i-1]) && (Zipcode_final$date_num[i-1] <= date_current)){
      list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-1]) #We add the previous entry smoke day only if the date and zipcode are the correct ones
    }
  }

  if (sum(list_sd_temp)>0){ #If any entry in the list of days to be checked is a smoke day, then the exposure is positive
    return(1)
  } else if (length(list_sd_temp)==2){ #Did we have enough information to check every day?
    return(0)
  } else { #If there was not enough days in our list of days, then we do not know the status of the exposure
    return(NA)
  }
})

Zipcode_final$Exposure_1 <- Exposure_1

Exposure_2 <- sapply((1:length(Zipcode_final$X)), function(i){
  zc <- Zipcode_final$zipcode[i] #We save the current Zipcode
```

```
    date_current <- Zipcode_final$date_num[i] # and the current date
    list_sd_temp <- c(Zipcode_final$Smoke.day[i]) #Create a list of days we will check for any
smoke day

    if (i > 1){ #We can only check the previous entry if i>1
      if ((Zipcode_final$zipcode[i-1]==zc) && ((date_current-2) <= Zipcode_final$date_num[i-
1]) && (Zipcode_final$date_num[i-1] <= date_current)){
        list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-1]) #We add the previous
entry smoke day only if the date and zipcode are the correct ones
      }
    }

    if (i > 2){ #We can only check two entries prior if i>2
      if ((Zipcode_final$zipcode[i-2]==zc) && ((date_current-2) <= Zipcode_final$date_num[i-
2]) && (Zipcode_final$date_num[i-2] <= date_current)){
        list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-2]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
      }
    }

    if (sum(list_sd_temp)>0){ #If any entry in the list of days to be checked is a smoke day,
then the exposure is positive
      return(1)
    } else if (length(list_sd_temp)==3){ #Did we have enough information to check every day?
      return(0)
    } else { #If there was not enough days in our list of days, then we do not know the status of
the exposure
      return(NA)
    }
  })

  Zipcode_final$Exposure_2 <- Exposure_2

  Exposure_3 <- sapply((1:length(Zipcode_final$X)), function(i){
    zc <- Zipcode_final$zipcode[i] #We save the current Zipcode
    date_current <- Zipcode_final$date_num[i] # and the current date
    list_sd_temp <- c(Zipcode_final$Smoke.day[i]) #Create a list of days we will check for any
smoke day

    if (i > 1){ #We can only check the previous entry if i>1
      if ((Zipcode_final$zipcode[i-1]==zc) && ((date_current-3) <= Zipcode_final$date_num[i-
1]) && (Zipcode_final$date_num[i-1] <= date_current)){
        list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-1]) #We add the previous
entry smoke day only if the date and zipcode are the correct ones
      }
    }
```

```
    if (i > 2){ #We can only check two entries prior if i>2
     if ((Zipcode_final$zipcode[i-2]==zc) && ((date_current-3) <= Zipcode_final$date_num[i-
2]) && (Zipcode_final$date_num[i-2] <= date_current)){
      list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-2]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
     }
    }

    if (i > 3){ #We can only check two entries prior if i>3
     if ((Zipcode_final$zipcode[i-3]==zc) && ((date_current-3) <= Zipcode_final$date_num[i-
3]) && (Zipcode_final$date_num[i-3] <= date_current)){
      list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-3]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
     }
    }

    if (sum(list_sd_temp)>0){ #If any entry in the list of days to be checked is a smoke day,
then the exposure is positive
     return(1)
    } else if (length(list_sd_temp)==4){ #Did we have enough information to check every day?
     return(0)
    } else { #If there was not enough days in our list of days, then we do not know the status of
the exposure
     return(NA)
    }
   })

   Zipcode_final$Exposure_3 <- Exposure_3
   Exposure_4 <- sapply((1:length(Zipcode_final$X)), function(i){
    zc <- Zipcode_final$zipcode[i] #We save the current Zipcode
    date_current <- Zipcode_final$date_num[i] # and the current date
    list_sd_temp <- c(Zipcode_final$Smoke.day[i]) #Create a list of days we will check for any
smoke day

    if (i > 1){ #We can only check the previous entry if i>1
     if ((Zipcode_final$zipcode[i-1]==zc) && ((date_current-4) <= Zipcode_final$date_num[i-
1]) && (Zipcode_final$date_num[i-1] <= date_current)){
      list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-1]) #We add the previous
entry smoke day only if the date and zipcode are the correct ones
     }
    }

    if (i > 2){ #We can only check two entries prior if i>2
     if ((Zipcode_final$zipcode[i-2]==zc) && ((date_current-4) <= Zipcode_final$date_num[i-
2]) && (Zipcode_final$date_num[i-2] <= date_current)){
```

```
            list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-2]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
          }
        }

        if (i > 3){ #We can only check two entries prior if i>3
          if ((Zipcode_final$zipcode[i-3]==zc) && ((date_current-4) <= Zipcode_final$date_num[i-
3]) && (Zipcode_final$date_num[i-3] <= date_current)){
            list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-3]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
          }
        }

        if (i > 4){ #We can only check two entries prior if i>3
          if ((Zipcode_final$zipcode[i-4]==zc) && ((date_current-4) <= Zipcode_final$date_num[i-
4]) && (Zipcode_final$date_num[i-4] <= date_current)){
            list_sd_temp <- c(list_sd_temp,Zipcode_final$Smoke.day[i-4]) #We add two entries
prior smoke day only if the date and zipcode are the correct ones
          }
        }

        if (sum(list_sd_temp)>0){ #If any entry in the list of days to be checked is a smoke day,
then the exposure is positive
          return(1)
        } else if (length(list_sd_temp)==5){ #Did we have enough information to check every day?
          return(0)
        } else { #If there was not enough days in our list of days, then we do not know the status of
the exposure
          return(NA)
        }
      })

      Zipcode_final$Exposure_4 <- Exposure_4

      Zipcode_final <- Zipcode_final[,-1] #Remove first column, which is not necessary

      write.csv(Zipcode_final,"D:/Washington_Files/Autumn2021/BIOST590/Daaniya - Final
Project/DataSets/Zip_code_WithExposuresk.csv")

      ## ---------------------------------------------------
      ##---------------------------------------------------
      #Making subgroups
      ##---------------------------------------------------

      ## Age ----------------------------------------------
```

```
waphis_patid$age.cat <- with(waphis_patid,
                        ifelse(age < 6, 1,
                              ifelse(age >= 6 & age < 13, 2,
                                    ifelse(age >= 13 & age < 20, 3, 0))))

table(waphis_patid$age.cat)

## Race ----------------------------------------------

table(waphis_patid$race_group)

waphis_patid$race.cat <- with(waphis_patid,
                         ifelse(race_group == "a. Non-Hisp White", 1,
                               ifelse(race_group == "b. Non-Hisp Black", 2,
                                     ifelse(race_group == "c. Hispanic", 3,
                                           ifelse(race_group == "d. Asian", 4,
                                                 ifelse(race_group == "e. Other", 5, 0))))))

table(waphis_patid$race.cat) #Matches table above

## Health Outcomes ------------------------------------

#Respiratory:113, 130 - 134, 137 - 144

class(waphis_patid$adrg_v32)

waphis_patid$resp.o <- with(waphis_patid,
                       ifelse(adrg_v32 %in% c(113, 130, 131, 132, 133,
                                              134, 137, 138, 139, 140,
                                              141, 142, 143, 144), 1, 0))

table(waphis_patid$resp.o) #27,270 cases total

#Respiratory infections subgroup: 113, 137 - 139

waphis_patid$resp_infections.o <- with(waphis_patid,
                                 ifelse(adrg_v32 %in% c(113, 137, 138, 139), 1, 0))

table(waphis_patid$resp_infections.o) #16,065 cases

#Dermal: 381, 385

waphis_patid$derm.o <- with(waphis_patid,
                       ifelse(adrg_v32 %in% c(381, 385), 1, 0))

table(waphis_patid$derm.o) #5802 cases total
```

```
#Trauma: 55 - 57, 135, 340 - 342, 347, 351, 384

waphis_patid$trauma.o <- with(waphis_patid,
                    ifelse(adrg_v32 %in% c(55, 56, 57, 135, 340, 341,
                                    342, 347, 351, 384), 1, 0))
table(waphis_patid$trauma.o) #26,619 cases total

#Mental health: 750 - 760

waphis_patid$mh.o <- with(waphis_patid,
                  ifelse(adrg_v32 %in% c(750, 751, 752, 753, 754, 755,
                                  756, 757, 758, 759, 760), 1, 0))

table(waphis_patid$mh.o)

##----------------------------------------------------
## Save file

waphis_cat.o <- waphis_patid
save(waphis_cat.o, file = "waphis_cat.o.rda")

##----------------------------------------------------
## Making a dataset of only unique patients (no repeat)
waphis_uni <- waphis_complete %>% distinct(hcia_pat_id, .keep_all = TRUE)
summary(waphis_uni)

## Load in patient ID dataset
phis_patid <- read_excel("phis_patid.xlsx")
summary(phis_patid)
View(phis_patid)

## Looking to see how many distinct patients there are --> there's 129529 :((((
phis_patid_uni <- phis_patid %>% distinct(patient_id, .keep_all = TRUE)
View(phis_patid_uni)

##----------------------------------------------------
## Merge subgroups
##----------------------------------------------------
## Try 1
waphis_patid <- merge(x = waphis_uni, y = phis_patid, by = "hcia_pat_id", all.x = TRUE)
summary(waphis_patid$patient_id)
View(waphis_patid)

## Save this dataset
save(waphis_patid, file = "waphis_patid.rda")
```

```
##----------------------------------------------------
## Keep only the unique
waphis_patid_uni <- waphis_patid %>% distinct(patient_id, .keep_all = TRUE)
summary(waphis_patid_uni)

## Keep this dataset
save(waphis_patid_uni, file = "waphis_patid_uni.rda")
##----------------------------------------------------
## Purpose: Create referent days

## ----------------------------------------------------
## Loading dataset that has duplicates removed and subgroups made
load(file = "waphis_cat.o.rda")
View(waphis_cat.o)

##----------------------------------------------------
## Make the First Data Set  --> need to add case
##----------------------------------------------------

waphis_cat.o2 <- waphis_cat.o #Just to have a copy of the original data

#MARIA COMMENT
#I realized waphis_cat.0 had duplicates to begin with. So it seems to me you actually have
197236 cases in total (not 197628)
waphis_cat.o <- waphis_cat.o[!duplicated(waphis_cat.o[,c("patient_id","date")]),]

waphis_mergeA <- waphis_cat.o

## Add case column

waphis_mergeA$case <- 1

##----------------------------------------------------
## Make the Second Data Set  --> Make referents
##----------------------------------------------------

## Making refs table
## Referent days for 3 weeks before and ahead

waphis_cat.o$ref.day1 <- as.factor(ymd(waphis_cat.o$date) - 28)
waphis_cat.o$ref.day2 <- as.factor(ymd(waphis_cat.o$date) - 21)
waphis_cat.o$ref.day3 <- as.factor(ymd(waphis_cat.o$date) - 14)
waphis_cat.o$ref.day4 <- as.factor(ymd(waphis_cat.o$date) - 7)
waphis_cat.o$ref.day5 <- as.factor(ymd(waphis_cat.o$date) + 7)
waphis_cat.o$ref.day6 <- as.factor(ymd(waphis_cat.o$date) + 14)
```

```
waphis_cat.o$ref.day7 <- as.factor(ymd(waphis_cat.o$date) + 21)
waphis_cat.o$ref.day8 <- as.factor(ymd(waphis_cat.o$date) + 28)
waphis_cat.o$index.month <- month(waphis_cat.o$date)

#Rename variable

setnames(waphis_cat.o, old=c("date", "ref.day1", "ref.day2", "ref.day3",
                             "ref.day4", "ref.day5", "ref.day6", "ref.day7",
                             "ref.day8"),
         new=c("lag0.index", "lag0.ref.day1", "lag0.ref.day2",
               "lag0.ref.day3", "lag0.ref.day4", "lag0.ref.day5",
               "lag0.ref.day6", "lag0.ref.day7", "lag0.ref.day8"))

## Checking attributes of each day --> they are all factors
class(waphis_cat.o$lag0.index) #POSICxt --> change to factor

waphis_cat.o$lag0.index <- as.factor(ymd(waphis_cat.o$lag0.index))

## ---------------------------------------------------
## Converting data format wide to long

# Need to relocate lag0.index first --> dplyr
waphis_cat.o <- waphis_cat.o %>% relocate(lag0.index, .before = lag0.ref.day1)

# Pivot longer
waphis_cases_longer <- waphis_cat.o %>%
  pivot_longer(70:78, names_to = "day", values_to = "date")

View(waphis_cases_longer)


## Using lag0.index to see how many cases I have --> still 197,628
cases_only <- waphis_cases_longer[waphis_cases_longer$day %in% "lag0.index",]

#MARIA COMMENT
#My own code to check how many of each type of day there are
table(waphis_cases_longer$day)

## ---------------------------------------------------
## Removing referent days in May and October
waphis_cases_longer$month <- month(waphis_cases_longer$date)

waphis_referents <- subset(waphis_cases_longer, waphis_cases_longer$month >5
                           & waphis_cases_longer$month <10)

## ---------------------------------------------------
```

```
## Removing referent days in May and October
waphis_ref_select <- subset(waphis_referents,
                    waphis_referents$month == waphis_referents$index.month)

#MARIA COMMENT
#My own review
table(waphis_ref_select$day, useNA = "ifany")

## Add row number
#MARIA COMMENT
waphis_ref_select$row_num <- 1:nrow(waphis_ref_select) #Same as what you had (do not
remember the original code)

##---------------------------------------------------
## Creating second dataset for merge
##---------------------------------------------------

waphis_mergeB <- waphis_ref_select

## Using lag0.index to see how many cases I have --> still 197,628
cases_only <- waphis_mergeB[waphis_mergeB$day %in% "lag0.index",]

table(waphis_mergeB$day, useNA = "ifany")

#MARIA COMMENT
waphis_mergeB$day <- factor(waphis_mergeB$day, levels = c("lag0.index", "lag0.ref.day1",
"lag0.ref.day2",
                                        "lag0.ref.day3", "lag0.ref.day4", "lag0.ref.day5",
                                        "lag0.ref.day6", "lag0.ref.day7", "lag0.ref.day8"))
#Factoring to control the order

waphis_mergeB <- waphis_mergeB[order(waphis_mergeB$day),] #Ordering per day type
waphis_mergeB <- waphis_mergeB[order(waphis_mergeB$patient_id),] #Ordering per
patient_id.

#This ensures that "lag0.index" are the first to show up, and will not be removed when
removing duplicates.

# Remove duplicated patients
waphis_mergeB <- waphis_mergeB[!duplicated(waphis_mergeB[,c("patient_id", "date")]),]


## Using lag0.index to see how many cases I have --> now 194,997
cases_only <- waphis_mergeB[waphis_mergeB$day %in% "lag0.index",] #194997

#MARIA COMMENT
```

#My own review<- now with the previous changes, the lag0.index remains the same number (197236 as in waphis_mergeA) :)

```
table(waphis_mergeB$day, useNA = "ifany")

##----------------------------------------------------
## M E R G E
##----------------------------------------------------
## Checking classes
class(waphis_mergeA$date)

class(waphis_mergeB$date) ##Needs to be factor

# Change class
waphis_mergeA$date <- as.factor(ymd(waphis_mergeA$date))

##----------------------------------------------------
## Merge

waphisC <- merge(waphis_mergeA[,c("patient_id", "date", "case")], waphis_mergeB,
            by = c("patient_id", "date"), all.x = TRUE, all.y = TRUE)

waphisC_all <- waphisC
waphisC_all <- waphisC %>% arrange(row_num)

#MARIA COMMENT
#My review<- It looks correct to me. Only cases are the 197236 corresponding to lag0.index

table(waphisC_all$day,waphisC_all$case, useNA = "ifany")

#MARIA COMMENT

## Using lag0.index to see how many cases I have --> now 195,384
cases_only <- waphisC_all[waphisC_all$day %in% "lag0.index",] #195384

## Using case == 1 to see how many cases I have --> 197,628
cases_only <- waphisC_all[waphisC_all$case %in% 1,]

## Now seeing how many cases are correctly attributed to lag0.index
summary(cases_only$day %in% "lag0.index" == TRUE)
# Result --> 195,384 of cases are lag0.index, we have 2244 misattributed ones

##----------------------------------------------------
## Seeing if one patient has no referent days

patients_list <- unique(waphis_mergeA$patient_id)
```

```
patients_list_controls <- unique(waphisC[is.na(waphisC$case),]$patient_id)

##----------------------------------------------------
## Assign 0 to NA

waphisC_all$case[is.na(waphisC_all$case)] = 0

waphis_refs_final <- waphisC_all

View(waphis_refs_final)

##----------------------------------------------------
## Save file
save(waphis_refs_final, file = "waphis_refs_final.rda")
View(waphis_refs_final)

##----------------------------------------------------
## ---------------------------------------------------
## Loading dataset that has referent days and row numbers
load(file = "waphis_refs_final.rda")
View(waphis_refs_final)

## Load zipcode and exposure dataset
zip_sd <- read_csv("Zipcode_exposure_sd.csv")

## Missing humidex --> need to add in these values for each date
## Load humidex
totalhumidex <- read_csv("master_humidex.csv")

## Drop useless columns
totalhumidex <- select(totalhumidex, -ZCTA)

##----------------------------------------------------
## Add Humidex vars to Zipcode exposure data
##----------------------------------------------------
## Check classes for all merge vars (date, zipcode)
class(totalhumidex$date)
class(totalhumidex$zipcode)

class(zip_sd$date)
class(zip_sd$zipcode)
##----------------------------------------------------
## Merge
zip_sd_humi <- merge(x = zip_sd, y = totalhumidex,
                 by = c("date", "zipcode"), all.x = TRUE)
```

```
View(zip_sd_humi)

## Drop useless columns
zip_sd_humi <- select(zip_sd_humi, -X1.y)

##-----------------------------------------------------
## Merge refs with Zip and env dataset
##-----------------------------------------------------
## Remove PM2.5, RH, Humidex, humidity so we can merge with exposure
waphis_refs_noexp <- select(waphis_refs_final, -c(pm25_zipwt, Humidex, RH_pct,
Temp_C))

## Checking classes
class(waphis_refs_noexp$date)
class(waphis_refs_noexp$zip) ##Needs to be numeric

class(zip_sd_humi$date) ##Needs to be factor
class(zip_sd_humi$zipcode)

# Change class
zip_sd_humi$date <- as.factor(ymd(zip_sd_humi$date))
waphis_refs_noexp$zip <- as.numeric(waphis_refs_noexp$zip)

## Rename zip_sd_humi zipcode
zip_sd_humi <- zip_sd_humi %>%
  rename(zip = zipcode)

##-----------------------------------------------------
## Merge
waphis_sd <- merge(x = waphis_refs_noexp, y = zip_sd_humi,
                   by = c("date", "zip"), all.x = TRUE)
View(waphis_sd)

## Reorder by row number to get cases and referents together again
# Confirm they are out of order --> yep!
View(waphis_sd$row_num)

# Reorder
waphis_sd <- waphis_sd %>% arrange(row_num)

##----------------------------------------------------
## Removing all the necessary cases
##-----------------------------------------------------
## Remove July 4th and 5th from referents
```

```
waphis_sd_noJul <- waphis_sd[!as.Date(waphis_sd$date) %in% as.Date(c("2006-07-04",
"2006-07-05",
                                    "2007-07-04", "2007-07-05",
                                    "2008-07-04", "2008-07-05",
                                    "2009-07-04", "2009-07-05",
                                    "2010-07-04", "2010-07-05",
                                    "2011-07-04", "2011-07-05",
                                    "2012-07-04", "2012-07-05",
                                    "2013-07-04", "2013-07-05",
                                    "2014-07-04", "2014-07-05",
                                    "2015-07-04", "2015-07-05",
                                    "2016-07-04", "2016-07-05",
                                    "2017-07-04", "2017-07-05",
                                    "2018-07-04", "2018-07-05",
                                    "2019-07-04", "2019-07-05",
                                    "2020-07-04", "2020-07-05")), ]

sum(is.na(waphis_sd_noJul$Humidex)) #183

##----------------------------------------------------
## Removing COVID cases
waphis_sd_nocov <- subset(waphis_sd_noJul, !dx1 %in% "U071")

##----------------------------------------------------
## Filling in the 183 NAs
##----------------------------------------------------
## Load fill dataset
## load in dataset
day_zip_fill_final <- read_csv("day_zip_fill_final.csv")

## Turn this into a date and then try factor conversion
day_zip_fill_final$date <- as.Date(day_zip_fill_final$date,format="%m/%d/%Y")
day_zip_fill_final$date <- as.factor(ymd(day_zip_fill_final$date))

##----------------------------------------------------
## There is one ref day exposure NA for hcia_pat_id 77568005 where data not found
## Since that case has 4 ref days, we will remove that one
waphis_sd_full <- waphis_sd_nocov[waphis_sd_nocov$row_num != 307691, ]

# remove it from zip too
day_zip_fill_final <- day_zip_fill_final[day_zip_fill_final$row_num != 308174, ]

## Checking the class
class(waphis_sd_full$date)
class(day_zip_fill_final$date)
```

```
class(waphis_sd_full$zip)
class(day_zip_fill_final$zip)

## Creating a new dataset to test
waphis_tofill <- waphis_sd_full
exp_fill <- day_zip_fill_final

##-----------------------------------------------------
## Use devtools package to.fill to fill in NAs
## Need to make dataset to prevent duplicate columns
fillpm25 <- exp_fill %>% select(date, zip, hcia_pat_id, pm25_zipwt)
fillrh <- exp_fill %>% select(date, zip, hcia_pat_id, rh_zipwt)
filltemp <- exp_fill %>% select(date, zip, hcia_pat_id, temp_zipwt)
fillhum <- exp_fill %>% select(date, zip, hcia_pat_id, hum_zipwt)

# Fill in PM2.5 missing values from waphis with values from exp
waphis_tofill <- FillIn(D1 = waphis_tofill, D2 = fillpm25,
                        Var1 = "pm25_zipwt", Var2 = "pm25_zipwt",
                        KeyVar = c("hcia_pat_id", "date", "zip"))

# Check!
sum(is.na(waphis_tofill$pm25_zipwt))

# Fill in RH missing values from waphis with values from exp
waphis_tofill <- FillIn(D1 = waphis_tofill, D2 = fillrh,
                        Var1 = "RH_pct", Var2 = "rh_zipwt",
                        KeyVar = c("hcia_pat_id", "date", "zip"))

# Fill in Temp_C missing values from waphis with values from exp
waphis_tofill <- FillIn(D1 = waphis_tofill, D2 = filltemp,
                        Var1 = "Temp_C", Var2 = "temp_zipwt",
                        KeyVar = c("hcia_pat_id", "date", "zip"))

# Fill in RH missing values from waphis with values from exp
waphis_tofill <- FillIn(D1 = waphis_tofill, D2 = fillhum,
                        Var1 = "Humidex", Var2 = "hum_zipwt",
                        KeyVar = c("hcia_pat_id", "date", "zip"))

sum(is.na(waphis_tofill$pm25_zipwt))
sum(is.na(waphis_tofill$RH_pct))
sum(is.na(waphis_tofill$Temp_C))
sum(is.na(waphis_tofill$Humidex))

waphis_tofill <- waphis_tofill %>% arrange(row_num)

## We now need to classify smoke day only
```

```
##----------------------------------------------------
## Saving final dataset
##----------------------------------------------------
## Check to remove useless columns
View(waphis_filled_nosd)

save(waphis_filled_nosd, file = "waphis_filled_nosd.rda")

##----------------------------------------------------
## Making and running the models to test for OR

load(file = "waphis_filled2_03.24.rda")
load(file = "waphis_filled_ED.rda")
load(file = "waphis_filled_nonED.rda")
waphism <- waphis_filled2_03.24
waphisED <- waphis_filled_ED
waphisIO <- waphis_filled_nonED

cases_onlyO <- waphism[waphism$case %in% 1,]
cases_onlyED <- waphisED[waphisED$case %in% 1,]
cases_onlyIO <- waphisIO[waphisIO$case %in% 1,]

table(cases_onlyO$age > 0)

##----------------------------------------------------
## Creating output table for summary
##----------------------------------------------------
## Function for OR table sig
OR.bin.summ <- function(model) {
  coef <- as.data.frame(coef(model))
  conf <- as.data.frame(confint(model))
  OR <- round(exp(coef[1, 1]), 3)
  LCL <- round(exp(conf[1, 1]), 3)
  UCL <- round(exp(conf[1, 2]), 3)
  p <- round(summary(model)$coef[1,5], 3)
  sig <- ifelse(p<0.001, "***", ifelse((p<0.01),"**", ifelse((p<0.05),"*",
ifelse((p<0.1),".",""))))
  table <- cbind(OR, LCL, UCL, p, sig)
  return(table)
}

##----------------------------------------------------
## Models for AIM 1
##----------------------------------------------------
```

```
model1.1O <- clogit(case ~ Smoke.day + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphism)

model1.2ED <- clogit(case ~ Smoke.day + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisED)

model1.3IO <- clogit(case ~ Smoke.day + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisIO)

overall <- OR.bin.summ(model1.1O)
ED <- OR.bin.summ(model1.2ED)
other <- OR.bin.summ(model1.3IO)

output.summary(model1.1O)

Service <- c("All cause", "ED only", "Inpatient and Observation")
table1 <- rbind(overall, ED, other)
table.serv <- cbind(Service, table1)
View(table.serv)

output.summary(model1.1O, tablenum = "Table1O", tablename = "Table overall")

##-----------------------------------------------------
## Models for AIM 2
##-----------------------------------------------------

##-----------------------------------------------------
## Model 2: Ages
##-----------------------------------------------------
##-----------------------------------------------------
## Overall

## Ages 0 - 5##
model2.1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = age.cat == 1)

model2.1O

## Ages 6 - 12##
model2.2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = age.cat == 2)

## Ages 13 - 19##
model2.3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = age.cat == 3)
```

```
age1O <- OR.bin.summ(model2.1O)
age2O<- OR.bin.summ(model2.2O)
age3O <- OR.bin.summ(model2.3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
table2O <- rbind(age1O, age2O, age3O)
table.ageO <- cbind(Age, table2O)
View(table.ageO)

table(cases_onlyO$age.cat)
cases_onlyO %>%
  group_by( age.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyO ) )

##------------------------------------------------------
## ED

## Ages 0 - 5##
model2.1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = age.cat == 1)

## Ages 6 - 12##
model2.2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = age.cat == 2)

## Ages 13 - 19##
model2.3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = age.cat == 3)

age1ED <- OR.bin.summ(model2.1ED)
age2ED <- OR.bin.summ(model2.2ED)
age3ED <- OR.bin.summ(model2.3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
table2ED <- rbind(age1ED, age2ED, age3ED)
table.ageED <- cbind(Age, table2ED)
View(table.ageED)

table(cases_onlyED$age.cat)
cases_onlyED %>%
  group_by( age.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyED ) )

##------------------------------------------------------
## IO
```

```
## Ages 0 - 5##
model2.1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = age.cat == 1)

## Ages 6 - 12##
model2.2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = age.cat == 2)

## Ages 13 - 19##
model2.3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = age.cat == 3)

age1IO <- OR.bin.summ(model2.1IO)
age2IO <- OR.bin.summ(model2.2IO)
age3IO <- OR.bin.summ(model2.3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
table2IO <- rbind(age1IO, age2IO, age3IO)
table.ageIO <- cbind(Age, table2IO)
View(table.ageIO)

table(cases_onlyIO$age.cat)
cases_onlyIO %>%
  group_by( age.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyIO ) )

##---------------------------------------------------
## Model 3: Races
##---------------------------------------------------
##---------------------------------------------------
## Overall

## Race: Non-Hisp white##
model3.1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                   data= waphism, subset = race.cat == 1)

## Race: Non-Hisp Black
model3.2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                   data= waphism, subset = race.cat == 2)

## Race: Hispanic
model3.3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                   data= waphism, subset = race.cat == 3)

## Race: Asian
model3.4O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
```

```
            data= waphism, subset = race.cat == 4)

## Race: Other
model3.5O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = race.cat == 5)

race1O <- OR.bin.summ(model3.1O)
race2O <- OR.bin.summ(model3.2O)
race3O <- OR.bin.summ(model3.3O)
race4O <- OR.bin.summ(model3.4O)
race5O <- OR.bin.summ(model3.5O)

RaceO <- c("Non-Hisp white", "Non-Hisp Black", "Hispanic", "Asian",
       "Other")

table3O <- rbind(race1O, race2O, race3O, race4O, race5O)
table.raceO <- cbind(RaceO, table3O)
View(table.raceO)

table(cases_onlyO$race.cat)
cases_onlyO %>%
  group_by( race.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyO ) )

##----------------------------------------------------
## ED

## Race: Non-hisp white##
model3.1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphisED, subset = race.cat == 1)

## Race: Non-Hisp Black
model3.2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphisED, subset = race.cat == 2)

## Race: Hispanic
model3.3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphisED, subset = race.cat == 3)

## Race: Asian
model3.4ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphisED, subset = race.cat == 4)

## Race: Other
model3.5ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphisED, subset = race.cat == 5)
```

```
race1ED <- OR.bin.summ(model3.1ED)
race2ED <- OR.bin.summ(model3.2ED)
race3ED <- OR.bin.summ(model3.3ED)
race4ED <- OR.bin.summ(model3.4ED)
race5ED <- OR.bin.summ(model3.5ED)

RaceED <- c("Non-Hisp white", "Non-Hisp Black", "Hispanic", "Asian",
        "Other")

table3ED <- rbind(race1ED, race2ED, race3ED, race4ED, race5ED)
table.raceED <- cbind(RaceED, table3ED)
View(table.raceED)

table(cases_onlyED$race.cat)
cases_onlyED %>%
  group_by( race.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyED ) )

##----------------------------------------------------
## Other
## Race: Non-hisp white##
model3.1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphisIO, subset = race.cat == 1)

## Race: Non-Hisp Black
model3.2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphisIO, subset = race.cat == 2)

## Race: Hispanic
model3.3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphisIO, subset = race.cat == 3)

## Race: Asian
model3.4IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphisIO, subset = race.cat == 4)

## Race: Other
model3.5IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphisIO, subset = race.cat == 5)

race1IO <- OR.bin.summ(model3.1IO)
race2IO <- OR.bin.summ(model3.2IO)
race3IO <- OR.bin.summ(model3.3IO)
race4IO <- OR.bin.summ(model3.4IO)
race5IO <- OR.bin.summ(model3.5IO)
```

```
RaceIO <- c("Non-Hisp white", "Non-Hisp Black", "Hispanic", "Asian",
          "Other")

table3IO <- rbind(race1IO, race2IO, race3IO, race4IO, race5IO)
table.raceIO <- cbind(RaceIO, table3IO)
View(table.raceIO)

table(cases_onlyIO$race.cat)
cases_onlyIO %>%
  group_by( race.cat ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyIO ) )

##---------------------------------------------------
## Model 4: SES Type
##---------------------------------------------------

##---------------------------------------------------
## Overall

## Government
model4.1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = pay_group == "a. Government")

## Private
model4.2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = pay_group == "b. Private")

## Other
model4.3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = pay_group == "c. Other")

SES1O <- OR.bin.summ(model4.1O)
SES2O<- OR.bin.summ(model4.2O)
SES3O <- OR.bin.summ(model4.3O)

SES <- c("Government", "Private", "Other")
table4O <- rbind(SES1O, SES2O, SES3O)
table.SESO <- cbind(SES, table4O)
View(table.SESO)

table(cases_onlyO$pay_group)
View(cases_onlyO)
cases_onlyO %>%
  group_by( pay_group ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyO ) )
```

```
##---------------------------------------------------
## ED

## Government
model4.1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = pay_group == "a. Government")

## Private
model4.2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = pay_group == "b. Private")

## Other
model4.3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = pay_group == "c. Other")

SES1ED <- OR.bin.summ(model4.1ED)
SES2ED <- OR.bin.summ(model4.2ED)
SES3ED <- OR.bin.summ(model4.3ED)

SES <- c("Government", "Private", "Other")
table4ED <- rbind(SES1ED, SES2ED, SES3ED)
table.SESED <- cbind(SES, table4ED)
View(table.SESED)

table(cases_onlyED$pay_group)
cases_onlyED %>%
  group_by( pay_group ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyED ) )

##---------------------------------------------------
## Other

## Government
model4.1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = pay_group == "a. Government")

## Private
model4.2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = pay_group == "b. Private")

## Other
model4.3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = pay_group == "c. Other")

SES1IO <- OR.bin.summ(model4.1IO)
```

```
SES2IO<- OR.bin.summ(model4.2IO)
SES3IO <- OR.bin.summ(model4.3IO)

SES <- c("Government", "Private", "Other")
table4IO <- rbind(SES1IO, SES2IO, SES3IO)
table.SESIO <- cbind(SES, table4IO)
View(table.SESIO)

table(cases_onlyIO$pay_group)
cases_onlyIO %>%
  group_by( pay_group ) %>%
  summarise( percent = 100 * n() / nrow( cases_onlyIO ) )

##-----------------------------------------------------
## Model 5: Health Outcomes
##-----------------------------------------------------
##-----------------------------------------------------
## Overall

## Outcome: Respiratory
model5.1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = resp.o == 1)

## Outcome: Respiratory Infections (Respiratory subset)
model5.2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = resp_infections.o == 1)

## Outcome: Dermal
model5.3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = derm.o == 1)

## Outcome: Trauma
model5.4O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = trauma.o == 1)

## Outcome: Mental health
model5.5O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = mh.o == 1)

cause1O <- OR.bin.summ(model5.1O)
cause2O <- OR.bin.summ(model5.2O)
cause3O <- OR.bin.summ(model5.3O)
cause4O <- OR.bin.summ(model5.4O)
cause5O <- OR.bin.summ(model5.5O)
allO <- OR.bin.summ(model1.1O)
```

```
CauseO <- c("Respiratory", "Respiratory Infections", "Dermal", "Trauma",
          "Mental Health", "All non-traumatic")

table5O <- rbind(cause1O, cause2O, cause3O, cause4O, cause5O, allO)
table.causeO <- cbind(CauseO, table5O)
View(table.causeO)

## Getting N for data table columns
sum(cases_onlyO$resp.o %in% 1)
(sum(cases_onlyO$resp.o %in% 1)/sum(cases_onlyO$case %in% 1))*100

sum(cases_onlyO$resp_infections.o %in% 1)
(sum(cases_onlyO$resp_infections.o %in% 1)/sum(cases_onlyO$case %in% 1))*100

sum(cases_onlyO$derm.o %in% 1)
(sum(cases_onlyO$derm.o %in% 1)/sum(cases_onlyO$case %in% 1))*100

sum(cases_onlyO$trauma.o %in% 1)
(sum(cases_onlyO$trauma.o %in% 1)/sum(cases_onlyO$case %in% 1))*100

sum(cases_onlyO$mh.o %in% 1)
(sum(cases_onlyO$mh.o %in% 1)/sum(cases_onlyO$case %in% 1))*100

##-----------------------------------------------------
## ED

## Outcome: Respiratory
model5.1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = resp.o == 1)

## Outcome: Respiratory Infections (Respiratory subset)
model5.2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = resp_infections.o == 1)

## Outcome: Dermal
model5.3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = derm.o == 1)

## Outcome: Trauma
model5.4ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = trauma.o == 1)

## Outcome: Mental health
model5.5ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = mh.o == 1)
```

```
cause1ED <- OR.bin.summ(model5.1ED)
cause2ED <- OR.bin.summ(model5.2ED)
cause3ED <- OR.bin.summ(model5.3ED)
cause4ED <- OR.bin.summ(model5.4ED)
cause5ED <- OR.bin.summ(model5.5ED)
allED <- OR.bin.summ(model1.2ED)

CauseED <- c("Respiratory", "Respiratory Infections", "Dermal", "Trauma",
        "Mental Health", "All non-traumatic for ED")

table5ED <- rbind(cause1ED, cause2ED, cause3ED, cause4ED, cause5ED, allED)
table.causeED <- cbind(CauseED, table5ED)
View(table.causeED)

## Getting N for data table columns
sum(cases_onlyED$resp.o %in% 1)
(sum(cases_onlyED$resp.o %in% 1)/sum(cases_onlyED$case %in% 1))*100

sum(cases_onlyED$resp_infections.o %in% 1)
(sum(cases_onlyED$resp_infections.o %in% 1)/sum(cases_onlyED$case %in% 1))*100

sum(cases_onlyED$derm.o %in% 1)
(sum(cases_onlyED$derm.o %in% 1)/sum(cases_onlyED$case %in% 1))*100

sum(cases_onlyED$trauma.o %in% 1)
(sum(cases_onlyED$trauma.o %in% 1)/sum(cases_onlyED$case %in% 1))*100

sum(cases_onlyED$mh.o %in% 1)
(sum(cases_onlyED$mh.o %in% 1)/sum(cases_onlyED$case %in% 1))*100

##----------------------------------------------------
## Other

## Outcome: Respiratory
model5.1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
             data= waphisIO, subset = resp.o == 1)

## Outcome: Respiratory Infections (Respiratory subset)
model5.2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
             data= waphisIO, subset = resp_infections.o == 1)

## Outcome: Dermal
model5.3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
             data= waphisIO, subset = derm.o == 1)

## Outcome: Trauma
```

```
model5.4IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = trauma.o == 1)

## Outcome: Mental health
model5.5IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = mh.o == 1)

cause1IO <- OR.bin.summ(model5.1IO)
cause2IO <- OR.bin.summ(model5.2IO)
cause3IO <- OR.bin.summ(model5.3IO)
cause4IO <- OR.bin.summ(model5.4IO)
cause5IO <- OR.bin.summ(model5.5IO)
allIO <- OR.bin.summ(model1.3IO)

CauseIO <- c("Respiratory", "Respiratory Infections", "Dermal", "Trauma",
          "Mental Health", "All non-traumatic for Inpt and Obs")

table5IO <- rbind(cause1IO, cause2IO, cause3IO, cause4IO, cause5IO, allIO)
table.causeIO <- cbind(CauseIO, table5IO)
View(table.causeIO)

## Getting N for data table columns
sum(cases_onlyIO$resp.o %in% 1)
(sum(cases_onlyIO$resp.o %in% 1)/sum(cases_onlyIO$case %in% 1))*100

sum(cases_onlyIO$resp_infections.o %in% 1)
(sum(cases_onlyIO$resp_infections.o %in% 1)/sum(cases_onlyIO$case %in% 1))*100

sum(cases_onlyIO$derm.o %in% 1)
(sum(cases_onlyIO$derm.o %in% 1)/sum(cases_onlyIO$case %in% 1))*100

sum(cases_onlyIO$trauma.o %in% 1)
(sum(cases_onlyIO$trauma.o %in% 1)/sum(cases_onlyIO$case %in% 1))*100

sum(cases_onlyIO$mh.o %in% 1)
(sum(cases_onlyIO$mh.o %in% 1)/sum(cases_onlyIO$case %in% 1))*100
##------------------------------------------------------
## Group
##------------------------------------------------------

## Overall -------------------

model6.1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = sex_group == "a. Male")

model6.2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
```

```
                    data= waphism, subset = sex_group == "b. Female")

    sex1O <- OR.bin.summ(model6.1O)
    sex2O<- OR.bin.summ(model6.2O)

    Sex <- c("Male", "Female")
    tableSO <- rbind(sex1O, sex2O)
    table.sexO <- cbind(Sex, tableSO)
    View(table.sexO)

    # Getting the N
    sum(cases_onlyO$sex_group %in% "a. Male")
    ((sum(cases_onlyO$sex_group %in% "a. Male")) / (sum(cases_onlyO$case %in% 1)))*100

    sum(cases_onlyO$sex_group %in% "b. Female")
    ((sum(cases_onlyO$sex_group %in% "b. Female")) / (sum(cases_onlyO$case %in%
1)))*100

    ## ED -------------------

    model6.1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = sex_group == "a. Male")

    model6.2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = sex_group == "b. Female")

    sex1ED <- OR.bin.summ(model6.1ED)
    sex2ED<- OR.bin.summ(model6.2ED)

    Sex <- c("Male", "Female")
    tableSED <- rbind(sex1ED, sex2ED)
    table.sexED <- cbind(Sex, tableSED)
    View(table.sexED)

    # Getting the N
    sum(cases_onlyED$sex_group %in% "a. Male")
    ((sum(cases_onlyED$sex_group %in% "a. Male")) / (sum(cases_onlyED$case %in%
1)))*100

    sum(cases_onlyED$sex_group %in% "b. Female")
    ((sum(cases_onlyED$sex_group %in% "b. Female")) / (sum(cases_onlyED$case %in%
1)))*100

    ## IO -------------------

    model6.1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
```

```
                    data= waphisIO, subset = sex_group == "a. Male")

    model6.2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = sex_group == "b. Female")

    sex1IO <- OR.bin.summ(model6.1IO)
    sex2IO<- OR.bin.summ(model6.2IO)

    Sex <- c("Male", "Female")
    tableSIO <- rbind(sex1IO, sex2IO)
    table.sexIO <- cbind(Sex, tableSIO)
    View(table.sexIO)

    # Getting the N
    sum(cases_onlyIO$sex_group %in% "a. Male")
    ((sum(cases_onlyIO$sex_group %in% "a. Male")) / (sum(cases_onlyIO$case %in% 1)))*100

    sum(cases_onlyIO$sex_group %in% "b. Female")
    ((sum(cases_onlyIO$sex_group %in% "b. Female")) / (sum(cases_onlyIO$case %in%
1)))*100

    ##-----------------------------------------------------
    ## Lag models
    ##-----------------------------------------------------
    ## Lag 1
    model1.1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism)

    model1.2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED)

    model1.3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO)

    overallL1 <- OR.bin.summ(model1.1OL1)
    EDL1 <- OR.bin.summ(model1.2EDL1)
    otherL1 <- OR.bin.summ(model1.3IOL1)

    Service_L1 <- c("All cause", "ED only", "Inpatient and Observation")
    tableL1 <- rbind(overallL1, EDL1, otherL1)
    table.servL1 <- cbind(Service_L1, tableL1)
    View(table.servL1)

    ## Getting the N
    sum(cases_onlyO$case %in% 1)- sum(is.na(cases_onlyO$Exposure_1))
    (1-((sum(is.na(cases_onlyO$Exposure_1))/sum(cases_onlyO$case %in% 1))))*100
```

```
sum(cases_onlyED$case %in% 1)- sum(is.na(cases_onlyED$Exposure_1))
(1-((sum(is.na(cases_onlyED$Exposure_1))/sum(cases_onlyED$case %in% 1))))*100

sum(cases_onlyIO$case %in% 1)- sum(is.na(cases_onlyIO$Exposure_1))
(1-((sum(is.na(cases_onlyIO$Exposure_1))/sum(cases_onlyIO$case %in% 1))))*100

##---------------------------------------------------
## Lag 2

model1.1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism)

model1.2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisED)

model1.3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisIO)

overallL2 <- OR.bin.summ(model1.1OL2)
EDL2 <- OR.bin.summ(model1.2EDL2)
otherL2 <- OR.bin.summ(model1.3IOL2)

Service_L2 <- c("All cause", "ED only", "Inpatient and Observation")
tableL2 <- rbind(overallL2, EDL2, otherL2)
table.servL2 <- cbind(Service_L2, tableL2)
View(table.servL2)

## Getting the N
sum(cases_onlyO$case %in% 1)- sum(is.na(cases_onlyO$Exposure_2))
(1-((sum(is.na(cases_onlyO$Exposure_2))/sum(cases_onlyO$case %in% 1))))*100

sum(cases_onlyED$case %in% 1)- sum(is.na(cases_onlyED$Exposure_2))
(1-((sum(is.na(cases_onlyED$Exposure_2))/sum(cases_onlyED$case %in% 1))))*100

sum(cases_onlyIO$case %in% 1)- sum(is.na(cases_onlyIO$Exposure_2))
(1-((sum(is.na(cases_onlyIO$Exposure_2))/sum(cases_onlyIO$case %in% 1))))*100

##---------------------------------------------------
## Lag 3

model1.1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism)

model1.2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisED)
```

```
model1.3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisIO)

overallL3 <- OR.bin.summ(model1.1OL3)
EDL3 <- OR.bin.summ(model1.2EDL3)
otherL3 <- OR.bin.summ(model1.3IOL3)

Service_L3 <- c("All cause", "ED only", "Inpatient and Observation")
tableL3 <- rbind(overallL3, EDL3, otherL3)
table.servL3 <- cbind(Service_L3, tableL3)
View(table.servL3)

## Getting the N
sum(cases_onlyO$case %in% 1)- sum(is.na(cases_onlyO$Exposure_3))
(1-((sum(is.na(cases_onlyO$Exposure_3))/sum(cases_onlyO$case %in% 1))))*100

sum(cases_onlyED$case %in% 1)- sum(is.na(cases_onlyED$Exposure_3))
(1-((sum(is.na(cases_onlyED$Exposure_3))/sum(cases_onlyED$case %in% 1))))*100

sum(cases_onlyIO$case %in% 1)- sum(is.na(cases_onlyIO$Exposure_3))
(1-((sum(is.na(cases_onlyIO$Exposure_3))/sum(cases_onlyIO$case %in% 1))))*100

##------------------------------------------------------
## Lag 4

model1.1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism)

model1.2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisED)

model1.3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                 data = waphisIO)

overallL4 <- OR.bin.summ(model1.1OL4)
EDL4 <- OR.bin.summ(model1.2EDL4)
otherL4 <- OR.bin.summ(model1.3IOL4)

Service_L4 <- c("All cause", "ED only", "Inpatient and Observation")
tableL4 <- rbind(overallL4, EDL4, otherL4)
table.servL4 <- cbind(Service_L4, tableL4)
View(table.servL4)

## Getting the N
sum(cases_onlyO$case %in% 1)- sum(is.na(cases_onlyO$Exposure_4))
```

```
(1-((sum(is.na(cases_onlyO$Exposure_4))/sum(cases_onlyO$case %in% 1))))*100

sum(cases_onlyED$case %in% 1)- sum(is.na(cases_onlyED$Exposure_4))
(1-((sum(is.na(cases_onlyED$Exposure_4))/sum(cases_onlyED$case %in% 1))))*100

sum(cases_onlyIO$case %in% 1)- sum(is.na(cases_onlyIO$Exposure_4))
(1-((sum(is.na(cases_onlyIO$Exposure_4))/sum(cases_onlyIO$case %in% 1))))*100

##--------------------------------------------------
##--------------------------------------------------
## Model 5: Health Outcomes
##--------------------------------------------------
##--------------------------------------------------
# Model 5.R: Resp Outcomes
##--------------------------------------------------
##--------------------------------------------------
## Age groups
## Overall -------------------

model5.RA1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = resp.o == 1 & age.cat == 1)

model5.RA2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = resp.o == 1 & age.cat == 2)

model5.RA3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = resp.o == 1 & age.cat == 3)

ageR1O <- OR.bin.summ(model5.RA1O)
ageR2O<- OR.bin.summ(model5.RA2O)
ageR3O <- OR.bin.summ(model5.RA3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRAO <- rbind(ageR1O, ageR2O, ageR3O)
table.RageO <- cbind(Age, tableRAO)
View(table.RageO)

# Getting the N
sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 1)
((sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 1)) /
   (sum(cases_onlyO$resp.o %in% 1)))*100

sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 2)
((sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 2)) /
   (sum(cases_onlyO$resp.o %in% 1)))*100
```

```
sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 3)
((sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$age.cat %in% 3)) /
   (sum(cases_onlyO$resp.o %in% 1)))*100

## ED -------------------

model5.RA1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp.o == 1 & age.cat == 1)

model5.RA2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp.o == 1 & age.cat == 2)

model5.RA3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp.o == 1 & age.cat == 3)

ageR1ED <- OR.bin.summ(model5.RA1ED)
ageR2ED<- OR.bin.summ(model5.RA2ED)
ageR3ED <- OR.bin.summ(model5.RA3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRAED <- rbind(ageR1ED, ageR2ED, ageR3ED)
table.RageED <- cbind(Age, tableRAED)
View(table.RageED)

# Getting the N
sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 1)
((sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 1)) /
   (sum(cases_onlyED$resp.o %in% 1)))*100

sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 2)
((sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 2)) /
   (sum(cases_onlyED$resp.o %in% 1)))*100

sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 3)
((sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$age.cat %in% 3)) /
   (sum(cases_onlyED$resp.o %in% 1)))*100

## IO -------------------

model5.RA1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = resp.o == 1 & age.cat == 1)

model5.RA2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = resp.o == 1 & age.cat == 2)

model5.RA3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
```

```
                data= waphisIO, subset = resp.o == 1 & age.cat == 3)

ageR1IO <- OR.bin.summ(model5.RA1IO)
ageR2IO<- OR.bin.summ(model5.RA2IO)
ageR3IO <- OR.bin.summ(model5.RA3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRAIO <- rbind(ageR1IO, ageR2IO, ageR3IO)
table.RageIO <- cbind(Age, tableRAIO)
View(table.RageIO)

# Getting the N
sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 1)
((sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 1)) /
   (sum(cases_onlyIO$resp.o %in% 1)))*100

sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 2)
((sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 2)) /
   (sum(cases_onlyIO$resp.o %in% 1)))*100

sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 3)
((sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$age.cat %in% 3)) /
   (sum(cases_onlyIO$resp.o %in% 1)))*100

##----------------------------------------------------
## Sex groups

## Overall -------------------

model5.RS1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = resp.o == 1 & sex_group == "a. Male")

model5.RS2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = resp.o == 1 & sex_group == "b. Female")

sexR1O <- OR.bin.summ(model5.RS1O)
sexR2O<- OR.bin.summ(model5.RS2O)

Sex <- c("Male", "Female")
tableRSO <- rbind(sexR1O, sexR2O)
table.RsexO <- cbind(Sex, tableRSO)
View(table.RsexO)

# Getting the N
sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")
((sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")) /
```

```
   (sum(cases_onlyO$resp.o %in% 1)))*100

sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")
((sum(cases_onlyO$resp.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")) /
   (sum(cases_onlyO$resp.o %in% 1)))*100

## ED -------------------

model5.RS1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp.o == 1 & sex_group == "a. Male")

model5.RS2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp.o == 1 & sex_group == "b. Female")

sexR1ED <- OR.bin.summ(model5.RS1ED)
sexR2ED<- OR.bin.summ(model5.RS2ED)

Sex <- c("Male", "Female")
tableRSED <- rbind(sexR1ED, sexR2ED)
table.RsexED <- cbind(Sex, tableRSED)
View(table.RsexED)

# Getting the N
sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")
((sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")) /
   (sum(cases_onlyED$resp.o %in% 1)))*100

sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")
((sum(cases_onlyED$resp.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")) /
   (sum(cases_onlyED$resp.o %in% 1)))*100

## IO -------------------

model5.RS1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = resp.o == 1 & sex_group == "a. Male")

model5.RS2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = resp.o == 1 & sex_group == "b. Female")

sexR1IO <- OR.bin.summ(model5.RS1IO)
sexR2IO<- OR.bin.summ(model5.RS2IO)

Sex <- c("Male", "Female")
tableRSIO <- rbind(sexR1IO, sexR2IO)
table.RsexIO <- cbind(Sex, tableRSIO)
View(table.RsexIO)
```

```
# Getting the N
sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")
((sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")) /
   (sum(cases_onlyIO$resp.o %in% 1)))*100

sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")
((sum(cases_onlyIO$resp.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")) /
   (sum(cases_onlyIO$resp.o %in% 1)))*100

##-------------------------------------------------------
## Lag groups

## Lag 1
model5.R1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = resp.o == 1)

model5.R2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                  data = waphisED, subset = resp.o == 1)

model5.R3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                  data = waphisIO, subset = resp.o == 1)

overallRL1 <- OR.bin.summ(model5.R1OL1)
EDRL1 <- OR.bin.summ(model5.R2EDL1)
otherRL1 <- OR.bin.summ(model5.R3IOL1)

Service_RL1 <- c("All cause", "ED only", "Inpatient and Observation")
tableRL1 <- rbind(overallRL1, EDRL1, otherRL1)
table.servRL1 <- cbind(Service_RL1, tableRL1)
View(table.servRL1)

## Getting the N
sum(cases_onlyO$resp.o %in% 1)- sum(cases_onlyO$resp.o %in% 1 &
is.na(cases_onlyO$Exposure_1))
(1-((sum(cases_onlyO$resp.o %in% 1 & is.na(cases_onlyO$Exposure_1))/
      sum(cases_onlyO$resp.o %in% 1))))*100

sum(cases_onlyED$resp.o %in% 1)- sum(cases_onlyED$resp.o %in% 1 &
is.na(cases_onlyED$Exposure_1))
(1-((sum(cases_onlyED$resp.o %in% 1 & is.na(cases_onlyED$Exposure_1))/
      sum(cases_onlyED$resp.o %in% 1))))*100

sum(cases_onlyIO$resp.o %in% 1)- sum(cases_onlyIO$resp.o %in% 1 &
is.na(cases_onlyIO$Exposure_1))
(1-((sum(cases_onlyIO$resp.o %in% 1 & is.na(cases_onlyIO$Exposure_1))/
```

```
        sum(cases_onlyIO$resp.o %in% 1))))*100

    ## Lag 2 --------------------

    model5.R1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = resp.o == 1)

    model5.R2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = resp.o == 1)

    model5.R3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = resp.o == 1)

    overallRL2 <- OR.bin.summ(model5.R1OL2)
    EDRL2 <- OR.bin.summ(model5.R2EDL2)
    otherRL2 <- OR.bin.summ(model5.R3IOL2)

    Service_RL2 <- c("All cause", "ED only", "Inpatient and Observation")
    tableRL2 <- rbind(overallRL2, EDRL2, otherRL2)
    table.servRL2 <- cbind(Service_RL2, tableRL2)
    View(table.servRL2)

    ## Getting the N
    sum(cases_onlyO$resp.o %in% 1)- sum(cases_onlyO$resp.o %in% 1 &
is.na(cases_onlyO$Exposure_2))
    (1-((sum(cases_onlyO$resp.o %in% 1 & is.na(cases_onlyO$Exposure_2))/
        sum(cases_onlyO$resp.o %in% 1))))*100

    sum(cases_onlyED$resp.o %in% 1)- sum(cases_onlyED$resp.o %in% 1 &
is.na(cases_onlyED$Exposure_2))
    (1-((sum(cases_onlyED$resp.o %in% 1 & is.na(cases_onlyED$Exposure_2))/
        sum(cases_onlyED$resp.o %in% 1))))*100

    sum(cases_onlyIO$resp.o %in% 1)- sum(cases_onlyIO$resp.o %in% 1 &
is.na(cases_onlyIO$Exposure_2))
    (1-((sum(cases_onlyIO$resp.o %in% 1 & is.na(cases_onlyIO$Exposure_2))/
        sum(cases_onlyIO$resp.o %in% 1))))*100

    ## Lag 3 --------------------
    model5.R1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = resp.o == 1)

    model5.R2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = resp.o == 1)

    model5.R3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
```

```
                    data = waphisIO, subset = resp.o == 1)

    overallRL3 <- OR.bin.summ(model5.R1OL3)
    EDRL3 <- OR.bin.summ(model5.R2EDL3)
    otherRL3 <- OR.bin.summ(model5.R3IOL3)

    Service_RL3 <- c("All cause", "ED only", "Inpatient and Observation")
    tableRL3 <- rbind(overallRL3, EDRL3, otherRL3)
    table.servRL3 <- cbind(Service_RL3, tableRL3)
    View(table.servRL3)

    ## Getting the N
    sum(cases_onlyO$resp.o %in% 1)- sum(cases_onlyO$resp.o %in% 1 &
is.na(cases_onlyO$Exposure_3))
    (1-((sum(cases_onlyO$resp.o %in% 1 & is.na(cases_onlyO$Exposure_3))/
        sum(cases_onlyO$resp.o %in% 1))))*100

    sum(cases_onlyED$resp.o %in% 1)- sum(cases_onlyED$resp.o %in% 1 &
is.na(cases_onlyED$Exposure_3))
    (1-((sum(cases_onlyED$resp.o %in% 1 & is.na(cases_onlyED$Exposure_3))/
        sum(cases_onlyED$resp.o %in% 1))))*100

    sum(cases_onlyIO$resp.o %in% 1)- sum(cases_onlyIO$resp.o %in% 1 &
is.na(cases_onlyIO$Exposure_3))
    (1-((sum(cases_onlyIO$resp.o %in% 1 & is.na(cases_onlyIO$Exposure_3))/
        sum(cases_onlyIO$resp.o %in% 1))))*100

    ## Lag 4 --------------------
    model5.R1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = resp.o == 1)

    model5.R2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = resp.o == 1)

    model5.R3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = resp.o == 1)

    overallRL4 <- OR.bin.summ(model5.R1OL4)
    EDRL4 <- OR.bin.summ(model5.R2EDL4)
    otherRL4 <- OR.bin.summ(model5.R3IOL4)

    Service_RL4 <- c("All cause", "ED only", "Inpatient and Observation")
    tableRL4 <- rbind(overallRL4, EDRL4, otherRL4)
    table.servRL4 <- cbind(Service_RL4, tableRL4)
    View(table.servRL4)
```

```
## Getting the N
sum(cases_onlyO$resp.o %in% 1)- sum(cases_onlyO$resp.o %in% 1 &
is.na(cases_onlyO$Exposure_4))
(1-((sum(cases_onlyO$resp.o %in% 1 & is.na(cases_onlyO$Exposure_4))/
    sum(cases_onlyO$resp.o %in% 1))))*100

sum(cases_onlyED$resp.o %in% 1)- sum(cases_onlyED$resp.o %in% 1 &
is.na(cases_onlyED$Exposure_4))
(1-((sum(cases_onlyED$resp.o %in% 1 & is.na(cases_onlyED$Exposure_4))/
    sum(cases_onlyED$resp.o %in% 1))))*100

sum(cases_onlyIO$resp.o %in% 1)- sum(cases_onlyIO$resp.o %in% 1 &
is.na(cases_onlyIO$Exposure_4))
(1-((sum(cases_onlyIO$resp.o %in% 1 & is.na(cases_onlyIO$Exposure_4))/
    sum(cases_onlyIO$resp.o %in% 1))))*100

##--------------------------------------------------
# Model 5.RI: Resp Infections Outcomes
##--------------------------------------------------
##--------------------------------------------------
## Age groups

## Overall -------------------

model5.RIA1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = resp_infections.o == 1 & age.cat == 1)

model5.RIA2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = resp_infections.o == 1 & age.cat == 2)

model5.RIA3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = resp_infections.o == 1 & age.cat == 3)

ageRI1O <- OR.bin.summ(model5.RIA1O)
ageRI2O<- OR.bin.summ(model5.RIA2O)
ageRI3O <- OR.bin.summ(model5.RIA3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRIAO <- rbind(ageRI1O, ageRI2O, ageRI3O)
table.RIageO <- cbind(Age, tableRIAO)
View(table.RIageO)

# Getting the N
sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 1)
((sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 1)) /
  (sum(cases_onlyO$resp_infections.o %in% 1)))*100
```

```
sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 2)
((sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 2)) /
   (sum(cases_onlyO$resp_infections.o %in% 1)))*100

sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 3)
((sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$age.cat %in% 3)) /
   (sum(cases_onlyO$resp_infections.o %in% 1)))*100

## ED -------------------

model5.RIA1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp_infections.o == 1 & age.cat == 1)

model5.RIA2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp_infections.o == 1 & age.cat == 2)

model5.RIA3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = resp_infections.o == 1 & age.cat == 3)

ageRI1ED <- OR.bin.summ(model5.RIA1ED)
ageRI2ED<- OR.bin.summ(model5.RIA2ED)
ageRI3ED <- OR.bin.summ(model5.RIA3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRIAED <- rbind(ageRI1ED, ageRI2ED, ageRI3ED)
table.RIageED <- cbind(Age, tableRIAED)
View(table.RIageED)

# Getting the N
sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 1)
((sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 1)) /
   (sum(cases_onlyED$resp_infections.o %in% 1)))*100

sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 2)
((sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 2)) /
   (sum(cases_onlyED$resp_infections.o %in% 1)))*100

sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 3)
((sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$age.cat %in% 3)) /
   (sum(cases_onlyED$resp_infections.o %in% 1)))*100

## IO -------------------

model5.RIA1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = resp_infections.o == 1 & age.cat == 1)
```

```
model5.RIA2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = resp_infections.o == 1 & age.cat == 2)

model5.RIA3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = resp_infections.o == 1 & age.cat == 3)

ageRI1IO <- OR.bin.summ(model5.RIA1IO)
ageRI2IO<- OR.bin.summ(model5.RIA2IO)
ageRI3IO <- OR.bin.summ(model5.RIA3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableRIAIO <- rbind(ageRI1IO, ageRI2IO, ageRI3IO)
table.RIageIO <- cbind(Age, tableRIAIO)
View(table.RIageIO)

# Getting the N
sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 1)
((sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 1)) /
   (sum(cases_onlyIO$resp_infections.o %in% 1)))*100

sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 2)
((sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 2)) /
   (sum(cases_onlyIO$resp_infections.o %in% 1)))*100

sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 3)
((sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$age.cat %in% 3)) /
   (sum(cases_onlyIO$resp_infections.o %in% 1)))*100

##-----------------------------------------------------
## Sex groups

## Overall -------------------

model5.RIS1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = resp_infections.o == 1 & sex_group == "a. Male")

model5.RIS2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = resp_infections.o == 1 & sex_group == "b. Female")

sexRI1O <- OR.bin.summ(model5.RIS1O)
sexRI2O<- OR.bin.summ(model5.RIS2O)

Sex <- c("Male", "Female")
tableRISO <- rbind(sexRI1O, sexRI2O)
table.RIsexO <- cbind(Sex, tableRISO)
```

```
View(table.RIsexO)

# Getting the N
sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")
((sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$sex_group %in% "a. Male"))
/
   (sum(cases_onlyO$resp_infections.o %in% 1)))*100

sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")
((sum(cases_onlyO$resp_infections.o %in% 1 & cases_onlyO$sex_group %in% "b.
Female")) /
   (sum(cases_onlyO$resp_infections.o %in% 1)))*100

## ED -------------------

model5.RIS1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = resp_infections.o == 1 & sex_group == "a. Male")

model5.RIS2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = resp_infections.o == 1 & sex_group == "b.
Female")

sexRI1ED <- OR.bin.summ(model5.RIS1ED)
sexRI2ED<- OR.bin.summ(model5.RIS2ED)

Sex <- c("Male", "Female")
tableRISED <- rbind(sexRI1ED, sexRI2ED)
table.RIsexED <- cbind(Sex, tableRISED)
View(table.RIsexED)

# Getting the N
sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")
((sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$sex_group %in% "a.
Male")) /
   (sum(cases_onlyED$resp_infections.o %in% 1)))*100

sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$sex_group %in% "b.
Female")
((sum(cases_onlyED$resp_infections.o %in% 1 & cases_onlyED$sex_group %in% "b.
Female")) /
   (sum(cases_onlyED$resp_infections.o %in% 1)))*100

## IO -------------------

model5.RIS1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = resp_infections.o == 1 & sex_group == "a. Male")
```

```
    model5.RIS2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                      data= waphisIO, subset = resp_infections.o == 1 & sex_group == "b.
Female")

    sexRI1IO <- OR.bin.summ(model5.RIS1IO)
    sexRI2IO<- OR.bin.summ(model5.RIS2IO)

    Sex <- c("Male", "Female")
    tableRISIO <- rbind(sexRI1IO, sexRI2IO)
    table.RIsexIO <- cbind(Sex, tableRISIO)
    View(table.RIsexIO)

    # Getting the N
    sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")
    ((sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$sex_group %in% "a.
Male")) /
       (sum(cases_onlyIO$resp_infections.o %in% 1)))*100

    sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$sex_group %in% "b.
Female")
    ((sum(cases_onlyIO$resp_infections.o %in% 1 & cases_onlyIO$sex_group %in% "b.
Female")) /
       (sum(cases_onlyIO$resp_infections.o %in% 1)))*100

    ##-----------------------------------------------------
    ## Lag groups

    ## Lag 1
    model5.RI1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                      data= waphism, subset = resp_infections.o == 1)

    model5.RI2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                      data = waphisED, subset = resp_infections.o == 1)

    model5.RI3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                      data = waphisIO, subset = resp_infections.o == 1)

    overallRIL1 <- OR.bin.summ(model5.RI1OL1)
    EDRIL1 <- OR.bin.summ(model5.RI2EDL1)
    otherRIL1 <- OR.bin.summ(model5.RI3IOL1)

    Service_RIL1 <- c("All cause", "ED only", "Inpatient and Observation")
    tableRIL1 <- rbind(overallRIL1, EDRIL1, otherRIL1)
    table.servRIL1 <- cbind(Service_RIL1, tableRIL1)
    View(table.servRIL1)
```

```
## Getting the N
sum(cases_onlyO$resp_infections.o %in% 1)- sum(cases_onlyO$resp_infections.o %in% 1
& is.na(cases_onlyO$Exposure_1))
(1-((sum(cases_onlyO$resp_infections.o %in% 1 & is.na(cases_onlyO$Exposure_1))/
    sum(cases_onlyO$resp_infections.o %in% 1))))*100

sum(cases_onlyED$resp_infections.o %in% 1)- sum(cases_onlyED$resp_infections.o %in%
1 & is.na(cases_onlyED$Exposure_1))
(1-((sum(cases_onlyED$resp_infections.o %in% 1 & is.na(cases_onlyED$Exposure_1))/
    sum(cases_onlyED$resp_infections.o %in% 1))))*100

sum(cases_onlyIO$resp_infections.o %in% 1)- sum(cases_onlyIO$resp_infections.o %in% 1
& is.na(cases_onlyIO$Exposure_1))
(1-((sum(cases_onlyIO$resp_infections.o %in% 1 & is.na(cases_onlyIO$Exposure_1))/
    sum(cases_onlyIO$resp_infections.o %in% 1))))*100

## Lag 2 --------------------

model5.RI1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = resp_infections.o == 1)

model5.RI2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                  data = waphisED, subset = resp_infections.o == 1)

model5.RI3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                  data = waphisIO, subset = resp_infections.o == 1)

overallRIL2 <- OR.bin.summ(model5.RI1OL2)
EDRIL2 <- OR.bin.summ(model5.RI2EDL2)
otherRIL2 <- OR.bin.summ(model5.RI3IOL2)

Service_RIL2 <- c("All cause", "ED only", "Inpatient and Observation")
tableRIL2 <- rbind(overallRIL2, EDRIL2, otherRIL2)
table.servRIL2 <- cbind(Service_RIL2, tableRIL2)
View(table.servRIL2)

## Getting the N
sum(cases_onlyO$resp_infections.o %in% 1)- sum(cases_onlyO$resp_infections.o %in% 1
& is.na(cases_onlyO$Exposure_2))
(1-((sum(cases_onlyO$resp_infections.o %in% 1 & is.na(cases_onlyO$Exposure_2))/
    sum(cases_onlyO$resp_infections.o %in% 1))))*100

sum(cases_onlyED$resp_infections.o %in% 1)- sum(cases_onlyED$resp_infections.o %in%
1 & is.na(cases_onlyED$Exposure_2))
(1-((sum(cases_onlyED$resp_infections.o %in% 1 & is.na(cases_onlyED$Exposure_2))/
```

```
        sum(cases_onlyED$resp_infections.o %in% 1))))*100

    sum(cases_onlyIO$resp_infections.o %in% 1)- sum(cases_onlyIO$resp_infections.o %in% 1
& is.na(cases_onlyIO$Exposure_2))
    (1-((sum(cases_onlyIO$resp_infections.o %in% 1 & is.na(cases_onlyIO$Exposure_2))/
        sum(cases_onlyIO$resp_infections.o %in% 1))))*100

    ## Lag 3 --------------------
    model5.RI1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = resp_infections.o == 1)

    model5.RI2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = resp_infections.o == 1)

    model5.RI3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = resp_infections.o == 1)

    overallRIL3 <- OR.bin.summ(model5.RI1OL3)
    EDRIL3 <- OR.bin.summ(model5.RI2EDL3)
    otherRIL3 <- OR.bin.summ(model5.RI3IOL3)

    Service_RIL3 <- c("All cause", "ED only", "Inpatient and Observation")
    tableRIL3 <- rbind(overallRIL3, EDRIL3, otherRIL3)
    table.servRIL3 <- cbind(Service_RIL3, tableRIL3)
    View(table.servRIL3)

    ## Getting the N
    sum(cases_onlyO$resp_infections.o %in% 1)- sum(cases_onlyO$resp_infections.o %in% 1
& is.na(cases_onlyO$Exposure_3))
    (1-((sum(cases_onlyO$resp_infections.o %in% 1 & is.na(cases_onlyO$Exposure_3))/
        sum(cases_onlyO$resp_infections.o %in% 1))))*100

    sum(cases_onlyED$resp_infections.o %in% 1)- sum(cases_onlyED$resp_infections.o %in%
1 & is.na(cases_onlyED$Exposure_3))
    (1-((sum(cases_onlyED$resp_infections.o %in% 1 & is.na(cases_onlyED$Exposure_3))/
        sum(cases_onlyED$resp_infections.o %in% 1))))*100

    sum(cases_onlyIO$resp_infections.o %in% 1)- sum(cases_onlyIO$resp_infections.o %in% 1
& is.na(cases_onlyIO$Exposure_3))
    (1-((sum(cases_onlyIO$resp_infections.o %in% 1 & is.na(cases_onlyIO$Exposure_3))/
        sum(cases_onlyIO$resp_infections.o %in% 1))))*100

    ## Lag 4 --------------------
    model5.RI1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = resp_infections.o == 1)
```

```
model5.RI2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisED, subset = resp_infections.o == 1)

model5.RI3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisIO, subset = resp_infections.o == 1)

overallRIL4 <- OR.bin.summ(model5.RI1OL4)
EDRIL4 <- OR.bin.summ(model5.RI2EDL4)
otherRIL4 <- OR.bin.summ(model5.RI3IOL4)

Service_RIL4 <- c("All cause", "ED only", "Inpatient and Observation")
tableRIL4 <- rbind(overallRIL4, EDRIL4, otherRIL4)
table.servRIL4 <- cbind(Service_RIL4, tableRIL4)
View(table.servRIL4)

## Getting the N
sum(cases_onlyO$resp_infections.o %in% 1)- sum(cases_onlyO$resp_infections.o %in% 1
& is.na(cases_onlyO$Exposure_4))
(1-((sum(cases_onlyO$resp_infections.o %in% 1 & is.na(cases_onlyO$Exposure_4))/
    sum(cases_onlyO$resp_infections.o %in% 1))))*100

sum(cases_onlyED$resp_infections.o %in% 1)- sum(cases_onlyED$resp_infections.o %in%
1 & is.na(cases_onlyED$Exposure_4))
(1-((sum(cases_onlyED$resp_infections.o %in% 1 & is.na(cases_onlyED$Exposure_4))/
    sum(cases_onlyED$resp_infections.o %in% 1))))*1000

sum(cases_onlyIO$resp_infections.o %in% 1)- sum(cases_onlyIO$resp_infections.o %in% 1
& is.na(cases_onlyIO$Exposure_4))
(1-((sum(cases_onlyIO$resp_infections.o %in% 1 & is.na(cases_onlyIO$Exposure_4))/
    sum(cases_onlyIO$resp_infections.o %in% 1))))*100

##---------------------------------------------------
# Model 5.D: Dermal Outcomes
##---------------------------------------------------
##---------------------------------------------------
## Age groups

## Overall -------------------

model5.DA1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphism, subset = derm.o == 1 & age.cat == 1)

model5.DA2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
              data= waphism, subset = derm.o == 1 & age.cat == 2)

model5.DA3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
```

```
                    data= waphism, subset = derm.o == 1 & age.cat == 3)

ageD1O <- OR.bin.summ(model5.DA1O)
ageD2O<- OR.bin.summ(model5.DA2O)
ageD3O <- OR.bin.summ(model5.DA3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableDAO <- rbind(ageD1O, ageD2O, ageD3O)
table.DageO <- cbind(Age, tableDAO)
View(table.DageO)

# Getting the N
sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 1)
((sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 1)) /
   (sum(cases_onlyO$derm.o %in% 1)))*100

sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 2)
((sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 2)) /
   (sum(cases_onlyO$derm.o %in% 1)))*100

sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 3)
((sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$age.cat %in% 3)) /
   (sum(cases_onlyO$derm.o %in% 1)))*100

## ED -------------------

model5.DA1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = derm.o == 1 & age.cat == 1)

model5.DA2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = derm.o == 1 & age.cat == 2)

model5.DA3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = derm.o == 1 & age.cat == 3)

ageD1ED <- OR.bin.summ(model5.DA1ED)
ageD2ED<- OR.bin.summ(model5.DA2ED)
ageD3ED <- OR.bin.summ(model5.DA3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableDAED <- rbind(ageD1ED, ageD2ED, ageD3ED)
table.DageED <- cbind(Age, tableDAED)
View(table.DageED)

# Getting the N
sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 1)
```

```
((sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 1)) /
   (sum(cases_onlyED$derm.o %in% 1)))*100

sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 2)
((sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 2)) /
   (sum(cases_onlyED$derm.o %in% 1)))*100

sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 3)
((sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$age.cat %in% 3)) /
   (sum(cases_onlyED$derm.o %in% 1)))*100

## IO -------------------

model5.DA1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = derm.o == 1 & age.cat == 1)

model5.DA2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = derm.o == 1 & age.cat == 2)

model5.DA3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = derm.o == 1 & age.cat == 3)

ageD1IO <- OR.bin.summ(model5.DA1IO)
ageD2IO<- OR.bin.summ(model5.DA2IO)
ageD3IO <- OR.bin.summ(model5.DA3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableDAIO <- rbind(ageD1IO, ageD2IO, ageD3IO)
table.DageIO <- cbind(Age, tableDAIO)
View(table.DageIO)

# Getting the N
sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 1)
((sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 1)) /
   (sum(cases_onlyIO$derm.o %in% 1)))*100

sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 2)
((sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 2)) /
   (sum(cases_onlyIO$derm.o %in% 1)))*100

sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 3)
((sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$age.cat %in% 3)) /
   (sum(cases_onlyIO$derm.o %in% 1)))*100

##----------------------------------------------------
## Sex groups
```

```
## Overall -------------------

model5.DS1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = derm.o == 1 & sex_group == "a. Male")

model5.DS2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphism, subset = derm.o == 1 & sex_group == "b. Female")

sexD1O <- OR.bin.summ(model5.DS1O)
sexD2O<- OR.bin.summ(model5.DS2O)

Sex <- c("Male", "Female")
tableDSO <- rbind(sexD1O, sexD2O)
table.DsexO <- cbind(Sex, tableDSO)
View(table.DsexO)

# Getting the N
sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")
((sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")) /
   (sum(cases_onlyO$derm.o %in% 1)))*100

sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")
((sum(cases_onlyO$derm.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")) /
   (sum(cases_onlyO$derm.o %in% 1)))*100

## ED -------------------

model5.DS1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = derm.o == 1 & sex_group == "a. Male")

model5.DS2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisED, subset = derm.o == 1 & sex_group == "b. Female")

sexD1ED <- OR.bin.summ(model5.DS1ED)
sexD2ED<- OR.bin.summ(model5.DS2ED)

Sex <- c("Male", "Female")
tableDSED <- rbind(sexD1ED, sexD2ED)
table.DsexED <- cbind(Sex, tableDSED)
View(table.DsexED)

# Getting the N
sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")
((sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")) /
   (sum(cases_onlyED$derm.o %in% 1)))*100
```

```
sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")
((sum(cases_onlyED$derm.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")) /
   (sum(cases_onlyED$derm.o %in% 1)))*100

## IO -------------------

model5.DS1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = derm.o == 1 & sex_group == "a. Male")

model5.DS2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = derm.o == 1 & sex_group == "b. Female")

sexD1IO <- OR.bin.summ(model5.DS1IO)
sexD2IO<- OR.bin.summ(model5.DS2IO)

Sex <- c("Male", "Female")
tableDSIO <- rbind(sexD1IO, sexD2IO)
table.DsexIO <- cbind(Sex, tableDSIO)
View(table.DsexIO)

# Getting the N
sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")
((sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")) /
   (sum(cases_onlyIO$derm.o %in% 1)))*100

sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")
((sum(cases_onlyIO$derm.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")) /
   (sum(cases_onlyIO$derm.o %in% 1)))*100

##----------------------------------------------------
## Lag groups

## Lag 1
model5.D1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                   data= waphism, subset = derm.o == 1)

model5.D2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                   data = waphisED, subset = derm.o == 1)

model5.D3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                   data = waphisIO, subset = derm.o == 1)

overallDL1 <- OR.bin.summ(model5.D1OL1)
EDDL1 <- OR.bin.summ(model5.D2EDL1)
otherDL1 <- OR.bin.summ(model5.D3IOL1)
```

```
Service_DL1 <- c("All cause", "ED only", "Inpatient and Observation")
tableDL1 <- rbind(overallDL1, EDDL1, otherDL1)
table.servDL1 <- cbind(Service_DL1, tableDL1)
View(table.servDL1)

## Getting the N
sum(cases_onlyO$derm.o %in% 1)- sum(cases_onlyO$derm.o %in% 1 &
is.na(cases_onlyO$Exposure_1))
(1-((sum(cases_onlyO$derm.o %in% 1 & is.na(cases_onlyO$Exposure_1))/
    sum(cases_onlyO$derm.o %in% 1))))*100

sum(cases_onlyED$derm.o %in% 1)- sum(cases_onlyED$derm.o %in% 1 &
is.na(cases_onlyED$Exposure_1))
(1-((sum(cases_onlyED$derm.o %in% 1 & is.na(cases_onlyED$Exposure_1))/
    sum(cases_onlyED$derm.o %in% 1))))*100

sum(cases_onlyIO$derm.o %in% 1)- sum(cases_onlyIO$derm.o %in% 1 &
is.na(cases_onlyIO$Exposure_1))
(1-((sum(cases_onlyIO$derm.o %in% 1 & is.na(cases_onlyIO$Exposure_1))/
    sum(cases_onlyIO$derm.o %in% 1))))*100

## Lag 2 --------------------

model5.D1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
              data= waphism, subset = derm.o == 1)

model5.D2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
              data = waphisED, subset = derm.o == 1)

model5.D3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
              data = waphisIO, subset = derm.o == 1)

overallDL2 <- OR.bin.summ(model5.D1OL2)
EDDL2 <- OR.bin.summ(model5.D2EDL2)
otherDL2 <- OR.bin.summ(model5.D3IOL2)

Service_DL2 <- c("All cause", "ED only", "Inpatient and Observation")
tableDL2 <- rbind(overallDL2, EDDL2, otherDL2)
table.servDL2 <- cbind(Service_DL2, tableDL2)
View(table.servDL2)

## Getting the N
sum(cases_onlyO$derm.o %in% 1)- sum(cases_onlyO$derm.o %in% 1 &
is.na(cases_onlyO$Exposure_2))
(1-((sum(cases_onlyO$derm.o %in% 1 & is.na(cases_onlyO$Exposure_2))/
```

```
        sum(cases_onlyO$derm.o %in% 1))))*100

    sum(cases_onlyED$derm.o %in% 1)- sum(cases_onlyED$derm.o %in% 1 &
is.na(cases_onlyED$Exposure_2))
    (1-((sum(cases_onlyED$derm.o %in% 1 & is.na(cases_onlyED$Exposure_2))/
        sum(cases_onlyED$derm.o %in% 1))))*100

    sum(cases_onlyIO$derm.o %in% 1)- sum(cases_onlyIO$derm.o %in% 1 &
is.na(cases_onlyIO$Exposure_2))
    (1-((sum(cases_onlyIO$derm.o %in% 1 & is.na(cases_onlyIO$Exposure_2))/
        sum(cases_onlyIO$derm.o %in% 1))))*100

    ## Lag 3 --------------------
    model5.D1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = derm.o == 1)

    model5.D2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = derm.o == 1)

    model5.D3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = derm.o == 1)

    overallDL3 <- OR.bin.summ(model5.D1OL3)
    EDDL3 <- OR.bin.summ(model5.D2EDL3)
    otherDL3 <- OR.bin.summ(model5.D3IOL3)

    Service_DL3 <- c("All cause", "ED only", "Inpatient and Observation")
    tableDL3 <- rbind(overallDL3, EDDL3, otherDL3)
    table.servDL3 <- cbind(Service_DL3, tableDL3)
    View(table.servDL3)

    ## Getting the N
    sum(cases_onlyO$derm.o %in% 1)- sum(cases_onlyO$derm.o %in% 1 &
is.na(cases_onlyO$Exposure_3))
    (1-((sum(cases_onlyO$derm.o %in% 1 & is.na(cases_onlyO$Exposure_3))/
        sum(cases_onlyO$derm.o %in% 1))))*100

    sum(cases_onlyED$derm.o %in% 1)- sum(cases_onlyED$derm.o %in% 1 &
is.na(cases_onlyED$Exposure_3))
    (1-((sum(cases_onlyED$derm.o %in% 1 & is.na(cases_onlyED$Exposure_3))/
        sum(cases_onlyED$derm.o %in% 1))))*100

    sum(cases_onlyIO$derm.o %in% 1)- sum(cases_onlyIO$derm.o %in% 1 &
is.na(cases_onlyIO$Exposure_3))
    (1-((sum(cases_onlyIO$derm.o %in% 1 & is.na(cases_onlyIO$Exposure_3))/
        sum(cases_onlyIO$derm.o %in% 1))))*100
```

```
## Lag 4 --------------------
model5.D1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = derm.o == 1)

model5.D2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisED, subset = derm.o == 1)

model5.D3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisIO, subset = derm.o == 1)

overallDL4 <- OR.bin.summ(model5.D1OL4)
EDDL4 <- OR.bin.summ(model5.D2EDL4)
otherDL4 <- OR.bin.summ(model5.D3IOL4)

Service_DL4 <- c("All cause", "ED only", "Inpatient and Observation")
tableDL4 <- rbind(overallDL4, EDDL4, otherDL4)
table.servDL4 <- cbind(Service_DL4, tableDL4)
View(table.servDL4)

## Getting the N
sum(cases_onlyO$derm.o %in% 1)- sum(cases_onlyO$derm.o %in% 1 &
is.na(cases_onlyO$Exposure_4))
(1-((sum(cases_onlyO$derm.o %in% 1 & is.na(cases_onlyO$Exposure_4))/
    sum(cases_onlyO$derm.o %in% 1))))*100

sum(cases_onlyED$derm.o %in% 1)- sum(cases_onlyED$derm.o %in% 1 &
is.na(cases_onlyED$Exposure_4))
(1-((sum(cases_onlyED$derm.o %in% 1 & is.na(cases_onlyED$Exposure_4))/
    sum(cases_onlyED$derm.o %in% 1))))*100

sum(cases_onlyIO$derm.o %in% 1)- sum(cases_onlyIO$derm.o %in% 1 &
is.na(cases_onlyIO$Exposure_4))
(1-((sum(cases_onlyIO$derm.o %in% 1 & is.na(cases_onlyIO$Exposure_4))/
    sum(cases_onlyIO$derm.o %in% 1))))*100

##----------------------------------------------------
# Model 5.T: Trauma Outcomes
##----------------------------------------------------
##----------------------------------------------------
## Age groups

## Overall -------------------

model5.TA1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
               data= waphism, subset = trauma.o == 1 & age.cat == 1)
```

```
model5.TA2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = trauma.o == 1 & age.cat == 2)

model5.TA3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = trauma.o == 1 & age.cat == 3)

ageT1O <- OR.bin.summ(model5.TA1O)
ageT2O<- OR.bin.summ(model5.TA2O)
ageT3O <- OR.bin.summ(model5.TA3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableTAO <- rbind(ageT1O, ageT2O, ageT3O)
table.TageO <- cbind(Age, tableTAO)
View(table.TageO)

# Getting the N
sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 1)
((sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 1)) /
   (sum(cases_onlyO$trauma.o %in% 1)))*100

sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 2)
((sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 2)) /
   (sum(cases_onlyO$trauma.o %in% 1)))*100

sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 3)
((sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$age.cat %in% 3)) /
   (sum(cases_onlyO$trauma.o %in% 1)))*100

## ED -------------------

model5.TA1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = trauma.o == 1 & age.cat == 1)

model5.TA2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = trauma.o == 1 & age.cat == 2)

model5.TA3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                data= waphisED, subset = trauma.o == 1 & age.cat == 3)

ageT1ED <- OR.bin.summ(model5.TA1ED)
ageT2ED<- OR.bin.summ(model5.TA2ED)
ageT3ED <- OR.bin.summ(model5.TA3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableTAED <- rbind(ageT1ED, ageT2ED, ageT3ED)
```

```
table.TageED <- cbind(Age, tableTAED)
View(table.TageED)

# Getting the N
sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 1)
((sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 1)) /
   (sum(cases_onlyED$trauma.o %in% 1)))*100

sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 2)
((sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 2)) /
   (sum(cases_onlyED$trauma.o %in% 1)))*100

sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 3)
((sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$age.cat %in% 3)) /
   (sum(cases_onlyED$trauma.o %in% 1)))*100

## IO -------------------

model5.TA1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = trauma.o == 1 & age.cat == 1)

model5.TA2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = trauma.o == 1 & age.cat == 2)

model5.TA3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisIO, subset = trauma.o == 1 & age.cat == 3)

ageT1IO <- OR.bin.summ(model5.TA1IO)
ageT2IO<- OR.bin.summ(model5.TA2IO)
ageT3IO <- OR.bin.summ(model5.TA3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableTAIO <- rbind(ageT1IO, ageT2IO, ageT3IO)
table.TageIO <- cbind(Age, tableTAIO)
View(table.TageIO)

# Getting the N
sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 1)
((sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 1)) /
   (sum(cases_onlyIO$trauma.o %in% 1)))*100

sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 2)
((sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 2)) /
   (sum(cases_onlyIO$trauma.o %in% 1)))*100

sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 3)
```

```
((sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$age.cat %in% 3)) /
   (sum(cases_onlyIO$trauma.o %in% 1)))*100

##-----------------------------------------------------
## Sex groups

## Overall -------------------

model5.TS1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = trauma.o == 1 & sex_group == "a. Male")

model5.TS2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = trauma.o == 1 & sex_group == "b. Female")

sexT1O <- OR.bin.summ(model5.TS1O)
sexT2O<- OR.bin.summ(model5.TS2O)

Sex <- c("Male", "Female")
tableTSO <- rbind(sexT1O, sexT2O)
table.TsexO <- cbind(Sex, tableTSO)
View(table.TsexO)

# Getting the N
sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")
((sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")) /
   (sum(cases_onlyO$trauma.o %in% 1)))*100

sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")
((sum(cases_onlyO$trauma.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")) /
   (sum(cases_onlyO$trauma.o %in% 1)))*100

## ED -------------------

model5.TS1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = trauma.o == 1 & sex_group == "a. Male")

model5.TS2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisED, subset = trauma.o == 1 & sex_group == "b. Female")

sexT1ED <- OR.bin.summ(model5.TS1ED)
sexT2ED<- OR.bin.summ(model5.TS2ED)

Sex <- c("Male", "Female")
tableTSED <- rbind(sexT1ED, sexT2ED)
table.TsexED <- cbind(Sex, tableTSED)
View(table.TsexED)
```

```
# Getting the N
sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")
((sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")) /
   (sum(cases_onlyED$trauma.o %in% 1)))*100

sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")
((sum(cases_onlyED$trauma.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")) /
   (sum(cases_onlyED$trauma.o %in% 1)))*100

## IO -------------------

model5.TS1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = trauma.o == 1 & sex_group == "a. Male")

model5.TS2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = trauma.o == 1 & sex_group == "b. Female")

sexT1IO <- OR.bin.summ(model5.TS1IO)
sexT2IO<- OR.bin.summ(model5.TS2IO)

Sex <- c("Male", "Female")
tableTSIO <- rbind(sexT1IO, sexT2IO)
table.TsexIO <- cbind(Sex, tableTSIO)
View(table.TsexIO)

# Getting the N
sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")
((sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")) /
   (sum(cases_onlyIO$trauma.o %in% 1)))*100

sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")
((sum(cases_onlyIO$trauma.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")) /
   (sum(cases_onlyIO$trauma.o %in% 1)))*100

##-----------------------------------------------------
## Lag groups

## Lag 1
model5.T1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                   data= waphism, subset = trauma.o == 1)

model5.T2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                   data = waphisED, subset = trauma.o == 1)

model5.T3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
```

```
                    data = waphisIO, subset = trauma.o == 1)

    overallTL1 <- OR.bin.summ(model5.T1OL1)
    EDTL1 <- OR.bin.summ(model5.T2EDL1)
    otherTL1 <- OR.bin.summ(model5.T3IOL1)

    Service_TL1 <- c("All cause", "ED only", "Inpatient and Observation")
    tableTL1 <- rbind(overallTL1, EDTL1, otherTL1)
    table.servTL1 <- cbind(Service_TL1, tableTL1)
    View(table.servTL1)

    ## Getting the N
    sum(cases_onlyO$trauma.o %in% 1)- sum(cases_onlyO$trauma.o %in% 1 &
is.na(cases_onlyO$Exposure_1))
    (1-((sum(cases_onlyO$trauma.o %in% 1 & is.na(cases_onlyO$Exposure_1))/
         sum(cases_onlyO$trauma.o %in% 1))))*100

    sum(cases_onlyED$trauma.o %in% 1)- sum(cases_onlyED$trauma.o %in% 1 &
is.na(cases_onlyED$Exposure_1))
    (1-((sum(cases_onlyED$trauma.o %in% 1 & is.na(cases_onlyED$Exposure_1))/
         sum(cases_onlyED$trauma.o %in% 1))))*100

    sum(cases_onlyIO$trauma.o %in% 1)- sum(cases_onlyIO$trauma.o %in% 1 &
is.na(cases_onlyIO$Exposure_1))
    (1-((sum(cases_onlyIO$trauma.o %in% 1 & is.na(cases_onlyIO$Exposure_1))/
         sum(cases_onlyIO$trauma.o %in% 1))))*100

    ## Lag 2 --------------------

    model5.T1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = trauma.o == 1)

    model5.T2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = trauma.o == 1)

    model5.T3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = trauma.o == 1)

    overallTL2 <- OR.bin.summ(model5.T1OL2)
    EDTL2 <- OR.bin.summ(model5.T2EDL2)
    otherTL2 <- OR.bin.summ(model5.T3IOL2)

    Service_TL2 <- c("All cause", "ED only", "Inpatient and Observation")
    tableTL2 <- rbind(overallTL2, EDTL2, otherTL2)
    table.servTL2 <- cbind(Service_TL2, tableTL2)
    View(table.servTL2)
```

```
## Getting the N
sum(cases_onlyO$trauma.o %in% 1)- sum(cases_onlyO$trauma.o %in% 1 &
is.na(cases_onlyO$Exposure_2))
(1-((sum(cases_onlyO$trauma.o %in% 1 & is.na(cases_onlyO$Exposure_2))/
    sum(cases_onlyO$trauma.o %in% 1))))*100

sum(cases_onlyED$trauma.o %in% 1)- sum(cases_onlyED$trauma.o %in% 1 &
is.na(cases_onlyED$Exposure_2))
(1-((sum(cases_onlyED$trauma.o %in% 1 & is.na(cases_onlyED$Exposure_2))/
    sum(cases_onlyED$trauma.o %in% 1))))*100

sum(cases_onlyIO$trauma.o %in% 1)- sum(cases_onlyIO$trauma.o %in% 1 &
is.na(cases_onlyIO$Exposure_2))
(1-((sum(cases_onlyIO$trauma.o %in% 1 & is.na(cases_onlyIO$Exposure_2))/
    sum(cases_onlyIO$trauma.o %in% 1))))*100

## Lag 3 --------------------
model5.T1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
              data= waphism, subset = trauma.o == 1)

model5.T2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
              data = waphisED, subset = trauma.o == 1)

model5.T3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
              data = waphisIO, subset = trauma.o == 1)

overallTL3 <- OR.bin.summ(model5.T1OL3)
EDTL3 <- OR.bin.summ(model5.T2EDL3)
otherTL3 <- OR.bin.summ(model5.T3IOL3)

Service_TL3 <- c("All cause", "ED only", "Inpatient and Observation")
tableTL3 <- rbind(overallTL3, EDTL3, otherTL3)
table.servTL3 <- cbind(Service_TL3, tableTL3)
View(table.servTL3)

## Getting the N
sum(cases_onlyO$trauma.o %in% 1)- sum(cases_onlyO$trauma.o %in% 1 &
is.na(cases_onlyO$Exposure_3))
(1-((sum(cases_onlyO$trauma.o %in% 1 & is.na(cases_onlyO$Exposure_3))/
    sum(cases_onlyO$trauma.o %in% 1))))*100

sum(cases_onlyED$trauma.o %in% 1)- sum(cases_onlyED$trauma.o %in% 1 &
is.na(cases_onlyED$Exposure_3))
(1-((sum(cases_onlyED$trauma.o %in% 1 & is.na(cases_onlyED$Exposure_3))/
    sum(cases_onlyED$trauma.o %in% 1))))*100
```

```
    sum(cases_onlyIO$trauma.o %in% 1)- sum(cases_onlyIO$trauma.o %in% 1 &
is.na(cases_onlyIO$Exposure_3))
    (1-((sum(cases_onlyIO$trauma.o %in% 1 & is.na(cases_onlyIO$Exposure_3))/
        sum(cases_onlyIO$trauma.o %in% 1))))*100

    ## Lag 4 --------------------
    model5.T1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = trauma.o == 1)

    model5.T2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisED, subset = trauma.o == 1)

    model5.T3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
                    data = waphisIO, subset = trauma.o == 1)

    overallTL4 <- OR.bin.summ(model5.T1OL4)
    EDTL4 <- OR.bin.summ(model5.T2EDL4)
    otherTL4 <- OR.bin.summ(model5.T3IOL4)

    Service_TL4 <- c("All cause", "ED only", "Inpatient and Observation")
    tableTL4 <- rbind(overallTL4, EDTL4, otherTL4)
    table.servTL4 <- cbind(Service_TL4, tableTL4)
    View(table.servTL4)

    ## Getting the N
    sum(cases_onlyO$trauma.o %in% 1)- sum(cases_onlyO$trauma.o %in% 1 &
is.na(cases_onlyO$Exposure_4))
    (1-((sum(cases_onlyO$trauma.o %in% 1 & is.na(cases_onlyO$Exposure_4))/
        sum(cases_onlyO$trauma.o %in% 1))))*100

    sum(cases_onlyED$trauma.o %in% 1)- sum(cases_onlyED$trauma.o %in% 1 &
is.na(cases_onlyED$Exposure_4))
    (1-((sum(cases_onlyED$trauma.o %in% 1 & is.na(cases_onlyED$Exposure_4))/
        sum(cases_onlyED$trauma.o %in% 1))))*100

    sum(cases_onlyIO$trauma.o %in% 1)- sum(cases_onlyIO$trauma.o %in% 1 &
is.na(cases_onlyIO$Exposure_4))
    (1-((sum(cases_onlyIO$trauma.o %in% 1 & is.na(cases_onlyIO$Exposure_4))/
        sum(cases_onlyIO$trauma.o %in% 1))))*100

    ##----------------------------------------------------
    # Model 5.mh: Mental Health Outcomes
    ##----------------------------------------------------
    ##----------------------------------------------------
    ## Age groups
```

```
## Overall -------------------

model5.MA1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = mh.o == 1 & age.cat == 1)

model5.MA2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = mh.o == 1 & age.cat == 2)

model5.MA3O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphism, subset = mh.o == 1 & age.cat == 3)

ageM1O <- OR.bin.summ(model5.MA1O)
ageM2O<- OR.bin.summ(model5.MA2O)
ageM3O <- OR.bin.summ(model5.MA3O)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableMAO <- rbind(ageM1O, ageM2O, ageM3O)
table.MageO <- cbind(Age, tableMAO)
View(table.MageO)

# Getting the N
sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 1)
((sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 1)) /
   (sum(cases_onlyO$mh.o %in% 1)))*100

sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 2)
((sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 2)) /
   (sum(cases_onlyO$mh.o %in% 1)))*100

sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 3)
((sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$age.cat %in% 3)) /
   (sum(cases_onlyO$mh.o %in% 1)))*100

## ED -------------------

model5.MA1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = mh.o == 1 & age.cat == 1)

model5.MA2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = mh.o == 1 & age.cat == 2)

model5.MA3ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                    data= waphisED, subset = mh.o == 1 & age.cat == 3)

ageM1ED <- OR.bin.summ(model5.MA1ED)
```

```
ageM2ED<- OR.bin.summ(model5.MA2ED)
ageM3ED <- OR.bin.summ(model5.MA3ED)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableMAED <- rbind(ageM1ED, ageM2ED, ageM3ED)
table.MageED <- cbind(Age, tableMAED)
View(table.MageED)

# Getting the N
sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 1)
((sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 1)) /
   (sum(cases_onlyED$mh.o %in% 1)))*100

sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 2)
((sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 2)) /
   (sum(cases_onlyED$mh.o %in% 1)))*100

sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 3)
((sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$age.cat %in% 3)) /
   (sum(cases_onlyED$mh.o %in% 1)))*100

## IO -------------------

model5.MA1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = mh.o == 1 & age.cat == 1)

model5.MA2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = mh.o == 1 & age.cat == 2)

model5.MA3IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphisIO, subset = mh.o == 1 & age.cat == 3)

ageM1IO <- OR.bin.summ(model5.MA1IO)
ageM2IO<- OR.bin.summ(model5.MA2IO)
ageM3IO <- OR.bin.summ(model5.MA3IO)

Age <- c("0 - 5", "6 - 12", "13 - 19")
tableMAIO <- rbind(ageM1IO, ageM2IO, ageM3IO)
table.MageIO <- cbind(Age, tableMAIO)
View(table.MageIO)

# Getting the N
sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 1)
((sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 1)) /
   (sum(cases_onlyIO$mh.o %in% 1)))*100
```

```
sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 2)
((sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 2)) /
   (sum(cases_onlyIO$mh.o %in% 1)))*100

sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 3)
((sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$age.cat %in% 3)) /
   (sum(cases_onlyIO$mh.o %in% 1)))*100

##-----------------------------------------------------
## Sex groups

## Overall -------------------

model5.MS1O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = mh.o == 1 & sex_group == "a. Male")

model5.MS2O <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = mh.o == 1 & sex_group == "b. Female")

sexM1O <- OR.bin.summ(model5.MS1O)
sexM2O<- OR.bin.summ(model5.MS2O)

Sex <- c("Male", "Female")
tableMSO <- rbind(sexM1O, sexM2O)
table.MsexO <- cbind(Sex, tableMSO)
View(table.MsexO)

# Getting the N
sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")
((sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$sex_group %in% "a. Male")) /
   (sum(cases_onlyO$mh.o %in% 1)))*100

sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")
((sum(cases_onlyO$mh.o %in% 1 & cases_onlyO$sex_group %in% "b. Female")) /
   (sum(cases_onlyO$mh.o %in% 1)))*100

## ED -------------------

model5.MS1ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                   data= waphisED, subset = mh.o == 1 & sex_group == "a. Male")

model5.MS2ED <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                   data= waphisED, subset = mh.o == 1 & sex_group == "b. Female")

sexM1ED <- OR.bin.summ(model5.MS1ED)
sexM2ED<- OR.bin.summ(model5.MS2ED)
```

```
Sex <- c("Male", "Female")
tableMSED <- rbind(sexM1ED, sexM2ED)
table.MsexED <- cbind(Sex, tableMSED)
View(table.MsexED)

# Getting the N
sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")
((sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$sex_group %in% "a. Male")) /
   (sum(cases_onlyED$mh.o %in% 1)))*100

sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")
((sum(cases_onlyED$mh.o %in% 1 & cases_onlyED$sex_group %in% "b. Female")) /
   (sum(cases_onlyED$mh.o %in% 1)))*100

## IO -------------------

model5.MS1IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = mh.o == 1 & sex_group == "a. Male")

model5.MS2IO <- clogit(case ~ Smoke.day + ns(Humidex, df=3) + strata(patient_id),
                 data= waphisIO, subset = mh.o == 1 & sex_group == "b. Female")

sexM1IO <- OR.bin.summ(model5.MS1IO)
sexM2IO<- OR.bin.summ(model5.MS2IO)

Sex <- c("Male", "Female")
tableMSIO <- rbind(sexM1IO, sexM2IO)
table.MsexIO <- cbind(Sex, tableMSIO)
View(table.MsexIO)

# Getting the N
sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")
((sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$sex_group %in% "a. Male")) /
   (sum(cases_onlyIO$mh.o %in% 1)))*100

sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")
((sum(cases_onlyIO$mh.o %in% 1 & cases_onlyIO$sex_group %in% "b. Female")) /
   (sum(cases_onlyIO$mh.o %in% 1)))*100

##----------------------------------------------------
## Lag groups

## Lag 1
model5.M1OL1 <-  clogit(case ~ Exposure_1 + ns(Humidex, df=3) + strata(patient_id),
                  data= waphism, subset = mh.o == 1)
```

```
model5.M2EDL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisED, subset = mh.o == 1)

model5.M3IOL1 <- clogit(case ~ Exposure_1 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisIO, subset = mh.o == 1)

overallML1 <- OR.bin.summ(model5.M1OL1)
EDML1 <- OR.bin.summ(model5.M2EDL1)
otherML1 <- OR.bin.summ(model5.M3IOL1)

Service_ML1 <- c("All cause", "ED only", "Inpatient and Observation")
tableML1 <- rbind(overallML1, EDML1, otherML1)
table.servML1 <- cbind(Service_ML1, tableML1)
View(table.servML1)

## Getting the N
sum(cases_onlyO$mh.o %in% 1)- sum(cases_onlyO$mh.o %in% 1 &
is.na(cases_onlyO$Exposure_1))
(1-((sum(cases_onlyO$mh.o %in% 1 & is.na(cases_onlyO$Exposure_1))/
    sum(cases_onlyO$mh.o %in% 1))))*100

sum(cases_onlyED$mh.o %in% 1)- sum(cases_onlyED$mh.o %in% 1 &
is.na(cases_onlyED$Exposure_1))
(1-((sum(cases_onlyED$mh.o %in% 1 & is.na(cases_onlyED$Exposure_1))/
    sum(cases_onlyED$mh.o %in% 1))))*100

sum(cases_onlyIO$mh.o %in% 1)- sum(cases_onlyIO$mh.o %in% 1 &
is.na(cases_onlyIO$Exposure_1))
(1-((sum(cases_onlyIO$mh.o %in% 1 & is.na(cases_onlyIO$Exposure_1))/
    sum(cases_onlyIO$mh.o %in% 1))))*100

## Lag 2 --------------------

model5.M1OL2 <-  clogit(case ~ Exposure_2 + ns(Humidex, df=3) + strata(patient_id),
                data= waphism, subset = mh.o == 1)

model5.M2EDL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisED, subset = mh.o == 1)

model5.M3IOL2 <- clogit(case ~ Exposure_2 + (ns(Humidex, df=3)) + strata(patient_id),
                data = waphisIO, subset = mh.o == 1)

overallML2 <- OR.bin.summ(model5.M1OL2)
EDML2 <- OR.bin.summ(model5.M2EDL2)
otherML2 <- OR.bin.summ(model5.M3IOL2)
```

```
Service_ML2 <- c("All cause", "ED only", "Inpatient and Observation")
tableML2 <- rbind(overallML2, EDML2, otherML2)
table.servML2 <- cbind(Service_ML2, tableML2)
View(table.servML2)

## Getting the N
sum(cases_onlyO$mh.o %in% 1)- sum(cases_onlyO$mh.o %in% 1 &
is.na(cases_onlyO$Exposure_2))
(1-((sum(cases_onlyO$mh.o %in% 1 & is.na(cases_onlyO$Exposure_2))/
    sum(cases_onlyO$mh.o %in% 1))))*100

sum(cases_onlyED$mh.o %in% 1)- sum(cases_onlyED$mh.o %in% 1 &
is.na(cases_onlyED$Exposure_2))
(1-((sum(cases_onlyED$mh.o %in% 1 & is.na(cases_onlyED$Exposure_2))/
    sum(cases_onlyED$mh.o %in% 1))))*100

sum(cases_onlyIO$mh.o %in% 1)- sum(cases_onlyIO$mh.o %in% 1 &
is.na(cases_onlyIO$Exposure_2))
(1-((sum(cases_onlyIO$mh.o %in% 1 & is.na(cases_onlyIO$Exposure_2))/
    sum(cases_onlyIO$mh.o %in% 1))))*100

## Lag 3 --------------------
model5.M1OL3 <-  clogit(case ~ Exposure_3 + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = mh.o == 1)

model5.M2EDL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
            data = waphisED, subset = mh.o == 1)

model5.M3IOL3 <- clogit(case ~ Exposure_3 + (ns(Humidex, df=3)) + strata(patient_id),
            data = waphisIO, subset = mh.o == 1)

overallML3 <- OR.bin.summ(model5.M1OL3)
EDML3 <- OR.bin.summ(model5.M2EDL3)
otherML3 <- OR.bin.summ(model5.M3IOL3)

Service_ML3 <- c("All cause", "ED only", "Inpatient and Observation")
tableML3 <- rbind(overallML3, EDML3, otherML3)
table.servML3 <- cbind(Service_ML3, tableML3)
View(table.servML3)

## Getting the N
sum(cases_onlyO$mh.o %in% 1)- sum(cases_onlyO$mh.o %in% 1 &
is.na(cases_onlyO$Exposure_3))
(1-((sum(cases_onlyO$mh.o %in% 1 & is.na(cases_onlyO$Exposure_3))/
    sum(cases_onlyO$mh.o %in% 1))))*100
```

```
sum(cases_onlyED$mh.o %in% 1)- sum(cases_onlyED$mh.o %in% 1 &
is.na(cases_onlyED$Exposure_3))
(1-((sum(cases_onlyED$mh.o %in% 1 & is.na(cases_onlyED$Exposure_3))/
  sum(cases_onlyED$mh.o %in% 1))))*100

sum(cases_onlyIO$mh.o %in% 1)- sum(cases_onlyIO$mh.o %in% 1 &
is.na(cases_onlyIO$Exposure_3))
(1-((sum(cases_onlyIO$mh.o %in% 1 & is.na(cases_onlyIO$Exposure_3))/
  sum(cases_onlyIO$mh.o %in% 1))))*100

## Lag 4 --------------------
model5.M1OL4 <-  clogit(case ~ Exposure_4 + ns(Humidex, df=3) + strata(patient_id),
            data= waphism, subset = mh.o == 1)

model5.M2EDL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
            data = waphisED, subset = mh.o == 1)

model5.M3IOL4 <- clogit(case ~ Exposure_4 + (ns(Humidex, df=3)) + strata(patient_id),
            data = waphisIO, subset = mh.o == 1)

overallML4 <- OR.bin.summ(model5.M1OL4)
EDML4 <- OR.bin.summ(model5.M2EDL4)
otherML4 <- OR.bin.summ(model5.M3IOL4)

Service_ML4 <- c("All cause", "ED only", "Inpatient and Observation")
tableML4 <- rbind(overallML4, EDML4, otherML4)
table.servML4 <- cbind(Service_ML4, tableML4)
View(table.servML4)

## Getting the N
sum(cases_onlyO$mh.o %in% 1)- sum(cases_onlyO$mh.o %in% 1 &
is.na(cases_onlyO$Exposure_4))
(1-((sum(cases_onlyO$mh.o %in% 1 & is.na(cases_onlyO$Exposure_4))/
  sum(cases_onlyO$mh.o %in% 1))))*100

sum(cases_onlyED$mh.o %in% 1)- sum(cases_onlyED$mh.o %in% 1 &
is.na(cases_onlyED$Exposure_4))
(1-((sum(cases_onlyED$mh.o %in% 1 & is.na(cases_onlyED$Exposure_4))/
  sum(cases_onlyED$mh.o %in% 1))))*100

sum(cases_onlyIO$mh.o %in% 1)- sum(cases_onlyIO$mh.o %in% 1 &
is.na(cases_onlyIO$Exposure_4))
(1-((sum(cases_onlyIO$mh.o %in% 1 & is.na(cases_onlyIO$Exposure_4))/
  sum(cases_onlyIO$mh.o %in% 1))))*100
```

```
##----------------------------------------------------
## Making descriptives table
## ---------------------------------------------------
## ---------------------------------------------------
load(file = "waphis_filled2_03.24.rda")
load(file = "waphis_filled_ED.rda")
load(file = "waphis_filled_nonED.rda")
waphism <- waphis_filled2_03.24
waphisED <- waphis_filled_ED
waphisIO <- waphis_filled_nonED

cases_onlyO <- waphism[waphism$case %in% 1,]
cases_onlyED <- waphisED[waphisED$case %in% 1,]
cases_onlyIO <- waphisIO[waphisIO$case %in% 1,]

View(cases_onlyO)

refs_onlyO <- waphism[waphism$case %in% 0,]
refs_onlyED <- waphisED[waphisED$case %in% 0,]
refs_onlyIO <- waphisIO[waphisIO$case %in% 0,]
## ---------------------------------------------------
## creating frequency table for county origins
dev.off()

tab1(cases_onlyED$county_name, sort.group = "decreasing", cum.percent = FALSE, graph
= FALSE) -> EDcounty
print(EDcounty)
write.csv(EDcounty$output.table, file="EDcounty.csv")

tab1(cases_onlyIO$county_name, sort.group = "decreasing", cum.percent = FALSE, graph =
FALSE) -> IOcounty
print(IOcounty)
write.csv(IOcounty$output.table, file="IOcounty.csv")

## creating frequency table for zip origins

tab1(cases_onlyED$zip, sort.group = "decreasing", cum.percent = FALSE, graph = FALSE)
-> EDzips
print(EDzips)
write.csv(EDzips$output.table, file="EDzips.csv")

tab1(cases_onlyIO$zip, sort.group = "decreasing", cum.percent = FALSE, graph = FALSE) -
> IOzips
print(IOzips)
write.csv(IOzips$output.table, file="IOzips.csv")
## ---------------------------------------------------
```

```
## Making the table Daily PM2.5 characteristics
summary(cases_onlyED$pm25_zipwt)
sd(cases_onlyED$pm25_zipwt)
quantile(cases_onlyED$pm25_zipwt)

summary(refs_onlyED$pm25_zipwt)
sd(refs_onlyED$pm25_zipwt)
quantile(refs_onlyED$pm25_zipwt)

summary(cases_onlyIO$pm25_zipwt)
sd(cases_onlyIO$pm25_zipwt)
quantile(cases_onlyIO$pm25_zipwt)

summary(refs_onlyIO$pm25_zipwt)
sd(refs_onlyIO$pm25_zipwt)
quantile(refs_onlyIO$pm25_zipwt)

# Wildfire Days
table(waphisED$Smoke.day %in% 1)

waphisEDsd <- waphisED[waphisED$Smoke.day %in% 1,]
summary(waphisEDsd$pm25_zipwt)
sd(waphisEDsd$pm25_zipwt)
summary(waphisEDsd$Humidex)
sd(waphisEDsd$Humidex)

View(waphisEDnosd)
waphisEDnosd <- waphisED[waphisED$Smoke.day %in% 0,]
summary(waphisEDnosd$pm25_zipwt)
sd(waphisEDnosd$pm25_zipwt)
summary(waphisEDnosd$Humidex)
sd(waphisEDnosd$Humidex)

waphisIOsd <- waphisIO[waphisIO$Smoke.day %in% 1,]
summary(waphisIOsd$pm25_zipwt)
sd(waphisIOsd$pm25_zipwt)
summary(waphisIOsd$Humidex)
sd(waphisIOsd$Humidex)

View(waphisIOnosd)
waphisIOnosd <- waphisIO[waphisIO$Smoke.day %in% 0,]
summary(waphisIOnosd$pm25_zipwt)
sd(waphisIOnosd$pm25_zipwt)
summary(waphisIOnosd$Humidex)
sd(waphisIOnosd$Humidex)
```

```
## ------------------------------------------------------
## Making figures of data
## ------------------------------------------------------
library(ggplot2)
library(tidyverse)
## ------------------------------------------------------
windowsFonts(A = windowsFont("Open Sans"))
theme_set(theme_bw(16))
## ----------------------------------------------------------------------
## Aim 1: Overall
## ----------------------------------------------------------------------

overall.ED <- data.frame(domain = c('All-cause'),
                         index=1:1,
                         OR = 1.00,
                         lower= 0.97,
                         upper= 1.03)

overall.IO <- data.frame(domain = c('All-cause'),
                         index=1:1,
                         OR = 1.07,
                         lower= 1.03,
                         upper= 1.12)

overall.gg <- ggplot() +
  geom_errorbar(data = overall.ED, aes(x = domain, y= OR, ymin=lower,
                                       ymax=upper, color = "ED"),
                width=0.2, cex=0.8) +
  geom_point(data = overall.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = overall.IO, aes(x = domain, y= OR, ymin=lower,
                                       ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
                position = position_nudge(x = 0.35)) +
  geom_point(data = overall.IO, aes(x = domain, y = OR),
             position = position_nudge(x = 0.35))+
  labs(title='ORs for All-Cause Hospital Services', x='', y = 'Odds Ratio (95% CI)', color =
'Hospital Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.95, 1.15)) +
  theme_minimal()

overall.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
                   axis.text = element_text(face="bold"))

## ----------------------------------------------------------------------
## Aim 1: Lag Plots
## ----------------------------------------------------------------------
```

```
lag.ED <- data.frame(domain = c('Lag 0', 'Lag 1', 'Lag 2', 'Lag 3', 'Lag 4'),
                     index=1:5,
                     OR=c(1.00, 1.00, 1.01, 1.01, 1.00),
                     lower=c(0.97, 0.97, 0.98, 0.98, 0.98),
                     upper=c(1.03, 1.03, 1.03, 1.03, 1.03))

lag.IO <- data.frame(domain = c('Lag 0', 'Lag 1', 'Lag 2', 'Lag 3', 'Lag 4'),
                     index=1:5,
                     OR=c(1.07, 1.08, 1.05, 1.06, 1.05),
                     lower=c(1.03, 1.04, 1.02, 1.03, 1.01),
                     upper=c(1.12, 1.12, 1.09, 1.10, 1.09))

lag.gg <- ggplot() +
  geom_errorbar(data = lag.ED, aes(x = domain, y= OR, ymin=lower,
                                ymax=upper, color = "ED"),
              width=0.2, cex=0.8) +
  geom_point(data = lag.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = lag.IO, aes(x = domain, y= OR, ymin=lower,
                                ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
              position = position_nudge(x = 0.35)) +
  geom_point(data = lag.IO, aes(x = domain, y = OR),
            position = position_nudge(x = 0.35))+
  labs(title='ORs for Lag Days 0 - 4', x='', y = 'Odds Ratio (95% CI)', color = 'Hospital
Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.95, 1.15)) +
  theme_minimal()

lag.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
             axis.text = element_text(face="bold"))

## ----------------------------------------------------
## Age Plots
## ----------------------------------------------------

ages.ED <- data.frame(domain = c('0 - 5', '6 - 12', '13 - 19'),
                      index=1:3,
                      OR=c(1.00, 1.02, 1.00),
                      lower=c(0.96, 0.96, 0.93),
                      upper=c(1.04, 1.07, 1.07))

ages.IO <- data.frame(domain = c('0 - 5', '6 - 12', '13 - 19'),
                      index=1:3,
                      OR=c(1.06, 1.08, 1.08),
                      lower=c(1.00, 1.00, 1.00),
```

```
                  upper=c(1.13, 1.12, 1.16))

## ---------------------------------------------------
## Plot
ages.gg <- ggplot() +
  geom_errorbar(data = ages.ED, aes(x = domain, y= OR, ymin=lower,
                                    ymax=upper, color = "ED"),
                width=0.2, cex=0.8) +
  geom_point(data = ages.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = ages.IO, aes(x = domain, y= OR, ymin=lower,
                                    ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
                position = position_nudge(x = 0.35)) +
  geom_point(data = ages.IO, aes(x = domain, y = OR),
             position = position_nudge(x = 0.35))+
  labs(title='ORs by Age Groups', x='Age Groups', y = 'Odds Ratio (95% CI)', color =
'Hospital Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.9, 1.20)) +
  theme_minimal()

ages.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
                axis.text = element_text(face="bold"))

## ---------------------------------------------------
## Cause Plots
## ---------------------------------------------------

cause0.ED <- data.frame(domain = c('Mental Health', 'Trauma', 'Dermal', 'Resp Infections',
                                   'Respiratory', 'All Cases'),
                        index=1:6,
                        OR=c(0.94, 0.99, 1.10, 1.11, 1.09, 1.00),
                        lower=c(0.80, 0.92, 0.95, 1.01, 1.01, 0.97),
                        upper=c(1.10, 1.05, 1.27, 1.21, 1.17, 1.03))

cause0.IO <- data.frame(domain = c('Mental Health', 'Trauma', 'Dermal', 'Resp Infections',
                                   'Respiratory', 'All Cases'),
                        index=1:6,
                        OR=c(1.15, 1.44, 1.61, 1.14, 0.97, 1.07),
                        lower=c(0.98, 1.03, 0.88, 0.89, 0.84, 1.03),
                        upper=c(1.35, 2.02, 2.94, 1.45, 1.11, 1.12))

cause0.gg <- ggplot() +
  geom_errorbar(data = cause0.ED, aes(x = domain, y= OR, ymin=lower,
                                      ymax=upper, color = "ED"),
                width=0.2, cex=0.8) +
  geom_point(data = cause0.ED, aes(x = domain, y = OR)) +
```

```
    geom_errorbar(data = cause0.IO, aes(x = domain, y= OR, ymin=lower,
                               ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
             position = position_nudge(x = 0.35)) +
    geom_point(data = cause0.IO, aes(x = domain, y = OR),
            position = position_nudge(x = 0.35))+
    labs(title='Cause of Admission ORs', x='', y = 'Odds Ratio (95% CI)', color = 'Hospital
Service') +
    geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
    scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.75, 3.00)) +
    theme_minimal()

  cause0.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
              axis.text = element_text(face="bold"))

  ## ---------------------------------------------------
  ## Sex Plots
  ## ---------------------------------------------------

  sexes.ED <- data.frame(domain = c('Both', 'Male', 'Female'),
                  index=1:3,
                  OR=c(1.00, 1.00, 1.00),
                  lower=c(0.97, 0.96, 0.96),
                  upper=c(1.03, 1.04, 1.04))

  sexes.IO <- data.frame(domain = c('Both', 'Male', 'Female'),
                  index=1:3,
                  OR=c(1.07, 1.05, 1.10),
                  lower=c(1.03, 0.99, 1.04),
                  upper=c(1.12, 1.11, 1.17))

  ## ---------------------------------------------------
  ## Plot
  sexes.gg <- ggplot() +
    geom_errorbar(data = sexes.ED, aes(x = domain, y= OR, ymin=lower,
                              ymax=upper, color = "ED"),
             width=0.2, cex=0.8) +
    geom_point(data = sexes.ED, aes(x = domain, y = OR)) +
    geom_errorbar(data = sexes.IO, aes(x = domain, y= OR, ymin=lower,
                              ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
             position = position_nudge(x = 0.35)) +
    geom_point(data = sexes.IO, aes(x = domain, y = OR),
            position = position_nudge(x = 0.35))+
    labs(title='ORs by Sexes', x='', y = 'Odds Ratio (95% CI)', color = 'Hospital Service') +
    geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
    scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.95, 1.20)) +
    theme_minimal()
```

```
sexes.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
             axis.text = element_text(face="bold"))
## --------------------------------------------------
## Insurance Plots
## --------------------------------------------------

ins.ED <- data.frame(domain = c('All', 'Government', 'Private', 'Other'),
                index=1:4,
                OR=c(1.00, 1.01, 0.99, 1.05),
                lower=c(0.97, 0.97, 0.95, 0.90),
                upper=c(1.03, 1.05, 1.04, 1.23))

ins.IO <- data.frame(domain = c('All', 'Government', 'Private', 'Other'),
                index=1:4,
                OR=c(1.07, 1.07, 1.07, 1.23),
                lower=c(1.03, 1.01, 1.00, 0.99),
                upper=c(1.12, 1.13, 1.13, 1.51))

## --------------------------------------------------
## Plot
ins.gg <- ggplot() +
  geom_errorbar(data = ins.ED, aes(x = domain, y= OR, ymin=lower,
                          ymax=upper, color = "ED"),
           width=0.2, cex=0.8) +
  geom_point(data = ins.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = ins.IO, aes(x = domain, y= OR, ymin=lower,
                          ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
           position = position_nudge(x = 0.35)) +
  geom_point(data = ins.IO, aes(x = domain, y = OR),
          position = position_nudge(x = 0.35))+
  labs(title='ORs by Insurance Type', x='Insurance Type', y = 'Odds Ratio (95% CI)', color =
'Hospital Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.80, 1.60)) +
  theme_minimal()

ins.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
           axis.text = element_text(face="bold"))
## --------------------------------------------------
## Race Plots
## --------------------------------------------------

race.ED <- data.frame(domain = c('Non-Hispanic White', 'Non-Hispanic Black', 'Hispanic',
                        'Asian', 'Other', 'All'),
                 index=1:6,
```

```
              OR=c(1.00, 1.02, 1.02, 0.98, 0.97, 1.00),
              lower=c(0.96, 0.94, 0.96, 0.88, 0.91, 0.97),
              upper=c(1.05, 1.11, 1.09, 1.09, 1.04, 1.03))

race.IO <- data.frame(domain = c('Non-Hispanic White', 'Non-Hispanic Black', 'Hispanic',
                    'Asian', 'Other', 'All'),
              index=1:6,
              OR=c(1.11, 1.30, 1.07, 0.86, 1.00, 1.07),
              lower=c(1.04, 1.11, 0.98, 0.71, 0.91, 1.03),
              upper=c(1.18, 1.53, 1.16, 1.03, 1.09, 1.12))

## ---------------------------------------------------
## Plot

race.gg <- ggplot() +
  geom_errorbar(data = race.ED, aes(x = domain, y= OR, ymin=lower,
                      ymax=upper, color = "ED"),
         width=0.2, cex=0.8) +
  geom_point(data = race.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = race.IO, aes(x = domain, y= OR, ymin=lower,
                      ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
         position = position_nudge(x = 0.35)) +
  geom_point(data = race.IO, aes(x = domain, y = OR),
        position = position_nudge(x = 0.35))+
  labs(title='Race Group ORs', x='Race Groups', y = 'Odds Ratio (95% CI)', color = 'Hospital
Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.60, 1.60)) +
  theme_minimal()

race.gg + theme(text = element_text(family = "A"), plot.title = element_text(hjust = 0.5),
         axis.text = element_text(face="bold"))

## ---------------------------------------------------
## Age x Cause
## ---------------------------------------------------

agec.ED <- data.frame(domain = c('Mental Health', 'Trauma', 'Dermal', 'Resp Infections',
                    'Respiratory', 'All Cases'),
              index=1:6,
              OR=c(0.94, 0.99, 1.10, 1.11, 1.09, 1.00),
              lower=c(0.80, 0.92, 0.95, 1.01, 1.01, 0.97),
              upper=c(1.10, 1.05, 1.27, 1.21, 1.17, 1.03))

agec.IO <- data.frame(domain = c('Mental Health', 'Trauma', 'Dermal', 'Resp Infections',
                    'Respiratory', 'All Cases'),
```

```
                    index=1:6,
                    OR=c(1.15, 1.44, 1.61, 1.14, 0.97, 1.07),
                    lower=c(0.98, 1.03, 0.88, 0.89, 0.84, 1.03),
                    upper=c(1.35, 2.02, 2.94, 1.45, 1.11, 1.12))

agec.gg <- ggplot() +
  geom_errorbar(data = agec.ED, aes(x = domain, y= OR, ymin=lower,
                            ymax=upper, color = "ED"),
          width=0.2, cex=0.8) +
  geom_point(data = agec.ED, aes(x = domain, y = OR)) +
  geom_errorbar(data = agec.IO, aes(x = domain, y= OR, ymin=lower,
                            ymax=upper, color = "Inpt./Obvs."), width=0.2, cex=0.8,
          position = position_nudge(x = 0.35)) +
  geom_point(data = agec.IO, aes(x = domain, y = OR),
        position = position_nudge(x = 0.35))+
  labs(title='ORs for All Ages by Admission Cause', x='Cause of Admission', y = 'Odds Ratio
(95% CI)', color = 'Hospital Service') +
  geom_hline(color='black', yintercept = 1, linetype='solid', alpha=.5) +
  scale_y_continuous(labels = function(x) format(x, nsmall = 1), limits = c(0.80, 3.00)) +
  theme_minimal()
```