Norwegian University of Life Sciences

Master's Thesis 2022 30 ECTS
Faculty for Landscape and Society

Understanding the role of socioeconomic conditions on temperature exposure and cardiorespiratory mortality in Oslo

Sammenhengen mellom sosioøkonomiske faktorer, temperatureksponering og kardiorespiratorisk dødelighet i Oslo

Camilla Skau Nyland
Master in Public Health Science

Camilla Skau Nyland

Understanding the role of socioeconomic conditions on temperature exposure and cardiorespiratory mortality in Oslo

Master's in public health science

Spring 2022

Supervised by Shilpa Rao-Skirbekk and Suleman Atique

Norwegian University of Life Sciences

Faculty for Landscape and Society

Postboks 5003

1432 Ås

Norway

Preface

This thesis marks the end of two interesting and learning full years in Public Health Science at NMBU. I have really enjoyed the experience.

There are many people for whom I want to thank for contributing in various ways to the process of writing this thesis. I would like to say a huge thank you to my main supervisor Shilpa Rao at Norwegian Institute of Public Health (NIPH). Thank you for giving me the opportunity to be part of the EXHAUSTION project and for so warningly including me in the warm and exciting environment at NIPH, and in many of the interesting seminars, conferences and other happenings both at NIPH and outside. I have learned so much by working with you these last months and I have really enjoyed being at NIPH and part of the EXHAUSTION project. Your support and guidance along the way have meant a lot. I am truly grateful. I would also like to thank Suleman Atique, my main supervisor from NMBU, for providing supervision, concrete advice and support along the way. Furthermore, I want to thank my co-supervisor Bente Oftedal at NIPH. I am highly grateful for your guidance, concrete feedback, help and support. I would also like to thank Norun Krog at NIPH for contributing with valuable guidance and input in meetings along the way.

The research in this thesis has received data from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820655 – EXHAUSTION. I would like to offer my acknowledgment to the EXHAUSTION project. There are many people connected to the project I want to thank. Terese Bekkevold NIPH, thank you for providing and helping with data, and input in meetings. Thank you to Alfonso Palomares for all the efforts with GIS. A special thank you to Siqi Zhang, for providing R-scripts and help with the statistical analyses. I would not have been able to do this thesis without any of you.

A big thank you to my co-student Kicki Svennson, for great collaboration and support throughout this process. It would not have been the same without you. I would also like to thank Meizu Chen and Christian Rikter-Ydse for all your help and guidance in R-studio. It made my R-journey manageable and is highly appreciated. Furthermore, thank you to Geir Aamodt at NMBU for supervision regarding methods and results. Lastly, I would like to thank my friends and family for supporting and cheering me on, especially on less motivated days.

Camilla Skau Nyland, Oslo, May 2023

Abstract

Background: Climate change is the greatest public health threat of our time. Increasing temperatures may increase mortality related to CRDs, particularly in urban areas. People of lower socioeconomic status may be more vulnerable to temperature related health impacts. In Oslo, there are large social health inequalities, including prevalence of CRDs. The role of socioeconomic conditions on temperature related mortality in Oslo is little explored, and research is needed to develop equitable climate adaptation strategies.

Objective: Investigate the association between ambient high and low temperature exposure and cardiorespiratory mortality in Oslo, and how this association may be modified by socioeconomic conditions.

Methods: This study used a time-stratified case-crossover design with cohort and -registry based data. Daily mortality and temperature data for the period 2000-2018 were used to model the temperature-mortality association for CVD and respiratory mortality, and household income and education level were used as effect modifiers. I estimated the OR of mortality for cold and heat effects and interaction of effect modifiers using conditional logistic regression and distributed lag non-linear model (DLNM) including various lag-structures. Additional analyses were conducted, stratified by lag-structures, age, sex and seasonality.

Results: The results showed no statistically significant association of exposure to extreme cold and hot temperatures on CVD or respiratory mortality except for respiratory mortality when including lag days 0-3. However, the results indicated higher ORs of respiratory mortality from heat effects, while CVD seemed to have higher ORs from cold effects. Females had higher ORs from heat effects (respiratory mortality), while males had higher ORs of mortality from cod effects. No statistically significant effect modification of education or income were found. Though, it indicated a tendency to a trend of increasing risk of CVD mortality from heat effects with lower levels of education. There were similar tendencies for the respiratory mortality, however with extremely wide confidence intervals.

Conclusion: Further research on this topic is needed for Oslo, with a larger population sample. Nevertheless, it will be important to consider socioeconomic conditions when developing climate adaptation strategies for Oslo considering the existing and increasing social health inequalities in the population**.**

Sammendrag

Bakgrunn: Klimaendringer er vår tids største folkehelsetrussel. Økende temperaturer kan øke dødelighet relatert til kardiorespiratorisk sykdom, særlig i urbane områder. Personer med lavere sosioøkonomisk status kan være mer sårbare for temperaturerelaterte påvirkninger. I Oslo er det store sosiale helseforskjeller, inkludert for prevalens av kardiorespiratorisk sykdom. Hvordan sosioøkonomiske faktorer påvirker temperaturerelatert dødelighet i Oslo er lite utforsket, og forskning er nødvendig for å utvikle rettferdige klimatilpasningsstrategier.

Formål: Undersøke sammenhengen mellom utendørstemperatur relatert til kulde og hete og kardiorespiratorisk dødelighet i Oslo, og hvordan denne sammenhengen kan modereres av sosioøkonomiske faktorer.

Metode: Denne studien brukte et tidsstratifisert case-crossover design med kohort- og registerbaserte data. Dødelighet og temperaturdata for perioden 2000-2018 ble brukt til å modellere sammenhengen mellom temperatur og kardiovaskulær (CVD) og respiratorisk dødelighet. Husholdningsinntekt og utdanningsnivå ble brukt som effektmoderatorer. Jeg estimerte OR for dødelighet relatert til kulde- og hette og interaksjon av effektmoderatorene ved å bruke betinget logistisk regresjon og Distributed Lag Non-Linear Model (DLNM) inkludert ulike lag-strukturer. Tilleggsanalyser ble utført, stratifisert etter lag-strukturer, alder, kjønn og sesonger.

Resultater: Resultatene viste ingen statistisk signifikant sammenheng mellom eksponering for ekstrem kulde og hete på CVD eller respiratorisk dødelighet, bortsett fra respiratorisk dødelighet ved inklusjon av 0-3 lag-dager. Imidlertid indikerte resultatene høyere OR av respiratorisk dødelighet fra heteeffekter, mens OR for CVD var høyere for kuldeeffekter. Kvinner hadde høyere OR av heteeffekter (respiratorisk dødelighet), mens menn hadde høyere OR av kulde. Ingen statistisk signifikant effektnodifikasjon av utdanning eller inntekt ble funnet. Resultatene indikerte en tendens til en trend med økende risiko for kardiovaskulær dødelighet fra hete for lavere utdanningsnivåer. Det var lignende tendenser for respiratorisk dødelighet, men med ekstremt brede konfidensintervaller.

Konklusjon: Det er behov for videre forskning på dette temaet for Oslo, med et større befolkningsutvalg. Det vil likevel være viktig å vurdere sosioøkonomiske faktorer ved utvikling av klimatilpasningsstrategier for Oslo med tanke på de eksisterende og økende sosiale helseforskjellene i befolkningen.

Abbreviations

CHD	Coronary heart disease
CI	Confidence interval
CONOR	Cohort of Norway
COPD	Chronic obstructive pulmonary disease
CRD	Cardiorespiratory diseases
CVD	Cardiovascular diseases
DLNM	Distributed lag non-linear model
DÅR	Dødsårsaksregisteret (Cause of Death Registry)
GIS	Geographical Information Systems
HUBRO	Helseundersøkelsen i Oslo (Oslo health study)
ICD	International Classification of Diseases
IPCC	Intergovernmental Panel on Climate Change
MI	Myocardial infarction
MMT	Minimum mortality temperature
NCD	Non-communicable disease
NDI	Neighborhood deprivation index
NIPH	Norwegian Institute of Public Health
NOU	Norges offentlige utredninger
OR	Odds ratio
RD	Respiratory diseases
REK	Regionale Komiteer for medisinsk og helsefaglig forskningsetikk
SD	Standards deviation
SES	Socioeconomic status
SDG	Sustainable Development Goal
SDH	Social determinants of health
SSB	Statistisk Sentralbyrå (Statistics Norway)
UHI	Urban heat island
UN	United Nations
WHO	World Health Organization

Table of content

List of figures

List of tables

List of appendixes

1. Introduction

Due to man-made climate change, the world is experiencing a warmer climate, with an increase in global mean temperatures of 1,2°C since pre-industrial time (Romanello et al., 2021). The past seven years are reported as the hottest seven years on record, globally, and Europe has warmed faster overall than any other continent (Copernicus Climate Change Service, 2023). Climate change is considered the greatest public health threat of our time (World Health Organization, 2021b). A warmer climate has many direct and indirect impacts for health, of which more frequent and intense temperature events are one of them. Cold and hot temperatures are considered important risk factors for disease, particularly for cardiorespiratory diseases (CRDs), which consist of cardiovascular (CVDs) and respiratory diseases (RDs). CRDs are some of the leading causes of death worldwide, including Norway (Ariansen et al., 2020; World Health Organization, 2021a). Projections indicate an increase in excess mortality related to temperature exposure, though with great geographical variations between and within countries, and dependent on the extent of global temperature increases (Cissé et al., 2022).

Populations living in urban areas more vulnerable to climate change impacts, such as temperature deviations, due to the built environment and urbanization (World Health Organization, 2021c). More than 55% of the global population now lives in urban areas, and this number is expected to increase to 68% by 2050. In Norway, around 82% of the population lives in towns or cities (Statistisk sentralbyrå, 2022c). Other factors that make people vulnerable to climate change impacts are socioeconomic conditions, with social inequalities being considered an important vulnerability factor for the health impacts of climate change (Cissé et al., 2022). Several studies have shown an increased mortality risk from temperature exposure among people with lower socioeconomic status (SES), although the evidence is inconsistent (Son et al., 2019).

The 2030 UN Sustainable Development Goals (SDGs) is a global call for action to fight a range of social, environmental and development issues while at the same time tackle climate change (United Nations, s.a.-a). Climate adaptation is one way to tackle climate change and reducing vulnerability, by adjusting to the current and expected impacts (Leichenko & O'Brien, 2019). The UN SDGs work as a framework for Norwegian policy, including public health work (Meld. St. 19 (2018-2019))**.** Social health inequalities are a great public health concern in Norway, particularly in Oslo, where there are also great social health inequalities

between the districts. Reducing social health inequalities is one of the main goals of the Norwegian public health work. Prevalence of CRDs is higher among lower SES groups, and in Oslo lower SES groups also experience higher rates of CRD mortality (Oslo kommune, 2020b).

Already in 2010, The Official Norwegian Report (NOU) "Tilpasning til et klima i endring" recommended research on impacts of heat exposure on mortality and the significance this may have for society (NOU 2010: 10). Furthermore, the Norwegian Environment Agency underlines that the social dimensions and climate change vulnerability detected in the IPCC report from 2022, is highly relevant for Norway (Miljødirektoratet, 2022b). The potential skewed effects of climate change for lower SES populations, may contribute to further enhance the existing social health inequalities. Increased mortality risk related to temperature may be one of them. However, this topic is little explored in Norway. So far, limited knowledge exists on the role of socioeconomic conditions related to temperature exposure and impacts on public health in Norway.

The insights from research are crucial for policymakers in the development of equitable climate adaptation strategies. Considering the concentrated population and urbanization in the urban area of Oslo and the built environment, the great social health inequalities among the people living there, especially with regards to CRDs, it will be relevant to enhance knowledge about the role of socioeconomic conditions in the temperature-mortality relationship in Oslo. Therefore, this thesis will investigate how socioeconomic conditions impact the association of temperature exposure and mortality in Oslo, focusing on short-term exposure and CRD mortality.

1.1 Structure of thesis

This thesis is written as a monography, consisting of seven chapters. After this introduction chapter follows chapter 2 with a presentation of empirics and chosen theory relevant for the thesis aim. Aim and research questions is presented in chapter 3 In chapter 4, the thesis' methods are explained, as well as a short evaluation of ethical considerations of this thesis. Following this, chapter 5 presents the results from the statistical analyses. In chapter 6 the results are discussed considering the theoretical background and existing literature, as well as a discussion of the strengths and limitations of this thesis. Lastly, in chapter 7 I present my conclusion based on the discussion in chapter 6 and the implications of this for further research and public health work.

2. Background

In this chapter I present the theoretical and empirical background that forms the backdrop of this thesis. Relevant terms and theories will be explored, as well as the existing knowledge of the thesis` theme. Furthermore, I present the political and legal guidelines for the public health work relevant for the theme.

2.1 Climate change

2.1.1 Climate change-related temperature deviations

Climate change has led to many regions experiencing higher daily minimum and maximum temperatures than they used to (Leichenko & O'Brien, 2019). This is especially evident during seasons such as hotter summers, where regions might experience temperatures that deviate from their average temperatures. An increase in the number of extreme temperature events such as heatwaves in the future will lead to more people being exposed to extreme temperatures (Pörtner et al., 2022). Several parts of the world have experienced heatwaves during the last two decades, with the summer of 2022 being the most recent example of this, affecting regions like India, Pakistan, and Europe. Europe experienced its hottest summer in 2022 at 1.4 °C above the 1991-2020 average, and it was the second warmest year on record (Copernicus Climate Change Service, 2023). Extreme heat during the late spring and summer of 2022 resulted in conditions that could be hazardous for human health.

When it comes to cold temperatures however, extreme cold events are reported to have decreased in frequency and intensity on a global level and are projected to consistently decrease for most warming levels (Cissé et al., 2022). However, cold weather events can still take place periodically and affect urban areas and their connected infrastructure.

2.1.2 Temperature exposure in an urban context

Urbanization is one of the greatest global trends with significant impact on health (World Health Organization, 2021c). According to Bednar-Friedl et al. (2022), the number of European residents living in urban areas was 547 million in 2015, corresponding to 74% of the total European population at the time. This number is expected to increase up to 84% by 2050.

Urban areas are considered drivers of climate change (World Health Organization, 2021c). The expansion and increasing densification of the built environment in urban areas can lead to the so-called *urban heat island effect* (UHI) (Reinwald et al., 2021). UHIs are

characterized by temperature differences between urban areas and the surrounding suburban and rural areas. The temperature differences can be up to several °C, usually around 2-4°C. The UHI effect is mainly being caused by the build-up and superstructure of natural permeable surfaces covered with vegetation. Urban areas usually consist of more materials like concrete and paving which absorb sun energy during the day, and slowly release this energy into the air as heat (Heaviside et al., 2017). Even though there is less incoming sun energy in the winter season, the UHI effect can be strong throughout the year.

The UHI effect can impact health for urban populations in several ways, whereby the most direct health impact is through exposure to increased temperatures (Heaviside et al., 2017). Additionally, it can interact with and worsen air pollution. This may lead to higher temperature-exposure for those living in urban areas, especially during extreme events, which potentially can exacerbate heat-related health impact, including mortality. Generally, people living in urban areas can have an increased heat related mortality risk due to the UHI effect. During the 2003-heatwave in Europe, mortality increased by three times among residents in London, and in France, Paris and other cities were particularly affected (Heaviside et al., 2017; The Lancet, 2018). According to European Environment Agency (2018), urban populations may experience twice as many days of heatwave compared to their surrounding rural areas.

2.1.3 Climate change vulnerability and adaptation

Climate change impacts are highly heterogonous depending on geographics, sociodemographics and socioeconomic conditions (Pörtner et al., 2022). *Climate change vulnerability* can be defined as *"the predisposition or likelihood of being adversely affected by a climatic event or circumstance"* (Leichenko & O'Brien, 2019, p.140). Furthermore, vulnerability is a result of the interaction between exposure to climate change and various factors that affect both susceptibility and capacity to adapt. Climate change vulnerability varies across time, countries and regions, and among groups and subgroups of communities. According to the report from the Intergovernmental Panel on Climate Change (IPCC) from 2022, vulnerability reflects variations and changes in macro-scale non-climatic factors like population changes, economic development, infrastructure, behavior, technology and ecosystems, as well as individual- or household-specific characteristics such as age, health status, access to livelihood amenities and socioeconomic factors like education, income and employment (Pörtner et al., 2022).

The risk and vulnerability to climate change impacts can be reduced by adaptation to the climate changes and the accompanying consequences. *Climate change adaptation* has been defined by the IPCC (2014, cited in Eriksen, 2015, p.4) as "*the process of adjustment to actual or expected climate and its effects. In human systems, adaptation seeks to moderate or avoid harm or exploit beneficial opportunities*". In 2015 all the member countries in The United Nations (UN) agreed on *The 2030 Agenda for Sustainable Development* and the 17 SDGs aiming at securing peace and wealth for all citizens of the world, while protecting the planet (United Nations, s.a.-a).

"Leave no one behind" is a central promise of the SDGs, to reduce inequalities and vulnerabilities among and within countries that leave people behind, by combating the root causes (United Nations, s.a.-b). At the UN Climate Change Conference (COP21) in Paris in 2015, world leaders agreed on the legally binding *Paris Agreement* (United Nations, s.a.-c). With The Agreement all the 193 Parties agreed to reduce global greenhouse gas emissions to limit the global temperature increase in this century to 2 degrees Celsius while also aiming for an even further reduction to 1,5 degrees, as well as reducing climate vulnerability, strengthen resilience and increase capacity to adapt to climate impacts.

The degree to which a society is capable to adapt to climate changes and it´s impacts is described as the society's *adaptive capacity (Oslo Kommune, 2020a).* The adaptive capacity depends on how activities are organized, available knowledge and resources, what is prioritized and how. To a large extent, solutions and adaptive capacity lie in the areas of land use such as green areas, buildings and infrastructure, which prevents climate change impacts in other areas of society as well (Oslo Kommune, 2020a).

2.1.4 Climate change in the Norwegian context and Oslo

Norway´s commitment to the Paris Agreement is enshrined in the Norwegian Climate Change Act (klimaloven, 2017). The Norwegian government is working towards cutting emissions by 50-55% by 2030 compared to 1990-levels and by 90-95% by 2050. In White Paper No 13 (2020-2021) it is stated that the adaptive capacity of Norway shall be strengthened (Meld. St. 13 (2020-2021)).

The NOU "Tilpasning til et klima I endring" from 2010 wrote about climate change impacting public health in many ways (NOU 2010: 10)**.** The report recommended a regular

update on the knowledge base for climate change and its impacts, including analysis of vulnerability and adaptation-needs and climate projections. In White Paper No.33 (2012-2013) “Climate adaptation in Norway”, Norwegian government stressed the importance of climate adaptation and underlines that it must be an integrated part of the work of actors in various fields (Meld. St. 33 (2012-2013)). It is essential that climate change considerations are included in planning, developing- and decision-making processes in all sectors in society and at all levels. Furthermore, the white paper places much of the future responsibility for adaptation on the municipalities. In 2014, the municipality of Oslo developed a climate adaptation strategy, in the woke of White Paper No.33 (2012-2013). The Climate adaptation strategy is updated regularly in line with increased knowledge about the future climate and based on the experiences that the municipality gains in the area (Oslo Kommune, 2013).

The average temperature in Norway has increased by approximately 1.2°C from 1901 to 2021, and the rate of changes has increased in recent decades (Miljødirektoratet, 2022a). Data from the Norwegian Meteorological Institute show an increase in the number and geographical distribution of hot days, defined as days where the average temperature is higher than 20 °C. In Norway as a whole, winters have become shorter. In Tromsø and Oslo, winters have become 22 days shorter between the two standard normal periods 1961-1990 and 1991-2020.

The temperature in Oslo has increased by 1,5°C and extreme events have occurred more often over the past century (Oslo Kommune, 2020a). The city is particularly vulnerable to climate change due to the dense population, urbanization and built environment. The basin-shaped topography and compact city center may lead to UHIs. Oslo is projected to have an increase in inhabitants of about 15,6% in 2050 compared to 2022 (Statistisk sentralbyrå, 2022a). Demographic changes like an expected increase in the elderly population will in itself impact general health outcomes in the population (Jore et al., 2022). Predictions on the future population in Oslo predicts that the number of people aged 80-89 years and 99 years and older, will double (Oslo Kommune, 2023). However, the population in Oslo is relatively young, and predictions for 2040 show that the population will remain relatively young.

Hanssen-Bauer et al. (2015) projected the future climate in Norway and impacts from climate change, using different climate change scenarios based on models from IPCC (2014). Due to the precautionary-principle they used the high-emission-scenario as a basis for the impact-

assessment. This projected a warmer climate towards 2071-2100, with an increase of annual mean-temperature of 4,5 °C, (span: 3,3 to 6,4 °C). The greatest warming is projected for winter-season, and least warming for the summer-season. However, the projections for the lower-emission-scenarios are significantly lower, respectively 2,7°C, and 1,5°C for the lowest-emission-scenario. There is a big difference between an increase of 1,5°C and 2°C, regarding impacts and adaptation-possibilities (Pörtner et al., 2022). Particularly a 4,5 °C-increase will have drastic consequences for health. Temperature exposure is associated with mortality, with CVDs and RDs being identified as some of the main causes of death (Ragettli et al., 2023).

2.2 Temperature exposure and cardiorespiratory mortality

2.2.1 Cardiorespiratory health

CRDs involves both CVDs and RDs, and they are considered *non-communicable diseases* (NCDs) (World Health Organization, 2021a). CVDs are a group of diseases that affect the heart and blood vessels. They include coronary heart disease (CHD), myocardial infarction (MI), cerebrovascular disease like stroke, rheumatic heart disease and other conditions. Chronic respiratory diseases are a group of diseases that affect the airways and other structures of the lungs (World Health Organization, s.a.-a). Asthma and chronic obstructive pulmonary disease (COPD) are some of the most common chronic respiratory diseases.

Health behavior factors, such as, unhealthy diet (especially high-fatty foods), physical inactivity, smoking tobacco and use of alcohol and substance-abuce are some of the most common risk factors for developing CVDs, as well as hypertension, high cholesterol and overweight (Oslo kommune, 2020b). COPD is caused by exposure to harmful particles or gasses, of which smoking is considered the most important cause and may explain around two thirds of COPD-cases in Norway. Air pollution is another important cause of COPD. Asthma in childhood is also a risk factor for developing COPD.

According to WHO, CVDs are the number one cause of death globally (World Health Organization, 2021a). Around 17,9 million people died from CVDs in 2019, which makes up for 32% of all global deaths. The same year, around 262 million people globally were affected by asthma and around 455 000 people died from asthma. COPD is the third leading cause of death globally and caused 3.23 million deaths in 2019.

In Norway, the burden of disease is dominated by NCDs, which cause approximately 87% of the total burden of disease (Meld. St. 15 (2022-2023)). An estimated 1/5 of the population is affected by CVDs in Norway, either having a CVD or having high risk of developing a CVD (Ariansen et al., 2020). The prevalence of CVDs in Norway has decreased the past decades, due to healthier living conditions and preventive work, as well as more effective treatment. However, risk factors for CVD, such as smoking, physical inactivity and unhealthy diets are still prevalent, which suggests that CVDs will still affect many people. According to the Cause of Death Registry in Norway (DÅR), RDs, including COPD, were the fourth most frequently registered cause of death in Norway in 2021, and the fraction of people living with COPD will remain high in the years to come due to the increase in of elderly (Nystad, 2022).. Both CVD and COPD, in addition to cancer and diabetes, causes most of the premature mortality-cases in Norway, which is defined as death before the age of 75 years (Meld. St. 19 (2018-2019)). In Oslo, CVDs are the most common cause of death for men and the second most common for women (Oslo kommune, 2020b).

2.2.2 Ambient temperature exposure and CRD mortality

Both cold and hot temperature-exposure are established as risk factors for health with studies showing important mortality effects in populations across the globe (Cissé et al., 2022; Masselot et al., 2023). The way in which exposure to temperatures impacts the human body is well documented and understood (Watts et al., 2019). An increased body temperature due to heat exposure can make the heart beat faster and harder, dilate the veins - particularly on the skin-surface and increase sweating to remove heat and regulate the body temperature. This can also lead to dehydration, which can be a critical condition (Liu et al., 2022). When exposed to cold temperatures however, the body tries to keep the warm body temperature, e.g., by constriction of veins, and cold exposure can increase peripheral resistance (Saucy et al., 2021).

Extreme deviations in temperature can exacerbate existing health conditions, potentially leading to or worsening disease and death (Heaviside et al., 2017; Liu et al., 2022). Thus, mortality associated with extreme temperatures is not necessarily caused directly by hypo, - or hyperthermia or heat stroke, but by other indirect causes such as CRDs that are triggered when the human body attempts to adapt to the excessive environment (Liu et al., 2015).

Exposure to heat can be a risk to all people, but people of older age and those with pre-existing health conditions, particularly CRDs, are more sensitive to temperature extremes

than the general population (Åström et al., 2016). Aging leads to physiological changes in the thermo-regulation and reduced cardiovascular capacity (Liu et al., 2022). Furthermore, there is a higher prevalence of pre-existing health conditions as well as associated medication use among elderly people, which also might explain increased risk of negative CRD outcomes when exposed to temperature.

Many studies have investigated the effects of ambient temperature on mortality, the so-called *temperature-mortality association*, of which many have found an increased mortality risk due to cold and hot temperatures (Gasparrini et al., 2015; Gasparrini et al., 2022; Masselot et al., 2023; Ragettli et al., 2023; Rocklöv et al., 2014). Several of the studies have investigated cause-specific mortality and have found a positive association of temperature exposure and CVD and respiratory mortality (Åström et al., 2018; Bunker et al., 2016; Liu et al., 2015; Liu et al., 2022; Saucy et al., 2021; Witt et al., 2015).

Several respiratory diseases are climate sensitive based on how people are exposed to the factors causing these diseases (Cissé et al., 2022). Several exposure pathways contribute to RDs in general, and some of these exposure pathways are climate related including changes in ambient air pollution concentrations. Temperature increase may alter concentration of air pollutants, and air pollution could modify the temperature-mortality association (Paavola, 2017).

The relationship between temperature and mortality is often described as a U or (inverse) J-V-shaped curve, with a trough at a so-called *minimum mortality temperature* (MMT), and then increasing slopes towards both hot and cold tails of the temperature distribution (Åström et al., 2016; Liu et al., 2015; Ruuhela et al., 2018). The MMT refers to the temperature at which the mortality is at a minimum, often referred to as the *optimum temperature.*

The risk of temperature related mortality has shown to be significant at even moderately high temperatures, although the most severe health impacts often happen during periods of more extreme temperatures such as heatwaves or cold spells, which have been associated with large excess mortality (Heaviside et al., 2017; Rocklöv et al., 2014). For instance, in 2003 the heatwave in Europe caused about 70 000 excess deaths (Saucy et al., 2021).

However, the temperature-mortality association is highly heterogeneous across geographical areas and climatic zones (Åström et al., 2016; Ruuhela et al., 2018). The MMT is often lower in cooler compared to warmer climatic zones. According to Åstrøm et al (2018), mortality

risks related to high and low temperature exposure tend to be higher in colder and warmer regions, respectively, suggesting partial adaptation of population to their own climate,

Furthermore, the association is a complex one due to the non-linear and delayed (lagged) impacts of thermal stress on health. Mortality related to heat effects usually occur relatively shortly after temperatures have begun to increase, often on the same day and last a few days (Gasparrini et al., 2015; Liu et al., 2015; Ruuhela et al., 2018), Cold effects might take longer to emerge, typically after a couple of days and last about 10 days till weeks. The excess mortality due to prolonged duration of extreme temperatures may be due to physiological exhaustion related to cumulative stress over many consecutive days (Rocklöv et al., 2014). Moreover, heat and cold stress can potentially lead to a displacement in mortality called *"harvesting",* with death occurring earlier than it would have otherwise (Liu et al., 2015). Health impacts of extreme temperatures also vary by seasonality, e.g., health impacts of extreme heat can be more severe in the spring than later in the season.

Winter mortality is mostly caused by CVDs and RDs (Åström et al., 2016). Cold temperatures lead to increased rates of several CVDs and RDs such as MI, hypothermia and influenza (Bunker et al., 2016). Despite increasing warming and decreasing cold events, there is inconsistent evidence of whether mortality related to extreme cold events will decrease in future decades in European urban areas (Cissé et al., 2022). This is partly due to various medical factors (such as increased cardiac risk factors and influenza seasons) also contribute to this excess mortality in winter periods. Higher regional variability in future climates means that extreme cold events still can be important mortality risks locally. This will be reinforced in many urban areas, particularly European cities, due to the expected increase of elderly and the high prevalence of CRDs.

Hajat et al. (2014) projected the risk of future heat-and cold-related mortality to be significantly raised across all regions in the UK in the 2050s. Despite the very high increase of heat related mortality; an increase of 257% compared to the current annual baseline of 2000 deaths and 2% decrease in cold related mortality from a baseline of around 41 000 deaths, the burden of cold temperature continued to be higher than the heat burden in all periods.

Although there is a projected temperature increase in Norway, extreme temperatures like heatwaves will not have the same mortality impacts as in the Southern Europe, due to the colder climate in Norway (Aamaas, 2019). In fact, the decrease in cold related mortality is

expected to be bigger than the increase in heat-related mortality in the Northern Europe (Gasparrini et al., 2017). However, they did not consider different vulnerable subgroups in their study, nor did they include data from Norway.

A study from Sweden found that heatwaves significantly increased both all-cause mortality and CHD-mortality (Åström et al., 2018). On the contrary, the heatwave in the summer of 2018 in Norway did not increase the net mortality among people over 75 and 85 years (Ranhoff et al., 2019), although they did not investigate cause-specific mortality.

Though it is uncertain how ambient temperature in Norway will change in the future and what health impacts this will have for the Norwegian population, it is certain that it will not impact all groups equally (Ólafsdóttir, 2021). Future temperature related mortality also depends on vulnerability factors in the population, such as socioeconomic conditions. Climate change can be considered a driver of social health inequalities through already vulnerable people being hit harder by climate change impacts (Oslo Kommune, 2023). To stem the effects of climate change on health, World Health Organization (WHO) stresses that one must act against the inequalities that are the fundamental causes of health challenges. Next chapter explains how socioeconomic conditions can affect the mortality risk related to temperature exposure. Prior to this, I will set the context by exploring the framework of social health inequalities. To understand how socioeconomic conditions can impact the temperature-mortality association, one needs to understand how socioeconomic conditions are associated with health in general.

2.3 Social inequalities in health

2.3.1 Social determinants of health

Health is not just about diseases such as CRDs or mortality, but rather the absence of diseases. In 1946 WHO defined health as *"a state of complete physical, mental and social well-being, not merely the absence of disease or infirmity*" (Naidoo & Wills, 2016p.4). The concept of health is holistic, consisting of individual, societal, environmental, and global dimensions. Within these dimensions there are many factors impacting the population's health, so-called *health determinants.* Social factors influencing health are called *social determinants of health* (SDH). WHO describes SDH as non -medical factors that impact health outcomes, and are the conditions in which people are born, grow, work, live and age (World Health Organization, s.a.-b). SDH are determined by the wider set of forces and

systems shaping the conditions of daily life. This includes economic policies and systems, development agendas, social policies and norms, and political systems. Some of the SDH have extensive and complex impacts on health, such as economy and education. Other SDH are simpler and more direct in the way they impact health, for instance physical environment and tobacco (Meld. St. 15 (2022-2023)). Usually, the determinants come first and then health follows because of the determinants.

A rainbow-model *"the main determinants of health"* by Dahlgren and Whitehead (1991) is a common representation of the SDH, aimed at illustrating how health determinants in various spheres of society and at different levels impact the public health, and how they impact each other (Øversveen & Rydland, 2021) (Figure 1). The outer layer of the model shows the organization of society and includes general socio-economic, cultural, and environmental conditions. In the second layer are the societal institutions that shape living- and working conditions, while the next layer shows social and community networks. The next layer consists of individual lifestyle factors such as many of the risk factors for diseases as mentioned above, namely smoking habits, physical activity, and diets. Finally, the inner layer consisting of age, sex and constitutional factors also influences health, but are considered fixed factors over which people have little control (Dahlgren & Whitehead, 1991).

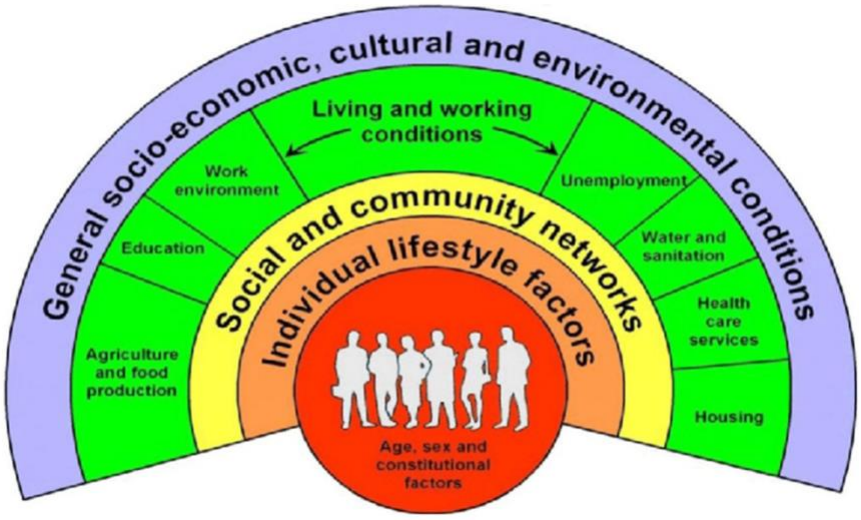

Figure 1*:* Social determinants of health. Adjusted model from Dahlgren and Whitehead (1991).

Poor health is not something that happens by chance or bad luck, rather there are systematic differences in health among the population. There are differences in health determinants, health outcomes such as mortality, and access to healthcare. These differences can be referred to as *inequalities* and *inequities*. Inequalities refer to differences between populations which to a large extent are avoidable. If these differences are also unfair and unjust, they can be described as inequities (Naidoo & Wills, 2016). Inequalities in health are in many cases unjustly distributed among populations based on socioeconomic status (SES) and can thus be described as inequities. *Socioeconomic status* can be described as *"the result of how much and what kind of resources an individual possesses"* Øversveen and Rydland (2021, p.38) Income, education, and employment status are common indicators of SES, and many studies investigating effects of socioeconomic factors involve these factors, including studies investigating effects of temperature on health.

Social health inequalities follow a gradient across society, called the *social gradient* (World Health Organization, s.a.-b). It is not just the poorest in society that experience poor health, rather slightly better SES gives (statistically speaking) slightly better health (Syse et al., 2022). Several studies have explored the association between income inequalities and mortality, suggesting that it is the relative inequalities in income and material resources between populations, rather than the country being rich or poor itself, that affects health (Pickett & Wilkinson, 2015; Rodgers, 1979; Wilkinson & Pickett, 2011). However, as the rainbow model from Dahlgren and Whitehead demonstrates, it is not just the material resources that impacts health. The following section explores why and how social health inequalities arise and persist.

2.3.2 Social conditions as fundamental causes of health inequalities

Various explanations seek to understand health inequalities, focusing on cultural, behavioral, material, structural and psychosocial dimensions, suggesting that adverse environmental conditions at different times during life can lead to poor health (Naidoo & Wills, 2016). The causes are complex, and an accumulation of these conditions contributes to creating social health inequalities.

Other explanations of the association between social inequalities and health are that it is health that impacts the social position, so-called *reversed causality (Meld. St. 19 (2018-2019)).* It means that poor health leads to lower educational level or unemployment and so forth.

Neither of the explanations completely explains the inequalities in health, however Naidoo and Wills (2016) underline that health behaviors should not be separated from their social context. This is in line with Link and Phelans (1995) theory of social conditions as fundamental causes of health inequalities. The theory seek to explain why the association of SES and mortality has persisted despite considerable changes in the disease and risk factors that presumably explain it (Phelan et al., 2010). SES inequalities in mortality now reflect new major causes of death such as CRDs and cancer. Health behaviors often follow the same pattern as education and income level, i.e., poor health behaviors are more common in lower SES populations (Syse et al., 2022).

According to Link and Phelans theory this long-lasting association persists because SES consists of many types of resources, i.e., money, knowledge, prestige, power, and beneficial social connections that protect health, in line with SDH and Øversveen and Rydland`s description of SES. These types of resources are flexible and beneficial in different situations and can be reused when necessary (Ólafsdóttir, 2021). They can enable people to prevent health risks by having a healthier lifestyle, or reduce impacts from disease or events, such as temperature extremes, Education influences employment and income level. But it is also presumed to measure, among other things, an individual's cognitive resources and knowledgebase, in addition to the ability to perceive health information, so-called *health literacy* (Dahll et al., 2014).

A central aspect in the fundamental causes-theory is that of contextualizing risk factors and looking at the bigger picture, i.e. the broader social context that people live in (Ólafsdóttir, 2021). It is important to understand why and how people make certain decisions affecting their health. Flexible resources can be conceptualized as *"the "causes of causes" or "risks of risks" that shape individual health behaviors by influencing whether people know about, have access to, can afford, and receive social support for their efforts to engage in health-enhancing or health protective behaviors"* (Phelan et al., 2010, p.30).

The mortality decline among higher SES populations has been evident in preventable diseases such as CRDs (Phelan et al., 2010). Furthermore, mortality rates for diseases that are not yet preventable have remained the same for low and high SES populations. Hence, the theory hypothesizes that people of higher SES have better capacity to prevent or reduce risk factors for health and promote health, due to flexible resources like knowledge, money, and power. As Ólafsdóttir (2021p. 32) writes "*in the context of climate change, the theory of*

fundamental causes of health inequalities anticipates that those with more resources will be better equipped to avoid the risks of extreme temperature events, as well as minimizing the consequences of them if they occur in their lives".

However, there are some limitations to the theory of Link and Phelan. Their theory has been empirically tested, including a study of 20 European populations from 1980-2010 by Mackenbach et al. (2017). They used mortality data by education level (low, medium, and high) for 22 causes of death. Their results support the fundamental causes-theory, however it also demonstrates other factors and mechanisms impacting mortality, than those implied in the theory. Furthermore, the findings of Mackenbach et al. (2017) also indicate that material resources are not the main hinderance for good health in a modern European context.

Social health inequalities are a key issue in Norwegian public health policy. Next chapter explains social health inequalities in the Norwegian context, including the legal and political framework that underlies the work on public health and social health differences.

2.3.3 Social health inequalities in the Norwegian context and Oslo

The public health of the Norwegian population is generally good and the life expectancy high, being 84.7 years for females and 81.6 years for males in 2021 (Meld. St. 15 (2022-2023)). Although social health inequalities in Norway in general are small compared to other countries, there are still considerable social health inequalities in Norway with an increasing trend since the 1980s (Syse et al., 2022). These inequalities have increased even further after the Covid-19 pandemic, due to the uneven provision of infection prevention measures. The proportion of financially secure households fell from 65% during Covid-19 to 49% in August 2022 (Meld. St. 15 (2022-2023)).

The Norwegian population with a high educational attainment and good economy live on average 5-6 years longer and have fewer health concerns than those with lower educational levels and poorer economy (Meld. St. 15 (2022-2023)). Those with the lowest household income has markedly shorter life expectancy. This group is characterized by a high proportion of people living alone and a low educational level.

Social inequalities in Norway are seen in several diseases, particularly in NCDs (Syse et al., 2022). According to The National Institute of Public Health (NIPH) health behaviors explain much of the inequalities in CVDs. Those with higher educational level smoke less, are more

physical active and generally have a healthier diet, as well as a lower blood pressure and are less obese, than those with lower educational level. It has also been shown that people with short educational level have a higher risk of dying after a heart attack. However, alcohol consumption is higher in higher SES groups, though it seems that the negative consequences related to alcohol consumption are smaller in higher SES groups than in lower SES groups (Syse et al., 2022).

There are clear geographical health inequalities between the counties in Norway, with Oslo standing out negatively (Meld. St. 15 (2022-2023)). Oslo is a city where the great social health inequalities manifest itself as a geographical accumulation of resources and burdens (Oslo Kommune, 2023). Although the living conditions in Oslo are generally good, the city is diverse with big differences between various groups and geographical areas. Neighboring districts, sub-districts and quarters can have very different composition in terms of population age, SES, ethnic background, and housing type. This means that there may be large social and economic differences across relatively small geographical areas, and that various areas in the city have different challenges and opportunities (Oslo kommune, 2020b). It is often referred to as a geographical divide between the eastern and western part of the city.

Generally, Oslo-residents have high educational attainments, however there are large differences among the districts (Oslohelsa, 2020). There is a bigger proportion of people with low educational status and household income among the districts in the eastern part of Oslo, with the opposite proportion-size for the western districts. For instance, the proportion of highly educated people is about double the size in the outer western districts than in the eastern district Groruddalen. The same goes for household income, with the average household income being much higher in the western districts, although low income is not just an east-side-phenomena. Likewise, the proportion of unemployed and people living in overcrowded housing is also higher in the eastern part of Oslo (Oslo Kommune, w.y). In the public health strategy for Oslo 2023-2030 it is stated that 92% of people (aged 20-66) with a university-or college degree of more than 4 years are employed, compared to only 55% of people with primary school as their highest level (Oslo Kommune, 2023). Living conditions and health literacy are also unevenly distributed in the population with a clear association to education, income, and employment (Syse et al., 2022).

In Oslo, people with high educational level, generally experience lower levels of chronic disease and lower mortality, than people with low education (Oslo kommune, 2020b). The

variations between districts in prevalence of CVDs, COPD, lung cancer and diabetes (type 2) essentially follow the same pattern as the district variations in income and education. The population in the western districts are more seldom in contact with their general practitioner or emergency doctor due to CVD and have lower CVD mortality, than people living in some of the eastern districts. There are also great differences between the districts in Oslo when it comes to premature mortality due to COPD, with the lowest proportion in the period 2009-2017 found in the western districts. Men in Oslo who has the highest educational level have around 5.4 years of longer life expectancy than men with the shortest educational attainment. For women, the differences are 4.4 years (Oslo Kommune, 2023)

Life expectancy differences varies between the districts and it can be up to around 7 years of difference for males, and 5 years of difference for females (Meld. St. 15 (2022-2023)). According to the report "Oslohelsa", the differences in life expectancy between districts can be related to the local environment and other characteristics of the local population than their educational status (Oslo kommune, 2020b)

The Norwegian public health work is anchored in the Norwegian Public Health Act, which defines public health work as: "*society's efforts to influence factors that directly or indirectly promote* the populations health and well-being*; prevent mental and somatic illness,* injury or suffering;*; or that protect against health threats; as well as efforts seeking more equal distribution of factors that directly or indirectly affect health"* (folkehelseloven, 2011). The Norwegian Ministry of Health and Welfare underlines that public health work must facilitate health for all people and (strive to) reduce social differences. Public health work should be a part of all sectors, in line with *"health in all policies" (Meld. St. 19 (2018-2019)).*

"Reducing social health inequalities" is the main priority in the newest White Paper (No 15 (2022-2023) and one of the three national goals (Meld. St. 15 (2022-2023)). This White Paper is to a large extent based on a report with updated information about SDH and social inequalities in Norway, conducted by Sir Michael Marmot in cooperation with The Norwegian University of Science and Technology. An important message in this report is that public health measures should be universally oriented, but work best for those who need it most, so-called *proportional universalism.* Furthermore, the Ministry of Health highlights several perspectives which, together with the efforts against social health inequalities, will be of considerable importance for future public health. Climate change is one of them, and they underline the increasingly clear association between climate and health and acknowledges

that climate crisis is also a health crisis. Furthermore, that people of lower SES are more vulnerable to the impacts of climate change. The Norwegian Ministry of Health aims to operationalize the public health goals in line with SDGs, which also forms the basis for the public health policy.

The Public Health Act lays the fundament for the public health work to be systematic (folkehelseloven, 2011). This means that each municipality must have an overview of their population's health status and the health determinants. Social health inequalities are also a main priority in Oslo's public health strategy for 2023-2030, with a focus on equalizing living- and environmental conditions for the residents (Oslo Kommune, 2023). The strategy also mentions climate change and its potential impacts on public health, including the potential skewed impacts due to the inequalities in living - and geographical conditions in the city.

But what does the literature say about the associations of SES, temperature exposure and mortality?

2.4 Social health inequalities and impacts on temperature related mortality

A large body of literature has investigated the potential vulnerability of socioeconomic conditions and inequalities in the temperature-mortality association, in different parts of the world. Effects modification occurs when an exposure has a different health effect among different sub-groups of a population, e.g., if hot or cold temperatures had a higher risk of mortality in people of lower SES than people of higher SES, then effect modification has occurred (Son et al., 2019). Some studies have investigated individual level SES, while others have explored community, - and neighborhood-level SES.

The evidence is inconsistent. Many studies show an increased mortality risk for people of lower SES, whereas other studies show no increased risk or effect modification of socioeconomic factors. A positive association of SES in the temperature-mortality relationship is often related to where and how people live. Social inequalities are also connected to environmental inequalities. In many cases, lower SES populations are disproportionately exposed to hazardous environmental exposures such as higher and lower temperatures, as well as air pollution (European Environment Agency, 2018; World Health Organization, 2019).

In the review by Liu and his colleagues (2015), they found that people with lower SES generally had a higher vulnerability to temperature-related CVD mortality and morbidity. Particularly education and occupation class seemed to be important socioeconomic factors that impacted the population´s vulnerability and capacity to adapt.

However, Son et al. (2019) found limited and inconsistent evidence of effect modification by individual level education and SES, as well as community-level SES in their systematic review and meta-analysis on all-cause temperature related mortality and potential effect modifiers, where they used studies before 2017 from around the world.

A study that investigated the susceptibility to mortality in weather extremes across 135 US cities found that areas where people with higher poverty and lower levels of education live, had a higher susceptibility to extreme heat during warm months (Zanobetti et al., 2013). This is in line with the findings from another US-study, which also found that people living in areas with greater poverty were more vulnerable to heat related mortality (Madrigano, J. et al. (2013). They concluded that local areal vulnerability should be taken in consideration when cities are planning climate adaptation strategies. However, in the study of Yu et al. (2010), they did not find sufficient evidence to conclude that people with lower SES in Brisbane, Australia had a higher risk of dying in hot temperatures.

In a European context, both the study of Gasparrini et al. (2022) and Masselot et al. (2023) suggested that socioeconomic inequalities could be a vulnerability factor for temperature related mortality for both heat and cold. The findings from Gasparrini et al. (2022) indicated that temperature-related mortality impacts were stronger in more deprived areas, in line with the studies from US.

The findings from a study in Stockholm, Sweden, indicated that individuals of lower wealth had higher risks of heat related mortality, however their findings also showed that this association was related to heat wave duration and was strongest for people younger than 65 (Rocklöv et al., 2014). Wealthier populations on the other hand, seemed to have a higher risk of cold related mortality.

The next chapter present the aim and research questions of this thesis.

3. Aims and research question

Drawing on what is explained in the background chapter, my hypothesis is that low and high temperatures will be associated with CRD mortality differently for people of low and high SES in Oslo, whereby lower SES groups will experience increased risk of CRD mortality compared to those with higher SES in Oslo. The findings from this study may aid the understanding of how socioeconomic conditions could modify the temperature-mortality association in Oslo and inform the development of equitable adaptation strategies that target vulnerable populations in Oslo.

Therefore, the following research questions will be investigated:

Main research question:

> *How do socioeconomic conditions impact the association between short-term temperature exposure and cardiovascular and respiratory mortality in Oslo?*

To answer the main research question, I will need to first investigate the following secondary research question:

> *How is short-term air temperature exposure associated with cardiovascular and respiratory mortality in Oslo?*

The objective of this thesis is thus to examine the temperature -mortality association for CVD and respiratory mortality in Oslo and if socioeconomic conditions are effect modifiers of this association, using individual level cohort and registry-based data for the adult population (over 18 years of age).

4. Methods

This chapter describes and explains the data and methods used in the thesis, including the methodological choices.

This thesis is part of an ongoing EU-funded project "Exposure to heat and air pollution in Europe – cardiopulmonary impacts and benefits of mitigation and adaptation" (EXHAUSTION). The EXHAUSTION-project quantifies the future burden of cardiopulmonary disease morbidity and mortality that is attributable to heat and air pollution (EXHAUSTION, 2020). It does so by investigating the association between climate-parameters, air pollution components and cardiopulmonary disease. EXHAUSTION aims to identify adaptation strategies that will help decrease premature death and disease among vulnerable groups in Europe (Zhang et al., 2020). The project started in 2019 and involves 14 partners from 10 European countries, including The Norwegian Institute of Public Health (NIPH), CICERO and The University of Oslo.

For the thesis I have collaborated with the NIPH and I have used data and methods from the EXHAUSTION project. In the EXHAUSTION project analyses on temperature exposure, CPD and effect modification by a variety of variables including socioeconomic variables, have already been conducted on a national level using Cohort of Norway (CONOR) by researchers connected to NIPH. For this thesis, I have applied the same methods and analyses as in EXHAUSTION, but using only an Oslo-cohort, The Oslo Health Study (HUBRO) for Oslo, which will be described further in chapter 4.2 Data material.

4.1 Study design

A time-stratified case-crossover design is used, which is especially suitable to investigate the effects of transient short-term exposures, like temperature, on the risk of health outcomes (Webb et al., 2017). In this design, each individual serves as their own control. Only participants who experienced mortality were included in the statistical analyses. With the time-stratified case-crossover design, one compares the exposure profile of each individual at the day of event (**case day**) with their exposure profile at proximate days before or after the event (**control days**) (Ragettli et al., 2023). In this study, day of death (case day) was matched with three control days, which were the same day of the week within the same month and year prior to the event of death. By making within-participant comparisons between case and control days within the same month, one controls for the potential

confounding effects of participant characteristics, seasonality, long term trend and day of week, by design.

4.2 Data material

I have used individual level data from various data sources, including epidemiological health data from HUBRO and the Norwegian Cause of Death Registry (DÅR), meteorological data from the Norwegian Meteorological Institute as well as socioeconomic data from Statistics Norway (SSB) and HUBRO. All data sources were linked to HUBRO. This data linkage was done beforehand of my thesis, as part of the CONOR analyses. The data sources were linked based on the participants unique personal identification number, and Geographical Information Systems (GIS) was used to link the participants to the meteorological data, based on their historical residential addresses that were updated annually. A dataset consisting of variables from all data sources were made for the CONOR analyses. For my thesis, participants from HUBRO were extracted from this linked dataset, creating a new dataset with only the HUBRO-participants. This was the dataset I used for my analyses.

HUBRO is part of CONOR, which is a collection of health data from several Norwegian health surveys (1994-2003) (Aamodt et al., 2010; Folkehelseinstituttet, 2019). HUBRO was conducted between 2000 and 2001. A total of 18 770 adults took part in HUBRO, which were 46 % of the invited. The participants' year of birth ranged from 1924 to 1955, where the mean year of birth was 1954. The final linked HUBRO dataset included 1070 participants who died from CVDs (N=1070) and 302 who died from respiratory diseases (N=203) during the study period. Participants with missing information on temperature data were excluded. This resulted in an analytical sample of N=1064 for CVD mortality and N=299 for respiratory mortality.

The linkage of HUBRO to the other data sources are illustrated in a flow chart diagram below (Figure 2).

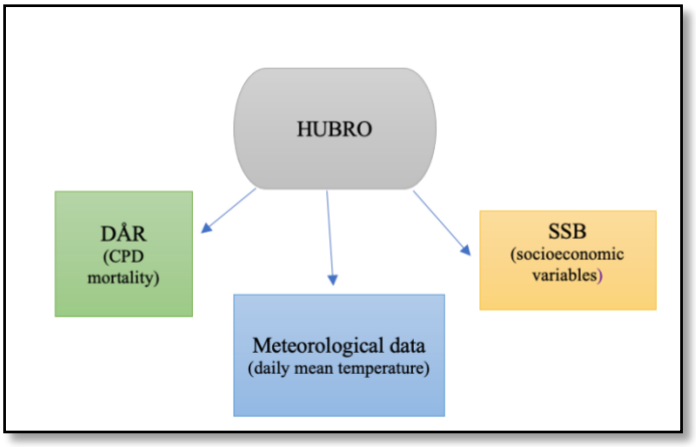

Figure 2: Flow chart diagram shows linkage of the data used in the statistical analyses. Green box shows mortality data, blue box shows exposure data and yellow box shows socioeconomic data.

In the following chapter I will describe the data and explain my choice of SES variables.

4.2.1 Health Data

From DÅR, Individual level daily mortality data for all HUBRO participants from January 1st 2000 to December 31st 2018 were obtained, which makes up the study period in this thesis. DÅR is a national register covering causes of death for all Norwegian residents (Folkehelseinstituttet, 2020). The causes of death are coded according to ICD-10, the 10th revision of the International Classification of Diseases (for the period of 1997 and onward). The mortality data were classified into death due to CVD (ICD-10: I00-I99) RD (ICD-10: J00-J99), and were used as separate health outcomes.

4.2.2 Meteorological data

The meteorological data is estimated from spatial-temporal temperature models. Daily mean, minimum and maximum air temperatures on a 1 km grid across mainland Norway from 1995 to 2018 are obtained from the SeNorge2 dataset, released by the Norwegian Meteorological Institute. The dataset goes back to 1957 and is updated daily based on measurement data.

Daily mean temperature is the exposure variable (independent variable) in this study. The possible delayed effects of temperature exposure on days prior to the event of death was also

accounted for in the analyses (lag days/periods). Variables were created for each exposure day, so temp_day0 for case day, temp_day1 for the day prior to death (lag day 1), temp_day2 for exposure two days prior and so on. These variables are continuous variables.

4.2.3. Socioeconomic data

The socioeconomic variables included as effect modifiers in the analyses are household income and education level. Education is a variable that stays relatively stable after 30 years of age when the majority has completed their education (Syse et al., 2022). This makes it a good variable to use to represent the participants' SES, particularly considering the age of the participants in this study. Income and employment status are also common variables to use as SES-indicators in research. In the CONOR analyses, unemployment was the only SES variable that showed a statistically significant effect modification of the temperature-mortality relationship, and so it would be relevant to also include employment status in this study. However, there were only four unemployed HUBRO participants that we know of. Most of them were pensioners. Thus, the employment variable was not relevant to include in the analyses. Therefore, I chose to include income level as well, as two variables gives a better representation of the role of SES in the temperature-mortality relationship.

Both the income and education level variable were part of the extracted HUBRO dataset. Thus, the decoding of the variables had already been done beforehand. I will explain how this was done.

Income

Household income is the total income for all household members after tax. Income after tax is defined by SSB as *"the sum of wages and salaries, income from self-employment, property income and transfers received minus total assessed taxes and negative transfers"* (Statistisk sentralbyrå, s.a). I had access to household income after tax from 2004 to 2018.

In the analyses the household income from the year of death was used. If this value was missing, it was replaced by the values closest to the year of death that was available. If the information was still missing (missing values), the participant was excluded from the interaction analyses.

The household income variable is a numeric variable that was dichotomized into a dummy-variable for the analyses, with the value 0 for income under median (low) and 1 for income above median (high).

Education

The education level variable is retrieved from both SSB and the HUBRO questionnaire. Originally this categorical variable had six categories both HUBRO and SSB, which was then decoded to three categories: "low" (<high school), "medium" (high school or vocational school) and "high" (university short/long). Similar coding has been done in other studies (Huang et al., 2015; Saucy et al., 2021). Missing information in the HUBRO variable was replaced by the SSB variable for the final education level variable that was used in the analyses.

4.3 Statistical analyses

The statistical analyses were performed using the R programming language (R software version 4.2.1). Original R-codes that were used for the CONOR analyses were used but adapted to HUBRO. These codes were provided by the researcher who conducted the CONOR-analyses. I also made some codes on my own. All statistical tests were two-sided with a 5% significance level.

Statistical analyses were conducted in two stages in line with the study objectives. In the first stage analyses were preformed to investigate the short-term effects of temperature exposure on mortality. In the second stage, the potential effect modification of income and education level on the association of temperature and CRD-mortality was investigated.

4.3.1 Stage 1 – Short term effects of temperature on CRD-mortality

Conditional logistic regression models were used to estimate the short-term effects of temperature on mortality using the analytical sample of participants who died during the study period. Conditional logistic regression takes the "matching" of the cases and controls into account, where the participants of this study are both cases and controls as described in chapter 4.1 and it estimates the adjusted odds ratio (OR) (Jaakkola, 2003). For the regression

analyses individual data used a dummy-variable with the value 1 for case days (= dead) and 0 for control days (= "alive"). CVD and respiratory mortality were analyzed separately.

Distributed lag non-linear model (DLNM) was applied to characterize the potentially non-linear and delayed association between temperature and mortality. Simple statistical models compare mortality and temperature from the same day (Ruuhela et al., 2018). However, distributed lag non-linear models (DLNMs) have been widely applied in recent years to account for the so-called *exposure-lag-response-associations* (Huang et al., 2015).

I conducted first a main analysis, where I included 0 to 10 lag days of temperature, which I will refer to as "the main exposure". However, the approach with 0-10 lag days may potentially underestimate heat effects (which have shorter lags) or cold (which might have longer lags) as demonstrated in the background chapter. Therefore, additional analyses were conducted, where different lag periods of temperature were applied, respectively 0-3 lag days for heat effects and 0-21 lag days for cold effects. Previous studies have used similar lag structures for heat and cold effects (Gasparrini et al., 2015; Huang et al., 2015; Saucy et al., 2021). The sensitivity analyses also included adjusting for seasonality by looking at the cold and heat-effects during the summer period (May-September) and the winter period (November-March) separately, as well as adjusting for sex (female, male) and age (65+, 75+). This was done by analyzing the various strata separately.

The effect of heat was estimated as the odds ratio (OR) of mortality for an increase in daily mean air temperature from the 75th to the 95th percentile of the temperature distribution, and the effect of cold was estimated for a decrease from the 25th to the 1st percentile. The uncertainty of the effect estimates was measured as confidence intervals (CI).

4.3.2 Stage 2 – Effect modification of income and education level

In the second stage, interaction analyses was performed to estimate the potential effect modification of income and education level on the association between temperature and CPD. Effect modification is a statistical technique where one investigates if the association between exposure X and outcome Y varies depending on a third variable M (Webb et al., 2017). The interaction analysis was performed by adding an interaction term between a cross-basis of air temperature and each of the socioeconomic variables to the conditional logistic regression model (temperature*income and temperature*education level). I did separate interaction analyses for income and education level stratified by the various categories (low, high for

income level and low, medium, high for education level), and stratified by cause of death (CVD and respiratory mortality).

4.4 Research ethics

This thesis is conducted in line with ethical principles of research according to the Declaration of Helsinki, a declaration developed by the World Medical Association in 1994 (Torp, 2019). The declaration works as a set of ethical principles for medical research. It particularly highlights the importance of informed consent of participants, the ethical responsibility that lies with the researcher and the considerations of vulnerable groups.

As this thesis is based on cohort and register data, it was necessary to apply for an approval from the Reginal Committee for Medical and Health Research Ethics (REK). A REK application was sent for the EXHAUSTION project. For my thesis, it was necessary with an updated application. The updated approval from REK was received prior to the start of the thesis (Appendix 1). A data managing agreement was established between NIPH and NMBU, to ensure data managing as approved by REK. The data managing agreement included requirements for non-disclosure and confidentiality.

In accordance with the ethical principles, proper and safe storage of the data is required. As I worked with sensitive data containing personal information in this thesis, all statistical analyses were performed in Services for Sensitive Data (TSD), a platform for collecting, storing, analyzing and sharing sensitive data in compliance with the Norwegian privacy regulation (Universitetet I Oslo, s.a.). It is developed by the Norwegian University of Oslo. The participants in HUBRO were given allotted numbers instead of their ID-number to anonymize them. The participants address levels that were used to link the temperature data to HUBRO was removed after the linkage was conducted. Thus, access to the participants addresses was not possible.

5. Results

This chapter presents and describes the results of the statistical analyses. Firstly, descriptive statistics of the population is presented in chapter 5.1, followed by the results of the main and additional analyses of short-term temperature effects on mortality in chapter 5.2. Finally, chapter 5.3 will present the results from the interaction analyses of income and education level on the association of short-term temperature exposure and mortality.

5.1 Descriptive statistics of population

Table 1 summarizes descriptive characteristics for HUBRO participants who died from CVDs and RDs during the study period, stratified by sex.

In total, 3527 HUBRO participants died from all causes of deaths (all-cause mortality) from 2000 to 2018, of which 1070 of these participants died from CVDs and 302 died from respiratory diseases, corresponding to respectively 30.3% and 8.6% of the all-cause mortality.

Table 1 Characteristics of HUBTO participants who died stratified by sex and cause of death[a]

	Sex			
	CVD (N=1070)		**RD** (N=302)	
	Female	**Male**	**Female**	**Male**
	n=533	n=537	n=158	n=144
Age at death				
Mean (SD)	83.8 (8.7)	79.7 (10.8)	83.5 (8.8)	81.7 (8.9)
Education				
Low level	221 (41.5)	161 (30.0)	75 (47.5)	55 (38.2)
Medium level	226 (42.4)	228 (42.5)	60 (38.0)	55 (38.2)
High level	86 (16.1)	146 (27.2)	22 (13.9)	34 (23.6)
Missing	0 (0)	2 ((0.3)	1 (0.6)	0 (0)
Income				
Mean (SD)	286 626 (169145)	418 900 (578622)	333118 (408870)	355043 (182233)
Low	306 (57.4)	161 (30.0)	90 (57.0)	48 (33.3)
High	169 (31.7)	299 (55.7)	56 (35.4)	83 (57.7)
Missing	58 (10.9)	77 (14.3)	12 (7.6)	13 (9.0)

[a] Data are presented as frequencies (column percentages) unless otherwise indicated.

The distribution of sex was even for CVD and respiratory mortality. For the CVD cases, there were 533 females (49.8%) and 537 males (50.2%), and for the respiratory cases there were 158 females (52.3%) and 144 males (47.7%). The mean age when the event of death occurred was 81.7 (SD: 9.9) for all CVD cases and 82.6 (SD: 8.9) for all respiratory cases. When

stratified by sex, females were generally a couple of years older when they died than males (Table 1),

For both CVD and RD there were most participants with low and medium level education. However, the proportion of males with high level education was higher than for females for both CVD and RD. The mean income was also much higher for males than females among the CVD cases, while it was similar for the RD cases. For both CVD and RD, there were higher proportions of males in the high-income category and highest proportion of females in the low-income category (table 1).

5.2 Association of short-term temperature exposure and CRD mortality

The exposure-response curves are illustrated in Figure 3, which is a result of the main analyses on short-term exposure and mortality, with a 10-day lag effect. The exposure-response relationship between short-term temperature exposure and mortality look non-linear, although the wide confidence intervals suggest linearity for CVD and respiratory mortality. The MMT for Oslo was around 15°C. The dashed lines in the exposure-response curves represent the cutoffs for what was considered extreme temperatures, which was the 1st percentile of temperature distribution for extreme cold and the 99th percentile for extreme heat.

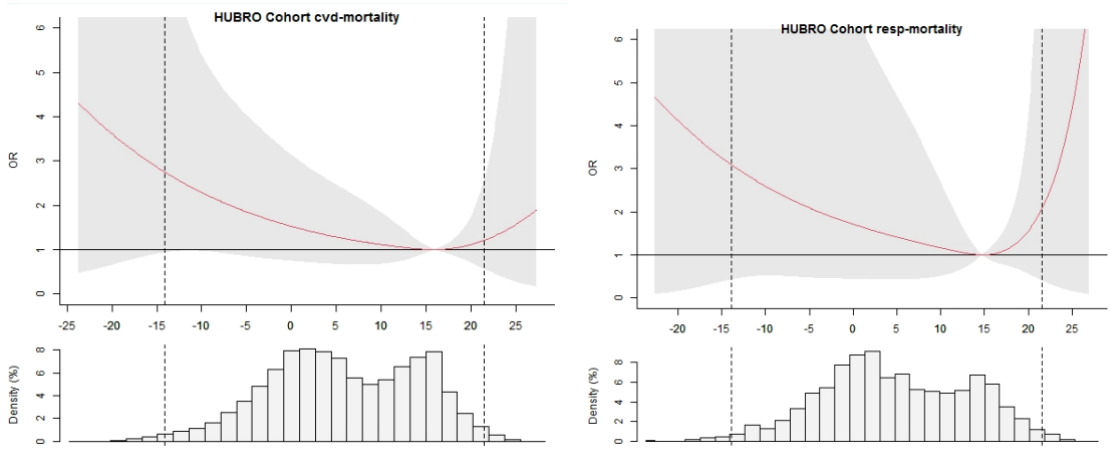

Figure 3. Exposure-response function between daily mean air temperature and mortality, and histograms with the air temperature distribution from 2000 to 2018 for cardiovascular and respiratory mortality. The dashed lines represent the cutoffs for extreme temperatures; 1st percentile for extreme cold and 99th percentile for extreme heat.

Table 2 summarizes the results from the main and additional analyses. It shows the OR of mortality for cold effect (decease in daily mean air temperature from the 25th to the 1st percentile of temperature distribution) and for heat effects (increase in daily mean air temperature from the 75th to the 99th percentile of temperature distribution). The table also shows the temperatures for the cold effects and heat effects for the main exposure (lag 0-10), and the extreme temperatures for the alternative lag-structures (99th percentile for lag 0-3, 1st percentile for lag 0-21) and seasons (99th percentile for warm, 1st percentile for cold). The median temperature (50th percentile of temperature distribution) for the study period was 5.3°C for CVD mortality and 4.11°C for respiratory mortality.

For the main analysis, the OR of CVD mortality was 1.79 (CI: 0.76-4.20) for cold effects and 1.17 (CI: 0.54-2.50) for heat effects, while the OR for respiratory mortality was 1.74 (CI: 0.36-8.29) for cold effects and 1.90 (CI: 0.38-9.51) for heat effects. Hence, the main analyses indicated that short-term exposure for heat effects seemed worse for respiratory mortality, while cold effects seemed to have similar mortality risk for both CVD and respiratory diseases, though the results did not reach statistical significance.

However, when I ran the cold and warm season separately in the additional analysis, it seemed that cold effects had a stronger adverse effect on CVD than on respiratory mortality, and the OR of respiratory mortality seemed to increase of cold effect with lag 0-21. In the section below I summarize the main results for CVD and respiratory mortality from the various stratums of the additional analysis. Only one of the findings reached statistical significance.

Lag days

I observed a significant increase in the OR of respiratory mortality for heat effects, when changing the lag structure to 0-3 days (OR: 4.92, CI: 1.48-16.4). The OR more than doubled compared to the main analysis. For the 21-day lag structure, there mainly was an increase in the OR of respiratory mortality from cold effects (OR: 11.42, CI: 0.94-139.56), however not significant.

Age

The OR for CVD mortality due to cold effects in the 65+ was 1.84 (CI: 0.75-4.50) and 1.72 (CI:0.68-4.40) for the 75+group. The OR for respiratory mortality in the 65+ group was 0.77

(CI: 0.15-3.94) and 1.02 (CI: 0.17-5.91) in the 75+ group. For heat effects however, the OR for CVD was 0.99 (CI: 0.66-1.47) in the 65+ group and 1.15 (CI: 0.51-2.60) in the 75+ group.

Sex

When stratifying by sex, it seemed that cold effects had the most impact on both CVD and respiratory mortality risk for males, while heat effects seemed to have the most impact for females (table 3). Particularly the OR for respiratory mortality differed greatly between the sexes, respectively ORs for cold effects of 3.12 (CI: 0.40-24.34) for males compared to 0.65 (CI: 0.06-7.42) for females, and ORs for heat effects; 0.36 (CI: 0.04-3.45) for males and 3.16 (CI: 0.20-51.08) for females. Additionally, for males the OR of CVD mortality was higher for cold effects (OR: 2.30, CI: 0.74-7.17) than heat effects (OR: 0.99, CI: 0.31-3.14), whereas the OR for CVD for females was approximately the same for both cold and heat effects.

Seasonality

When looking at the cold and warm season separately, it seemed that the warm season had more adverse effects on respiratory mortality, whereas the cold season seemed to have somewhat more effects on CVD mortality, however not significantly estimated. During the warm season (May-Sept) the ORs for heat effects were 5.90 (CI: 0.19-186.28) of respiratory mortality and 0.73 (CI: 0.18-2.95) for CVD mortality. During the cold period (Nov-March), the ORs for cold effects were 1.2 (CI: 0.10-13.62) for respiratory mortality and OR: 2.3 (CI: 0.66-8.01) for CVD mortality.

Table 2 ORs (95% CIs) and temperatures of mortality for a decrease in daily mean air temperatures from the 25th to the 1st percentile of temperature distribution (cold effects) and an increase in daily mean air temperatures from the 75th to the 99th percentile of temperature distribution (warm effects) in the main and additional analyses.

	Cold effect	**Temperature Decrease (°C)**	**Heat effect**	**Temperature Increase (°C)**
Main analysis				
Cardiovascular mortality	1.79 (0.76-4.20)	-0.17 - 14.1	1.17 (0.54-2.50)	13.02 - 21.5
Respiratory mortality	1.74 (0.36-8.29)	-0.1 - 14.0	1.90 (0.38-9.51)	11.81 - 21.6
Lag 0-3				
Cardiovascular mortality			1.15 (0.66-2.02)	21.51
Respiratory mortality			**4.92 (1.48-16.4)**	21.8
Lag 0-21				
Cardiovascular mortality	1.42 (0.37-5.55)	-14.0		
Respiratory mortality	11.46 (0.94-139.56)	-14.1		
Age 65+				
Cardiovascular mortality	1.84 (0.75-4.50)		0.99 (0.66-1.47)	
Respiratory mortality	0.77 (0.15-3.94)		2.1 (0.38-10.69)	
Age 75+				
Cardiovascular mortality	1.72 (0.68-4.40)		1.15 (0.51-2.60)	
Respiratory mortality	1.02 (0.17-5.91)		2.06 (0.35-12.01)	
Males				
Cardiovascular mortality	2.30 (0.74-7.17)		0.99 (0.31-3.14)	
Respiratory mortality	3.12 (0.4-24.34)		0.36 (0.04-3.45)	
Females				
Cardiovascular mortality	1.44 (0.42-4.93)		1.41 (0.51-3.88)	
Respiratory mortality	0.65 (0.06-7.42)		3.16 (0.2-51.08)	
Warm season (May-Sep)				
Cardiovascular mortality			0.73 (0.18-2.95)	22.9
Respiratory mortality			5.9 (0.19-186.28)	23.0
Cold season (Nov-March)				
Cardiovascular mortality	2.30 (0.66-8.01)	-15.9		
Respiratory mortality	1.20 (0.1-13.62)	-15.2		

Estimate(s) with $p < 0.05$ are marked as bold.

5.3 Results from interaction analyses

Table 3 summarizes the results of effect modification for income and education level for cold and heat effects for all-cause, - CVD and respiratory mortality, with a 10-day lag effect (main exposure). It shows the OR for each category of income and education level with p-value for the interaction between the groups. The section below presents the ORs of CVD and respiratory mortality for each category of income and education level compared to each other.

Income

The OR for CVD mortality from cold effects for the low-income group was 2.53 (CI: 0.76-8.39) and 1.74 (CI: 0.39-7.80) for the high-income group. For heat effects from CVD mortality, the OR was 1.09 (CI: 0.30-4.05) for the low-income group and 1.25 (CI: 0.43-3.66) for the high-income group. None of the results reached statistical significance.

For respiratory mortality, the OR for cold effect was 0.36 (CI: 0.03-4.40) for the low-income group and 2.0 (CI: 0.16-24.85) for the high-income group. The OR of heat effects was 2.97 (CI: 0.28-32.02) in the low-income group and 0.4 (0.02-6.35) in the high-income group. Once again, not statistically significant.

Education level

The OR of CVD mortality for heat effects was 2.34 (CI: 0.52-10.55) for the low-level education group, 1.2 (CI: 0.34-4.30) for the medium-level group and 0.63 (CI: 0.16-2.49) for the high-level education group. This might indicate a tendency towards a trend of stronger heat effects for the low-level to the high-level education group. However, the estimates were not statistically significant. For cold effects the OR for CVD mortality was 1.80 (CI: 0.47-7.01) for the low-level education group, 3.58 (CI: 1.02-12.61) for the medium-level group and 0.13 (CI:0.01-2.56) for the high-level education group. The OR for the medium-level group was statistically significant, though with a wide CI. The other ORs were not significant.

The OR of respiratory mortality for heat effects was 5.34 (CI: 0.31-91.85) for the low-level education group, 1.86 (CI:0.10-34.90) for the medium-level group and 1.37 (CI: 0.03-63.41) for the high education group. The ORs of respiratory mortality for cold effects was 1.63 (CI:

0.14-19.72) for the low-level education group, 2.01 (CI: 0.17-23.41) for the medium-level group and 0.02 (CI: 0.0-223.81) for the high-level group. These estimates were not significant, and the CI for the high-level education group was extremely wide, due to the very few respiratory mortality cases in the high-level education group.

Table 3 ORs (95% CIs) of mortality for a decrease in daily mean air temperatures from the 25th to the 1st percentile of temperature distribution (cold effects) and an increase in daily mean air temperatures from the 75th to the 99th percentile of temperature distribution (heat effects) modified by income and education level with p-value for the interaction between the level groups.

	Cardiovascular mortality		**Respiratory mortality**	
	Heat Effect	**Cold Effect**	**Heat Effect**	**Cold Effect**
Income				
Low	1.09 (0.30-4.05)	2.53 (0.76-8.39)	2.97 (0.28-32.02)	0.36 (0.03-4.40)
High	1.25 (0.43-3.66)	1.74 (0.39-7.80)	0.40 (0.02-6.35)	2.00 (0.16-24.85)
P-value for interaction	0,176		0.663	
Education Level				
Low	2.34 (0.52-10.55)	1.80 (0.47-7.01)	5.34 (0.31-91.85)	1.63 (0.14-19.72)
Medium	1.20 (0.34-4.30)	**3.58 (1.02-12.61)**	1.86 (0.10-34.90)	2.01 (0.17-23.41)
High	0.63 (0.16-2.49)	0.13 (0.01-2.56)	1.37 (0.03-63.41)	0.02 (0.0-223.81)
P-value for interaction	0.728		0.976	

Estimate(s) with p <0.05 are marked as bold.

6. Discussion

In this chapter I will discuss my findings in light of the theoretical background and existing literature. First, I will discuss how short-term temperature exposure is associated with CRD mortality in Oslo. Following this, I will discuss how socioeconomic conditions impact the temperature-mortality association from CVDs and respiratory diseases in Oslo. Lastly, follows a method discussion where I will discuss the methodological strengths and limitations of this study.

6.1 Associations of short-term temperature exposure and CRD mortality in Oslo

Summarized, the results from the main and additional analyses indicated that heat effects had a more adverse effect on respiratory mortality than cold effects although the CIs overlapped considerably in the main analysis. However, in the additional analyses with a 21-day lag structure, the OR of respiratory mortality for cold effects was very high. Furthermore, the OR of respiratory mortality for cold effects was also high for males, while it seemed that heat had an adverse effect on respiratory mortality in females. However, none of the results reached statistical significance, except the OR of respiratory mortality with the 3-day lag structure which increased of heat effects compared to lags 0-10.

The ORs of CVD mortality were generally higher for cold effects than heat effects both in the main and additional analyses, for males included. The OR of cold effects did not seem to increase with a 21-day lag structure compared to the main exposure, however the OR remained above 1. With lags 0-3, the OR for CVD mortality for heat effects remained approximately the same as in the main analyses. Neither of the results were statistically significant. The estimated CIs for both CVD and respiratory mortality were generally very wide, particularly for respiratory mortality.

Although the results did not reach statistical significance, the observed trends in the associations of heat effects on respiratory mortality are in line with other studies that have found increased risk of respiratory mortality when exposed to hot temperatures. Bunker et al. (2016) found in their systematic review that a 1°C temperature *rise* led to a 3.6% (95%CI: 3.18-4.02) increase in respiratory mortality, while a 1°C temperature *fall* led to a 2.9% (CI: 1.84-3.97) increase in respiratory mortality. Witt et al. (2015) found that the excess mortality risk

due to chronic lung disease from exposure to heatwaves was 1.8-2% higher per day than on days with average summer temperatures.

However, in a time-stratified case-crossover study from Denmark, Copenhagen they found an apparent modest effect of increasing temperature on respiratory mortality during the warm period (April-September), when also accounting for a 6-day lag effect (cumulative average of temperature), corresponding to 0.9% per 1°C increase (Wichmann et al., 2011). They also found a decrease in respiratory mortality risk during cold season (also with a 6-day lag effect), though, the results were not statistically significant. A study from Stockholm found that gradual temperature increases during summer were associated with mortality among people with COPD in the population younger than 65 years (Rocklöv et al., 2014).

A study from 2001 investigated temperature effects on all-cause, CVD and respiratory mortality in Oslo for the period 1990-1995 (Nafstad et al., 2001). The results only showed a statistically significant increase in daily mortality above 10°C for respiratory mortality, which increased by 4.7% per 1°C increase in the last seven days average temperature. Furthermore, at temperatures below 10°C, a 1°C fall in the last 7 days average temperature increased the daily mortality from CVDs by 1.7% and respiratory diseases by 2.1%. Hence, a temperature increase seemed to increase the relative risk of respiratory mortality more than a temperature decrease, in line with the trend of associations observed in this study. However, the study is over 20 years old. Though, it is, to the best of my knowledge, the only similar study investigating temperature effects on cause-specific mortality in Oslo using individual level data.

The significant OR of respiratory mortality from heat effects when including a 3 -day lag period may indicate delayed and cumulative effects of heat on respiratory mortality of the included participants. This is in line with findings from other studies, which have shown increased risk of heat effects on mortality when accounting for short-term delayed effects (Liu et al., 2015; Ragettli et al., 2023). Furthermore, the observed heat effects on respiratory mortality among females, are in line with previous findings from literature, of females having a higher risk of mortality related to hot temperatures (Gasparrini et al., 2022; Saucy et al., 2021; Son et al., 2019). The trend of higher OR of cold effects among males compared to females in this study (2.30, CI: 0.74-7.17 for CVD and 3.12, 0.4-24-34 for respiratory mortality), were also observed in the study of Rocklöv et al. (2014) who found higher death rates among males (over 65 years) compared to females when temperatures decreased (OR:1.019, 95%CI: 1.002-1.016 for a one

unit decrease of maximum temperature with lag 0-6). Literature suggests that differences in temperature-mortality associations among sex, can be related to differences in physiology, as well as exposure patterns and occupational exposure.

The trend in my associations indicating stronger heat effects on respiratory mortality might be related to the synergistic effect of air pollution and heat combined. Heat stress combined with increased ambient air pollution such can lead to acute as well as chronic damage to the lung tissue (Witt et al., 2015). The synergistic interaction of heat and air pollution can exacerbate respiratory conditions such as asthma and COPD. Chen et al. (2018) found an effect modification of air pollution and air temperature on mortality in their study of eight European urban areas, although they did not investigate respiratory mortality. The associations between air pollutants and mortality were generally stronger at high ambient temperatures compared to low temperatures (> 75^{th} percentile vs <25^{th} percentile of temperature distribution, respectively). On the other hand, the EXHAUSTUION-analyses with CONOR (including HUBRO) found no modifying effects of air pollution on the temperature-mortality association (EXHAUSTION project, 2022). However, there could potentially be other impacts of air pollution on the temperature-mortality relationship in Oslo, than observed in the CONOR analyses. Though, I did not include air pollution in my analyses and thus cannot know how this might have impacted the effects of temperature on respiratory mortality for only the HUBRO-participants.

The lag period of 21 days for cold effects also seemed to increase the risk of respiratory mortality, to 11.46 (0.94-139.56), although the CI is extremely wide. These findings might be related to influenza, which is more common during colder periods in Norway (Folkehelseinstituttet, 2022). Influenza can cause respiratory infections like pneumonia, which can be especially hazardous for elderly people and people with pre-existing diseases. Pneumonia was also one of the most common respiratory causes of death for the HUBRO cases during the study period. Bunker et al. (2016) found that the greatest risk of respiratory mortality was associated with cold-induced pneumonia (6.89%, CI: 20-12.99), as well as respiratory morbidity (4.93%, CI: 1.54-8.44). Considering this, one might have expected an increase in the OR of respiratory mortality during the cold season in the additional analyses. On the other hand, the analyses for winter used the main exposure of 0-10 lags day and did not account for the potential longer delay of cold effects, which might could have underestimated this association.

The lag-structures of 0-3 and 0-21 days did not seem to increase the OR of CVD mortality (OR: 1.15, CI: 0.66-2.02 for lag 0-3 and 1.42, 0.37-5.55 with lag 0-21) indicating no adverse delayed effects of heat or cold on CVD for the HUBRO participants. However, not statistically significant. The unobserved association of delayed effects of cold on CVD with a prolonged-lag period contrasts with the previous studies showing delayed effects of cold on CVD mortality (Liu et al., 2015). Furthermore, it is not in line with the trend of stronger effects of cold on CVD mortality in the main analyses and for cold season in this study.

CVD mortality has been associated with both cold and hot temperatures (Bunker et al., 2016). In a review from different parts of the world, including Sweden, Liu and his colleagues (2015) found that cold and hot temperatures, in general, were associated with a positive mean excess CVD mortality, although the relative risk (RR) were highly heterogenous across the included studies. Nevertheless, they underline that the detrimental effects of cold on CVD mortality were consistent.

Wichmann et al. (2011) found an inverse association of temperature exposure and CVD mortality for both the summer and winter period, although not significant for the winter period. They found that a moderate temperature increase indicated a protective effect on CVD mortality. During the warm period, they found a significant decrease in CVD mortality. Moreover, they found that the population over 80 years seemed to have a stronger protective effect of the temperature increase during summer.

In the study of Saucy et al. (2021) from Zurich, Switzerland however, they found an increased risk of CVD mortality mainly from heat with a significant OR of 1.28 (95% CI: 1.11-1.25) and the CVD mortality was higher during summer. They found a tendency of increased cold effects, however not significantly estimated (OR:1.15, CI: 0.95-1.39). In Stockholm, Sweden, increased temperatures during summer were associated with mortality among people older than 80 years, and for people with a previous MI for those under 65 years old (Rocklöv et al., 2014). Previous MI also increased the risk of mortality during winter among those older than 65. Another study from Sweden showed that heatwaves significantly increased both all-cause mortality and coronary heart disease-mortality (Åström et al., 2018)

The unobserved association of CVD mortality from heat effects in my analyses might be related to the few periods/days of extremely high temperatures in Oslo during the study period, as shown in the exposure-response curves in chapter 5.2. The daily median temperature during the study period was 5.31°C for the CVD cases, which was lower than the

MMT of around 15°C. Gasparrini et al. (2015) found that the attributable fraction for mortality due to cold was 7.29% (95% CI: 7.02-7.49), whereas it was 0.42% (CI: 0.39-0.44) for heat. They explain this difference by the high MMT percentile with most of the mean daily temperatures being lower than this. Furthermore, they found that most of the mortality was due to moderately cold and hot temperatures, and the contribution of extremely cold and hot days were comparatively low despite increased RR. These findings are supported by Masselot et al. (2023) who also found a higher attributable fraction for cold related mortality than heat related mortality across 854 cities in 30 European countries. Their study included Norwegian cities (including Oslo), with an estimated attributable fraction of 9.89% ((95% CI: 7.14-12.36) for cold and 0.25% (CI: 0.09-12.36) for heat.

Furthermore, Masselot et al. (2023) found great differences between ages, with increased cold, - and hot related mortality risk among older age groups, whereas younger ages had lower vulnerability to cold. They found that the mortality increase was generally more pronounced with increasing age for cold effects, including Oslo. Many other studies have also found an increased mortality risk among elderly, particularly for people over 75 years for both CVD and respiratory mortality (Bunker et al., 2016; Liu et al., 2015; Liu et al., 2022). My findings of mortality risk among the 65+ and 75+ however, did not show an indication of differences between the age groups, with a high degree of overlap in the CIs.

6.2 The role of socioeconomic conditions in the temperature-mortality association

The interaction analysis did not show any significant effect modification of income level and education level on the temperature-mortality relationship for the HUBRO-participants. The CIs were generally very wide with a high degree of overlap between most of the OR-estimates, in addition to high p-values. The estimates for CVD mortality for heat effects showed tendency to a trend of increasing mortality risk for low level of education. The OR was 2.34 (CI: 0.52-10.55) for the low-level education group, 1.2 (CI: 0.34-4.30) for the medium-level group and 0.63 (CI: 0.16-2.49) for the high-level education group, however with high uncertainty. There were similar tendencies for the respiratory mortality, however with extremely wide CIs.

Generally, the OR for the high-education level group was lower than the other groups and below 1, except for the OR of respiratory mortality for heat effects. Again, the CIs were very wide and overlapping, Although the results indicated no effect modification of income or

education level nor a clear trend of differences between the groups, there could still be differences in these associations for low and high SES groups in Oslo, but due to the very small population size of this study, I could not detect this. Particularly the respiratory cases were very few, with only 299 cases in the analytical sample. The small population size causes the wide CIs and large uncertainty in the effect estimates. This and other methodological strengths and limitations of this study will be further elaborated in the chapter 6.3 of method discussion.

Nevertheless, I will compare my results to other studies who have investigated effect modification of socioeconomic conditions in the temperature-mortality relationship. Furthermore, in light of these studies and the theoretical background of this thesis, I will try to explain some potential mechanisms for socioeconomic inequalities in the temperature-mortality relationship which may also exist in Oslo despite my findings.

As demonstrated in the background chapter, literature on effect modification of socioeconomic conditions on the temperature-mortality association show inconsistent evidence. Like my study, Saucy et al. (2021) also used three education level categories, where individuals with the low-level of education had increased OR of both cold (1.37, 95%CI: 1.0-1.92) and heat effects (1.49, 1.17-1.91) and for heat effects for the medium-level education group (1.25, 1.02-1.55). The low-level education group had the highest OR for both cold and heat effects, although the estimates were not statistically significant for the high-level group (0.97, 0.65-1.45 for heat, and 0.77, 0.45-1.32 for cold). The results from this study also showed lower ORs for the higher education level group (CVD: 0.63, 0.16-2.49 for heat, 0.13, 0.01-2.56 for cold, and RD: 1.37, 0.03-63.41 for heat, 0.02-223.81 for cold), however the CIs were very wide, particularly for the respiratory mortality.

Furthermore, they found gender-differential effect modification by education, particularly relevant for females. This study had a substantially larger population sample than my study, respectively 8830 cases in the low-level education group, 12 353 in the medium-level education group and 3150 in the high-level group, with narrower CIs than for my estimates. Also Huang et al. (2015) used three education categories and found the strongest temperature effects for the low-level education group, however only for heat effects, not cold effects.

On the other hand, the review from Son et al. (2019) also identified studies that found no difference or higher mortality risks for those with higher educational attainments, in line with the findings from Yu et al. (2010). Socioeconomic factors can be correlated with each other,

and one specific indicator do not fully represent actual SES. SES can relate to a variety of factors such as multiple sources of income, family income and historical income (Son et al., 2019). Furthermore, both individual and community SES can impact health. Although Son et al. (2019) found limited or suggestive evidence for effect modification of community-level SES, the studies of and Madrigano et al. (2013), Zanobetti et al. (2013) and Gasparrini et al. (2022) indicated higher mortality risks for people living in lower SES neighborhoods.

This is supported by the findings of Åström et al. (2018). In their study of effect modification from neighborhood deprivation in Sweden, they used neighborhood deprivation index (NDI) as a measure. This is a summary measure of four variables that indicate deprivation: the proportion of inhabitants with low educational status, low income, unemployment and social welfare recipients. They found that that neighborhood deprivation may modify the mortality risk due to CHD during heat waves in Sweden (not for all-cause mortality). On a national level, their findings showed a significantly higher RR of mortality for the most deprived neighborhoods; 1.32 (CI:1.17-1.48) compared to the less deprived neighborhoods; 1.10 (CI: 1.01-1.20) and the least deprived neighborhoods; 1.02 (0.91-1.15). This is in line with the previous findings from Wichmann et al. (2011), who found a stronger association of temperature increase and CVD mortality in Copenhagen among lower SES groups, where the neighborhood SES was based on household income, educational and employment status.

These findings can be seen in light of Link and Phelans fundamental causes theory. Flexible resources influence access to the broader contexts such as the neighborhood one resides in, with various associated risk- and protective factors. Phelan et al. (2010) exemplify this with the possibility people of high SES have to live in affluent neighborhoods with other people of high SES. Moreover, where the neighbors jointly make an effort to reduce health risks such as traffic and air pollution, as well as having health promoting amenities like green and blue spaces and playgrounds nearby. One of the possible explanations of the long-lasting divide between the wealthier western districts of Oslo and the less wealthy eastern districts is that people of higher SES tend to be attracted to neighborhoods with other people of high SES (Elstad, 2017).

In many urban areas, people of lower SES tend to live in more deprived neighborhoods with less vegetation and access to green and blue spaces (European Environment Agency, 2018; World Health Organization, 2019). Green and blue spaces, such as parks, urban forests, street trees, green roofs/walls on buildings, lakes and ponds, can reduce heat by providing cooling

through shading and evapotranspiration, and are considered to be important climate adaptation measures in urban areas (Dodman et al., 2022).

Additionally, lower SES groups tend to live in more denser parts of the city, where the UHI effects is often more intense (European Environment Agency, 2018; Reinwald et al., 2021). They can therefore have higher exposure to hot temperatures. In a study of spatial patterns of heat related CVD mortality in the Czech Republic urban districts with low and high SES index were compared. They found that below a certain threshold, SES had a relevant impact on excess CVD mortality, however not when all districts were considered together (including rural areas) (Urban et al., 2016). The findings of the significant relationship between decreased SES and increased heat related mortality in the most deprived urban districts actually occurred when they accounted for the cumulative lagged effects of heat, which were rather small in the high SES urban groups compared to the lower SES urban groups. The increase in excess mortality in the lower SES urban districts was 26.3% compared to 8.4% in the higher SES urban districts. This underlines the importance of accounting for potential delayed effects of temperatures on mortality.

Furthermore, lower SES populations are also more likely to live in areas nearby main roads and industrial activity where concentrations of air pollution are higher, due to cheaper rents in these areas (Paavola, 2017). The eastern districts of Oslo, with a higher proportion of lower SES groups, are shown to be disproportionately exposed to hazardous air pollution levels, compared to the higher SES western districts (Venter et al., 2023).

On the other hand, in some areas higher SES populations live in more central parts of the city and can thus be more exposed to UHIs (European Environment Agency, 2018). According to the recent study of Venter et al (2023), the portion of people in Oslo that are socioeconomically and environmentally disadvantaged are concentrated in the inner and outer eastern regions of the city. The proportion of people with a high educational attainment (university or college degree) in Oslo are similar in inner city/central districts like Sagene and Grunerløkka as the western districts such as Ullern, Frogner and Vestre Aker (Statistisk sentralbyrå, 2022b).

Furthermore, Oslo has around 800 parks and approximately 98% of the population lives less than 300 meters from a green area (Oslo kommune, 2020b). Nevertheless, Venter et al (2023) found that lower SES districts in Oslo had less available blue-green spaces than the higher

SES districts. However, the authors did not find increased exposure to heat among the lower SES districts, and income was not spatially correlated with heat hazard.

It is not just what area or neighborhood people live in that impacts how they are affected by temperatures and the ability to adapt. Housing conditions can also influence health impacts, such as mortality. According to World Health Organization (2019), housing conditions can have a direct impact on heat related mortality, particularly during extreme temperatures. Houses can offer protection against hazardous temperature exposures, or it can be responsible for increased exposure.

Inequalities in housing quality are often related to socioeconomic conditions associated with income (World Health Organization, 2019). People of lower SES can often live in housing segments of lower quality, due to lower financial resources (World Health Organization, 2019). This can be associated with less adequate living conditions in terms of building quality, thermal efficiency and equipment and amenities, such as air conditions and heating facilities. Geometry of buildings and how they are built and designed impact exposure to heat, with top floor apartments experiencing more thermal stress than ground floor apartments. Impacts of extreme temperatures increases when buildings do not cool down during the night but keep accumulating heat. Hoses of lower quality can overheat at lower temperatures (The Lancet, 2018).

According to a report on environmental inequalities in Europe, populations with lower household income in urban areas had greater difficulties with keeping the household cool during summer and adequately warm during winter (World Health Organization, 2019) An adequately warm household with temperatures above at least 18°C is essential to life expectancy and reduces excess winter mortality and health risks associated with CRDs, particularly for elderly. Many of the household struggling to adapt to low temperatures, are in the lowest income quantile. The highest inequality ratios by income related to keeping the household warm, was found in several countries including Norway, reporting income-related inequalities beyond a ratio of 6:1 (World Health Organization, 2019). Inability to keep the household adequately cool during periods of higher temperatures, were also found in Norway. Furthermore, increasing living costs are reinforcing health, social and economic inequalities in Oslo (Meld. St. 15 (2022-2023)). Many Norwegian residents today are experiencing the burden of increased electricity costs, which again hits lower SES groups extra harder. This also increases the number of households that are struggling. In the study of

(Rocklöv et al., 2014) however, it seemed that the wealthier populations in Stockholm had a higher risk of cold related mortality, which the authors related to wealthier municipalities might being characterized to a larger extent by older, less energy efficient houses.

In Oslo, the less wealthy residents tend to live in high rise apartments to the east and north, which might increase exposure to heat (Venter et al., 2023). Lower financial resources and a scarce housing market in Oslo, may lead to lower SES populations living in poorer quality houses and potentially a more deprived neighborhood. In Oslo the differences in people who own or rent their home follows the social gradient, with fewer people of lower SES owning their own home (Elstad, 2017). When owning the home, one may have more control and power over the housing situation and can thereby have better chances to adapt to climate changes, like temperature deviations. Seen as the temperatures are projected to increase further in Norway, although with high uncertainty of how much, it will be important with capacity to adapt to such temperature changes, particularly in an urban are like Oslo. Although, cold temperatures may remain the greatest mortality burden in Norway (Masselot et al., 2023).

The temperature-mortality relationship is also impacted by the individual's susceptibility to temperature exposure (Cissé et al., 2022). Poorer health status among lower SES groups in Oslo, with higher prevalence of pre-existing CRDs, can make these populations more susceptible to heat and cold effects to begin with. When lower SES groups additionally may be more exposed to ambient temperature deviations and less capable of adapting, this can enhance their vulnerability to climate change health impacts. Furthermore, it can increase existing health inequalities, contrary to the "Leaving No one behind"-principle.

Health literacy may also impact temperature related health effects (Paavola, 2017). For instance, educational status may influence ability to translate health education (e.g., information about temperature exposure and health impacts) into action. Again, this demonstrates the impacts of flexible resources, such as knowledge and power. In light of this, the higher health literacy among higher SES populations in Oslo may lead them to taking better action to protect themselves against harmful temperature impacts, and thus strengthen their adaptive capacity.

However, it is important to underline that I have not examined these underlying mechanisms myself. The unobserved effect modification of education and income level in this study may

be related to the small size of the population sample. In the next section follows a discussion of methodological considerations, strengths and weaknesses of this study.

6.3 Methodological considerations, strengths and limitations

In all research it is important to ensure quality inn all parts of the process. A study's quality relies on its *reliability*, *internal* and *external validity (Ringdal, 2014).* In the following section I will explain these terms and evaluate the quality of this study by discussing its reliability, internal and external validity.

6.3.1 Reliability

Reliability refers to whether the results of the study are reliable; if it is possible to replicate the study and verify the results (Ringdal, 2014). High reliability is a prerequisite for high validity.

Transparency is an important principle to ensure reliability (Ringdal, 2014). I have strived to ensure transparency in the thesis by providing a step-by-step description and explanation of the choices and considerations throughout the process, as best I could given that the data collection and linkage was done beforehand by others. Furthermore, I have attached the R-scripts of the codes used in the analyses to make it easier to replicate exactly what was done in the analyses (Appendix 2, 3 and 4).

For the results to be reliable, this requires that the methods used to investigate the research questions, are feasible for the study aim (Jaakkola, 2003p. 81). In this study I used the time-stratified case-crossover design which is a feasible design for studying short-term effects of transient exposures like temperature. This study design has been used in many similar studies investigating temperature effects on mortality, of which many have also used similar methodological approaches as in this study, including the conditional logistic regression and DLNM (Åström et al., 2018; Huang et al., 2015; Saucy et al., 2021; Wichmann et al., 2011).

Furthermore, the data used in this study is based on high-resolution temperature data, objective registry data and detailed cohort data. However, there may be potential biases that can have impacted the findings of this study, i.e. impacting the validity of the study.

6.3.2 Internal validity

Internal validity is defined as *«the extent to which the results of a study reflect the true situation in the study sample"* (Webb et al., 2017, p.251). For a study to have high internal validity, this implies that what one wants to investigate aligns with what one actually investigates (Ringdal, 2014).

To evaluate if the results of a study is really valid for the study population, one has to consider potential alternative explanations for the results, such as *bias* (Webb et al., 2017). Overall, there are usually three types of bias that may occur in epidemiological studies, and these are *confounding*, *selection bias* and *information bias.*

6.3.2.1 Selection bias

Selection bias refers to the selection of participants in the study, and it occurs when the participants included in the study sample are not *representative* for the for the target population, i.e., the population one wants to learn about and the population which the sample is supposed to represent (Webb et al., 2017).

There may be a risk of selection bias if there are systematic differences between those who participate in a study and those who do not. For instance, if less advantageous and stigmatized groups are less represented in the study sample, or there is an overrepresentation of people with higher SES. On a general basis, participants in baseline examinations of cohort studies tend to be females of high SES with healthier lifestyles in comparison to non-participants (Enzenbach et al., 2019). As my study is based on the HUBRO cohort, the possibility of selection bias cannot be eliminated. Although, the distribution of sex was even in the sample population of this study, and there were fewer participants with high level of education. Furthermore, in an article about CONOR, which HUBRO is a part of, it is stated that investigations of variables such as age, sex, education, smoking habits and alcohol use indicate that CONOR corresponds to the general population in Norway (unpublished material according to the article) (Aamodt et al., 2010). However, variables such as e.g., smoking habits and physical activity may be affected by under- or overreporting, a potential measurement error or information bias.

6.3.2.2 Information bias

Measurement error or *information bias* means that information is being misclassified, due to random or systematic errors, which can impact the effect estimate (Webb et al., 2017).

Differential misclassification impacts the study groups unequally, while *non-differential misclassification* affects the study groups study groups equally.

When information is based on self-reporting, though e.g., questionnaires, this could lead to underreporting of information about e.g., alcohol-use, smoking habits or overreporting about e.g., physical activity (Aamodt et al., 2010). For self-reported data, *recall bias* must be considered, i.e., if there are systematic differences in under -, or overestimating exposures between the comparing groups (Webb et al., 2017). Such information bias may produce under-, or overestimates of the associations. In the HUBRO questionnaire some of the questions may be affected by under-, or overreporting, such as smoking and alcohol use. However, I have not included such variables in my analysis. Furthermore, in my study the cases are their own controls, and the exposure data is objective and not based on self-reporting, thereby mitigating the possibility of recall bias. Through, the education level variable in my study was mostly based on the HUBRO questionnaire (with missing information replaced by information from SSB) and some participants of this study might have been misclassified into the wrong education category due to measurement error. If so, this might have impacted the observed association of effect modification (of education).

A strength of using education as a measure of SES is that education level remains relatively stable after 30 years of age when the majority has completed their education (Syse et al., 2022). This was suitable for my study populations as most of the participants were pensioners at the time of death. Moreover, there were only 0.2% missing information on education level for the CVD cases and 0.3% for the RD cases, which probably have not impacted the observed association.

For income however, the proportion of missing information was 12.6% for the CVD cases and 8.3% for the respiratory cases, which are considered high proportions that probably impact the effect estimation. It may over- or underestimate the association of effectmodification. Moreover, income can, to a greater extent than education, change over time, can be defined in many ways and not all income is reported to the tax authorities (Syse et al., 2022). In this study, the income variable used was the household income from the same year as death occurred, and the income level could have changed from e.g., the prior year. However, this is less likely for this study population as the majority were pensioners. Furthermore, income level does not capture potential fortunes/wealth of participants. One could have a low income but have a high fortune, in which case the income level would not

be an accurate measure of financial resources. For instance, maybe some of the participants in this study had saved up a private pension, perhaps in form of stocks or funds, which is not counted as their income. This can be considered a weakness of using income to represent SES. However, I still considered it to be a relevant variable to include in the analyses, together with education level to strengthen the association of SES in the temperature-mortality relationship.

The use of registry-based data mitigates the issue of information bias (Webb et al., 2017). The Cause of Death Registry provides precise information on cause of death, although some non-differential misclassification is possible. Some challenges in using mortality data related to cause-specific death may be related to coding of the actual cause of death. It may not be straightforward to evaluate the actual cause of death if the person has many coexisting diseases, which is the case for many elderlies. The average age at death among the participants of this study was 83.7 (9.69) for the CVD cases and 80.13 (10.37) for the respiratory cases, and it is likely that some of the participants had several diseases. However, approximately 39% of all the participants who died during the study period had a CVD or RD as cause of death. Thus, it may be unlikely with many cases of misclassification of the cause of death in this study population.

Even though the temperature data is based on high-resolution data, it may be subject to measurement errors as it is based on data from measuring stations and calculations. Furthermore, the temperature exposure is estimated for the participant's residential addresses (updated annually) and temperature data did not include exposures while away from home. This could lead to some non-differential misclassification of exposure, which could produce an underestimation or overestimation of temperature effects (Saucy et al., 2021). However, elderly people tend to spend more time home. On the other hand, with elderly people spending more time indoor could also affect the cold related mortality for the lowest temperatures, but probably not at higher temperatures (Nafstad et al., 2001). This will probably apply to most elderly people, but perhaps to a greater extent to those who have very poor health and have a high risk of dying, which can be the case for elderly with existing CRDs. If more people with CRDs spend more time home and are thus less exposed to cold temperatures than those without pe-existing CRDs, this maybe lead to differences in the estimated ORs among the groups. If this to a larger extent is the case for lower SES groups, who often hive a higher prevalence of CRDs, this may lead to lower estimated ORs among the low SEs groups and potentially impact the association of effect modification.

6.3.2.3. Confounding

Confounding occurs when the investigated association is confused by the effect of other factors (Webb et al., 2017). It refers to a mixing of effects, and can occur when the comparing groups are not completely exchangeable and thus differ by other factors than their exposure. If such factors are the cause of the outcome and is associated with the exposure, then parts of or all of the observed association between the exposure and outcome might be due to these factors. In the case-crossover deign the participants are "matched" with themselves and serves as their own controls, thereby reducing potential confounding of individual characteristics such as sex, age and health status, The minimizing confounding of individual characteristics in the case-crossover design, also strengthens effect modification analyses to identify vulnerable individuals (Saucy et al., 2021).

However, there may be other factors that could have confounded the observed associations in this study. Air pollution may impact the temperature-mortality association. The study of Rocklöv et al. (2014) from Sweden would include air pollution as a confounder if it changed the estimated relative effect by more than 10%, however it did not and was thus not included. Other studies have included air pollution as effect modifiers in their analyses of the temperature-mortality association, including the CONOR-analyses in EXHAUSTION (Chen et al., 2018). However, the CONOR analyses did not find any significant effect modification of air pollution (EXHAUSTION project, 2022). On the other hand, these results could potentially be different for Oslo, and air pollution may be a confounder of the observed temperature-mortality association.

Other meteorological parameters such as humidity and wind were not included in the analyses, nor did I not have data on these variables. These are also potential confounders of the temperature-mortality relationship (Nafstad et al., 2001). Furthermore, influenza may also impact temperature related mortality during wintertime. However, it is not likely that influenza impacts temperature, but rather cold temperatures that leads to influenza. In that way, influenza could potentially mediate the temperature-mortality association.

Even though one controls for potential confounders in the analyses, there will usually be some potential confounding variables that one does not know of or does not have the possibility to control for, known as *residual confounding* (Webb et al., 2017).

6.3.3. External validity

External validity refers to the study's *generalizability;* whether the results can be generalized and apply for other populations than the sample population (Webb et al., 2017). The

Although education is a relatively stable SES indicator at an individual level, the pattern of education in Oslo has changed over time. Since 1990 to 2018 the proportion of people with a higher educational attainment of more than 4 years has increased (Oslo kommune, 2020b). In 2000 and 2001 when the HUBRO questionnaire was conducted, the proportion of people with higher educational attainment of more than four year was greater among males, while this trend has turned and now there is a greater proportion of females with a high educational attainment. The education level category with the highest proportion in Oslo today is the high level education category (university or college) with 54.1% in 2022, 26.8% with high school or vocational school and 19.1% with primary school as their highest educational level (Statistisk sentralbyrå, 2022b). In my population sample there were approximately 35.7% and 43.1 % in the low education category, 42.4% and 38.1 % in the medium level category and 21.7% and 18.5% in the high level category for CVD and respiratory cases, respectively. Hence, the educational pattern of the study population is not the same as the educational pattern of Oslo's residents today, which also makes the sample population less representative for Oslo populations today.

Furthermore, as I have investigated the short-term effects of temperature exposure on mortality with a month prior to death as the hazard/exposure period and the average age at death was 81.7 (SD: 9.9) and 82.2 (SD: 8.9) for CVD and respiratory cases respectively, this means that this study is mostly representative for the elderly population of Oslo. Older age is considered a risk factor for the temperature-mortality relationship, which could have contributed to the observed ORs. It thus makes it harder to generalize the findings to younger populations. Moreover, the small population size makes the results of this study less generalizable beyond the study population, because of the high uncertainty in the effect estimates. Hence, it weakens the external validity of the study.

6.3.4 Strengths and limitations

A strength of this study is that it is based on high resolution and detailed data. Furthermore, that it included the DLNM model with various lag-structures. This could potentially produce more accurate estimates for the temperature-mortality information, when accounting for potential delayed effects of cold and hot temperatures.

However, a limitation of the study is the small population, particularly the cases of respiratory mortality and the various groups of education, - and income level for both CVD and respiratory mortality ranging from 25 to 454. The estimated CIs were generally very wide, indicating poor precision of the effect estimation and high uncertainty (Webb et al., 2017). Additionally, there were high p-values in the interaction analysis, which also indicates a higher probability that the observed association of effect modification could be due to chance.

A small population size can weaken the statistical power of a study, i.e., if "*a study has enough power to detect a true association with sufficient precision"* (Webb et al., 2017, p.176). It might be that there truly are associations of temperature exposure, CRD mortality and socioeconomic effect modification among the HUBRO-participants, but the study was too small to detect this with any certainty, i.e., weak statistical power. Although my findings were not statistically significant, the trend in the associations of heat effects on respiratory mortality and increased OR of respiratory mortality among females compared to males, is to a large extent consistent with existing literature, which might give some strength to these observed associations potentially being real and not only due to chance (Bunker et al., 2016).

The use of both income level and education level to represent SES can be a more accurate indicator for real SES than using just one variable. The use of three education level categories will probably capture more variation and nuances than using fewer categories. However, due to the small population size of this study it might have been more suitable to use only two categories, to have more participants in each education category for each cause of death-stratum. Particularly, considering effect modification requires a lot of statistical power to detect a statistically significant effect, often due to the lower number of participants in the various stratums (Webb et al., 2017).

In a high-income city with considerable income inequalities like Oslo, it could be more relevant to use more than two income categories to capture more of the income variation. However, due to the few cases of CVD and respiratory mortality in HUBRO, this was not suitable to do in this study. Furthermore, stratifying by seasonality could provide more useful information about the effects of temperature on health for a multi-season country like Norway. In this study it was considered relevant to see if the estimates differed by summer and winter (May-Sept and Nov-March). This could be a strength of the study. On the other hand, the season stratums for each cause of death were very small, leading to even wider CIs.

Particularly the CI for respiratory mortality during the cold season was extremely wide, which could be a limitation of stratifying by seasonality with the small population size. Moreover, the analyses of the season stratums included the main exposure (0-10 lag days), which could underestimate the cold effects during cold season and heat effects during the warm season. With a larger population sample, it could have been relevant to stratify by four seasons to potentially capture more of the varying effects of temperature on CRD mortality.

7. Conclusion and implications for public health work

In this study I did not find a statistically significant association of temperature and mortality, except for respiratory mortality when accounting for a 3-day lag period. Though I did find some trends indicating more adverse heat effects for respiratory mortality including females, while males seemed to be more impacted by cold effects. Furthermore, I did not find evidence of higher CVD or respiratory mortality risk related to high and low temperatures among lower SES groups compared to higher SES groups. This may be related to the small population sample.

Although not detected in this study, there could be an association of socioeconomic conditions in the temperature-mortality relationship in Oslo, particularly considering the existing socioeconomic health inequalities. Socioeconomic conditions impact the greater context of people's lives. Although I did not investigate underlying mechanisms, it is possible that lower SES populations in Oslo may have less access to flexible resources such as knowledge, money and power which may impact their level of exposure to temperature, their adaptive capacity and susceptibility. This might lead to a predisposition of being adversely affected by climate change circumstances, such as increased CRD mortality associated with temperature exposure, potentially making them more vulnerable to these impacts.

Therefore, despite the lack of evidence in this study with the small population sample, I still argue that socioeconomic conditions and potential unequal impacts of temperature exposure and other climate change consequences, should be considered when developing and implementing equitable climate adaptation strategies in Oslo. This is essential to achieve the 2030 UN SDGs and one of the main goals of Norwegian public health policy of reducing social health inequalities, in line with the "leave no one behind-principle" and the proportional universalism-strategy. Socioeconomic conditions should be considered in all sectors in the urban planning of Oslo. This may include housing policy and development, the use of green areas and providing of information and knowledge to the population.

However, this topic needs further investigation. Further research should be conducted on climate related temperature impacts and social health inequalities in Oslo, with a larger population sample. It would be interesting to investigate the association of temperature exposure and housing quality, if one owns or rents the household, and lives alone or not. Furthermore, studies including air pollution in the temperature-mortality relationship could also be relevant for Oslo.

References

Aamaas, B., Aaheim, A. Alnes, K. van Oort, B. Dannevig, H. Hønsi, T. (2019). *Oppdatering av kunnskap om konsekvenser av klimaendringer i Norge*. Report 2018:14. Available at: https://www.miljodirektoratet.no/globalassets/publikasjoner/m1209/m1209.pdf (accessed: 10.10.2022).

Aamodt, G., Søgaard, A. J., Næss, Ø., Beckstrøm, A. C. & Samuelsen, S. O. (2010). CONOR-databasen - et lite stykke Norge. *Tidsskriftet Den Norske Legeforening*, 130: 264 (3): 264-265. doi: doi: 10.4045/tidsskr.09.0369.

Ariansen, I., Selmer, R., Graff-Iversen, S., Egeland, G. M., Sakshaug, S. & Løchen, M.-L. (2020). *Cardiovascular disease in Norway*: National Institute of Public Health. Available at: https://www.fhi.no/nettpub/hin/ikke-smittsomme/Hjerte-kar/ (accessed: 30/08).

Åström, D. O., Tornevi, A., Ebi, K. L., Rocklöv, J. & Forsberg, B. (2016). Evolution of Minimum Mortality Temperature in Stockholm, Sweden, 1901-2009. *Environ Health Perspect*, 124 (6): 740-4. doi: 10.1289/ehp.1509692.

Åström, D. O., Åström, C., Forsberg, B., Vicedo-Cabrera, A. M., Gasparrini, A., Oudin, A. & Sundquist, K. (2018). Heat wave–related mortality in Sweden: A case-crossover study investigating effect modification by neighbourhood deprivation. *Scandinavian Journal of Public Health*, 48 (4): 428-435. doi: 10.1177/1403494818801615.

Bednar-Friedl, B., Biesbroek, R., Schmidt, D. N., Alexander, P., Børsheim, K. Y., Carnicer, J., Georgopoulou, E., Haasnoot, M., Le Cozannet, P. Lionello, G., et al. (2022). Europe. In Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al. (eds) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 1817–1927: Cambridge University Press.

Bunker, A., Wildenhain, J., Vandenbergh, A., Henschke, N., Rocklöv, J., Hajat, S. & Sauerborn, R. (2016). Effects of Air Temperature on Climate-Sensitive Mortality and Morbidity Outcomes in the Elderly; a Systematic Review and Meta-analysis of Epidemiological Evidence. *EBioMedicine*, 6: 258-268. doi: 10.1016/j.ebiom.2016.02.034.

Chen, K., Wolf, K., Breitner, S., Gasparrini, A., Stafoggia, M., Samoli, E., Andersen, Z. J., Bero-Bedada, G., Bellander, T., Hennig, F., et al. (2018). Two-way effect modifications of air pollution and air temperature on total natural and cardiovascular mortality in eight European urban areas. *Environ Int*, 116: 186-196. doi: 10.1016/j.envint.2018.04.021.

Cissé, G., R., McLeman, H., Adams, P., Aldunce, K., Bowen, D., Campbell-Lendrum, S., Clayton, K. L., Ebi, J., Hess, C., Huang, Q., et al. (2022). Health, Wellbeing, and the Changing Structure of Communities. In Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al. (eds) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.*, pp. 1041–1170: Cambridge University Press.

Copernicus Climate Change Service. (2023). *European state of the climate 2022. European heat, widespread drought tyoify European climate in 2022.* . Available at: https://climate.copernicus.eu/extreme-heat-widespread-drought-typify-european-climate-2022 (accessed: 11.04.2023).

Dahlgren, G. & Whitehead, M. (1991). Policies and strategies to promote social equity in health. Background document to WHO – Strategy paper for Europe. *Institute for Futures Studies*: 4-54.

Dahll, E., Bergsli, H. & van der Wel, K. A. (2014). *Sosial ulikhet i helse: En norsk kunnskapsoversikt* Available at: https://oda.oslomet.no/oda-xmlui/bitstream/handle/20.500.12199/738/Sosial%20ulikhet%20i%20helse%20En%20norsk%20kunnskapsoversikt.%20Hovedrapport.pdf?sequence=6&isAllowed=y (accessed: 20.01.2023).

Dodman, D., B. , Hayward, M., Pelling, V., Castan Broto, W., Chow, E., Chu, R., Dawson, L., Khirfan, T., McPhearson, A., Prakash, Y., et al. (2022). Cities, Settlements and Key Infrastructure. . In Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al. (eds) *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 907–1040: Cambridge University Press,.

Elstad, J. I. (2017). Helseulikhetenes by. In Ljunggren, J. (ed.) *OSLO ulikhetenes by*, pp. 171-187. Oslo: CAPPELEN DAMM AS.

Enzenbach, C., Wicklein, B., Wirkner, K. & Loeffler, M. (2019). Evaluating selection bias in a population-based cohort study with low baseline participation: the LIFE-Adult-Study. *BMC Medical Research Methodology*, 19 (1): 135. doi: 10.1186/s12874-019-0779-8.

Eriksen, S., Indeberg, Tor Håkon., O´Brien, Karen., Sygna, Linda. (2015). Develeopment is not enough In Indeberg, T. H. (ed.) *Climate change adaptation and develeopment* pp. 1-15: Taylor and Francis

European Environment Agency. (2018). *Unequal exposure and unequal impacts: social vulnerability to air pollution, noise and extreme temperature in Europe* 22/2018. Luxenburg: European Environment Agency.

EXHAUSTION. (2020). *About the project* (accessed: 15.09.2022).

EXHAUSTION project. (2022). *Short-term effects of cardiopulmonary mortality in Norway*. Unpublished manuscript.

Folkehelseinstituttet. (2019). *Helseundersøkelsen i Oslo - HUBRO (2000-2001)*. Available at: https://www.fhi.no/div/helseundersokelser/landsomfattende-helseundersokelser-lhu/helseundersokelser/hubro---helseundersokelsen-i-oslo-i/ (accessed: 28.09.2022).

Folkehelseinstituttet. (2020). *Om Dødsårsaksregisteret*. Available at: https://www.fhi.no/hn/helseregistre-og-registre/dodsarsaksregisteret/dodsarsaksregisteret2/ (accessed: 14.03.23).

Folkehelseinstituttet. (2022). *Influenza - Fact sheet about seasonal influenza*. Available at: https://www.fhi.no/en/id/influensa/seasonal-influenza/influenza-advice/influenza---fact-sheet-about-season/#influenza-in-norway (accessed: 04.05.2023).

folkehelseloven. (2011). *Lov om folkehelsearbeid (LOV-2011-06-24-29)*. Available at: https://lovdata.no/dokument/NL/lov/2011-06-24-29

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklov, J., Forsberg, B., et al. (2015). Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet*, 386 (9991): 369-75. doi: 10.1016/S0140-6736(14)62114-0.

Gasparrini, A., Guo, Y., Sera, F., Vicedo-Cabrera, A. M., Huber, V., Tong, S., de Sousa Zanotti Stagliorio Coelho, M., Nascimento Saldiva, P. H., Lavigne, E., Matus Correa, P., et al. (2017). Projections of temperature-related excess mortality under climate change

scenarios. *Lancet Planet Health*, 1 (9): e360-e367. doi: 10.1016/s2542-5196(17)30156-0.
Gasparrini, A., Masselot, P., Scortichini, M., Schneider, R., Mistry, M. N., Sera, F., Macintyre, H. L., Phalkey, R. & Vicedo-Cabrera, A. M. (2022). Small-area assessment of temperature-related mortality risks in England and Wales: a case time series analysis. *The Lancet Planetary Health*, 6 (7): e557-e564. doi: 10.1016/s2542-5196(22)00138-3.
Hajat, S., Vardoulakis, S., Heaviside, C. & Eggen, B. (2014). Climate change effects on human health: projections of temperature-related mortality for the UK during the 2020s, 2050s and 2080s. *J Epidemiol Community Health*, 68 (7): 641-8. doi: 10.1136/jech-2013-202449.
Hanssen-Bauer, I., Førland, E. J., Haddeand, I., Hisdal, H., Mayer, S., Nesje, A., Nilsen, J. E. Ø., Sandven, S., Sandø, A. B., A., S., et al. (2015). *Klima i norge 2100*. NCCS report no. 2/2015. Available at: https://klimaservicesenter.no/kss/rapporter/kin2100 (accessed: 10.10.2022).
Heaviside, C., Macintyre, H. & Vardoulakis, S. (2017). The Urban Heat Island: Implications for Health in a Changing Environment. *Curr Environ Health Rep*, 4 (3): 296-305. doi: 10.1007/s40572-017-0150-3.
Huang, Z., Lin, H., Liu, Y., Zhou, M., Liu, T., Xiao, J., Zeng, W., Li, X., Zhang, Y., Ebi, K. L., et al. (2015). Individual-level and community-level effect modifiers of the temperature-mortality relationship in 66 Chinese communities. *BMJ Open*, 5 (9): e009172. doi: 10.1136/bmjopen-2015-009172.
Jaakkola, J. J. (2003). Case-crossover design in air pollution epidemiology. *Eur Respir J Suppl*, 40: 81s-85s. doi: 10.1183/09031936.03.00402703.
Jore, S., Schwarse, P. E. & Aakra, Å. (2022). *Klimaendringer og helse* Available at: https://www.fhi.no/nettpub/hin/miljo/klima-og-helse/ (accessed: 11.10.2014).
klimaloven. (2017). *Lov om klimamål (LOV-2017-06-16-60)*. Available at: https://lovdata.no/dokument/NL/lov/2017-06-16-60
Leichenko, R. & O'Brien, K. (2019). *Climate and society: Transforming the future*: John Wiley & Sons.
Link, B. G. & Phelan, J. (1995). Social conditions as fundamental causes of disease. *J Health Soc Behav*, Spec No: 80-94.
Liu, C., Yavar, Z. & Sun, Q. (2015). Cardiovascular response to thermoregulatory challenges. *Am J Physiol Heart Circ Physiol*, 309 (11): H1793-812. doi: 10.1152/ajpheart.00199.2015.
Liu, J., Varghese, B. M., Hansen, A., Zhang, Y., Driscoll, T., Morgan, G., Dear, K., Gourley, M., Capon, A. & Bi, P. (2022). Heat exposure and cardiovascular health outcomes: a systematic review and meta-analysis. *Lancet Planet Health*, 6 (6): e484-e495. doi: 10.1016/S2542-5196(22)00117-6.
Mackenbach, J. P., Looman, C. W. N., Artnik, B., Bopp, M., Deboosere, P., Dibben, C., Kalediene, R., Kovacs, K., Leinsalu, M., Martikainen, P., et al. (2017). 'Fundamental causes' of inequalities in mortality: an empirical test of the theory in 20 European populations. *Sociol Health Illn*, 39 (7): 1117-1133. doi: 10.1111/1467-9566.12562.
Madrigano, Mittleman, M. A., Baccarelli, A., Goldberg, R., Melly, S., von Klot, S. & Schwartz, J. (2013). Temperature, myocardial infarction, and mortality: effect modification by individual- and area-level characteristics. *Epidemiology*, 24 (3): 439-46. doi: 10.1097/EDE.0b013e3182878397.

Madrigano, J., Mittleman, M. A., Baccarelli, A., Goldberg, R., Melly, S., Von Klot, S. & Schwartz, J. (2013). Temperature, myocardial infarction, and mortality: effect modification by individual and area-level characteristics. *Epidemiology (Cambridge, Mass.)*, 24 (3): 439.

Masselot, P., Mistry, M., Vanoli, J., Schneider, R., Iungman, T., Garcia-Leon, D., Ciscar, J. C., Feyen, L., Orru, H., Urban, A., et al. (2023). Excess mortality attributed to heat and cold: a health impact assessment study in 854 cities in Europe. *Lancet Planet Health*, 7 (4): e271-e281. doi: 10.1016/S2542-5196(23)00023-2.

Meld. St. 13 (2020-2021). *Klimaplan for 2021-2030*. Oslo: Klima- og miljødepartementet. Available at: https://www.regjeringen.no/no/dokumenter/meld.-st.-13-20202021/id2827405/

Meld. St. 15 (2022-2023). *Folkehelsemeldingen 2022-2023 - Nasjonal strategi for utjevning av sosiale helseforskjeller*. Oslo: Helse - og omsorgsdepartementet. Available at: https://www.regjeringen.no/no/dokumenter/meld.-st.-15-20222023/id2969572/ (accessed: 15.04.2023).

Meld. St. 19 (2018-2019). *Folkehelsemeldinga - Gode liv i eit trygt samfunn*. Oslo: Helse - og omsorgsdepartementet. Available at: https://www.regjeringen.no/no/dokumenter/meld.-st.-19-20182019/id2639770/ (accessed: 12.01.2023).

Meld. St. 33 (2012-2013). *Climate change adaptation in Norway* Oslo Norwgian Minisrty of Climate and Enviroment Available at: https://www.regjeringen.no/contentassets/e5e7872303544ae38bdbdc82aa0446d8/en-gb/pdfs/stm201220130033000engpdfs.pdf (accessed: 10.10.2022).

Miljødirektoratet. (2022a). *Dagens og framtidens klima for Norge*. Available at: https://www.miljodirektoratet.no/ansvarsomrader/klima/for-myndigheter/klimatilpasning/klimatilpasning-krever-kunnskap/dagens-og-framtidens-klima/ (accessed: 17.02.2023).

Miljødirektoratet. (2022b). *Funn fra FNs klimapanel AR6: kunnskap med relevans for Norge fra andre delrapport om virkninger, sårbarhet og tilpasninger til klimaendringer*. M-2364. Available at: https://www.miljodirektoratet.no/publikasjoner/2022/oktober/funn-fra-fns-klimapanel-ar6/ (accessed: 19.01.2023).

Nafstad, P., Skrondal, A. & Bjertness, E. (2001). Mortality and temperature in Oslo, Norway, 1990–1995. *European Journal of Epidemiology*, 17 (7): 621-627. doi: 10.1023/A:1015547012242.

Naidoo, J. & Wills, J. (2016). *Foundations for health Promotion.* . 4. ed. Great Britain: ELSEVIER

NOU 2010: 10. *Tilpasning til eit klima i endring. Samfunnet si sårbarheit og behov for tilpasning til konsekvensar av klimeandringane*. Available at: https://www.regjeringen.no/contentassets/01c4638b3f3e4573929f3b375f4731e0/nn-no/pdfs/nou201020100010000dddpdfs.pdf (accessed: 10.10.2022).

Nystad, W. (2022). *Kronisk obstruktiv lungesykdom (kols) i Norge*. Available at: https://www.fhi.no/nettpub/hin/ikke-smittsomme/kols/ (accessed: 02/09/2022).

Ólafsdóttir, S. (2021). Health Inequalities in the Nordic Countries : What We Know and What We Can Do. (2021).

Oslo Kommune. (2013). *Klimatilpasningsstrategi for Oslo kommune*. Rapport utgitt av Bymiljøetaten, november 2013. Available at:

https://www.oslo.kommune.no/getfile.php/13108654-1455893001/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Miljø%20og%20klima/Slik%20jobber%20vi%20med%20miljø%20og%20klima/Styrende%20dokumenter/Klimatilpasningsstrategi%20for%20Oslo%20kommune.PDF (accessed: 11.02.2023).

Oslo Kommune. (2020a). *Climate Change Vulnerability Analysis for Oslo. Short version*: Klimaetaten. Available at: https://www.klimaoslo.no/wp-content/uploads/sites/88/2021/03/Climate-Change-Vulnerability-Analysis-for-Oslo-short-version.pdf (accessed: 09.10.2022).

Oslo kommune. (2020b). *Oslohelsa 2020. Oversikt over påvirkningsfaktorer og helsetisland i Oslo*: Helseetaten. Available at: https://www.oslo.kommune.no/getfile.php/13131998-1590734531/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Statistikk/Rapport%20Oslohelsa%202020.pdf (accessed: 21.01.23).

Oslo Kommune. (2023). *Utkast Folkehelsestrategi for Oslo-2023-2030*. Available at: https://www.oslo.kommune.no/getfile.php/13463925-1664451903/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Folkehelse/Ny%20folkehelsestrategi/Utkast%20Folkehelsestrategi%20for%20Oslo%20%282023-2030%29%20endelig.pdf (accessed: 21.01.23).

Øversveen, E. & Rydland, H. T. (2021). *Sosial ulikhet i helse. En sammfunnsvitenskapelig innføring.* . Bergen: Fagbokforlaget.

Paavola, J. (2017). Health impacts of climate change and health and social inequalities in the UK. *Environ Health*, 16 (Suppl 1): 113. doi: 10.1186/s12940-017-0328-z.

Phelan, J. C., Link, B. G. & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav*, 51 Suppl: S28-40. doi: 10.1177/0022146510383498.

Pickett, K. E. & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, 128: 316-326. doi: https://doi.org/10.1016/j.socscimed.2014.12.031.

Pörtner, H. O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., et al. (2022). *IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Sixth assessment Intergovernmental Panel on Climate Change.

Ragettli, M. S., Saucy, A., Flückiger, B., Vienneau, D., de Hoogh, K., Vicedo-Cabrera, A. M., Schindler, C. & Röösli, M. (2023). Explorative Assessment of the Temperature–Mortality Association to Support Health-Based Heat-Warning Thresholds: A National Case-Crossover Study in Switzerland. *International Journal of Environmental Research and Public Health*, 20 (6): 4958.

Ranhoff, A. H., Hygen, H. O., Di Ruscio, F., Rao, S. & Strand, B. H. (2019). The warm summer of 2018 - increased mortality among older people? *Tidsskr Nor Laegeforen*, 139 (10). doi: 10.4045/tidsskr.19.0167.

Reinwald, F., Haluza, D., Pitha, U. & Stangl, R. (2021). Urban Green Infrastructure and Green Open Spaces: An Issue of Social Fairness in Times of COVID-19 Crisis. *Sustainability*, 13 (19). doi: 10.3390/su131910606.

Ringdal, K. (2014). *Enhet og Mangfold. Samfunnsvitenskapelig forskning og kvantitativ metode, 3.utgave.* Bergen: Fagbokforlaget.

Rocklöv, J., Forsberg, B., Ebi, K. & Bellander, T. (2014). Susceptibility to mortality related to temperature and heat and cold wave duration in the population of Stockholm County, Sweden. *Glob Health Action*, 7: 22737. doi: 10.3402/gha.v7.22737.

Rodgers, G. B. (1979). Income and inequality as determinants of mortality: An international cross-section analysis. *Population Studies*, 33 (2): 343-351. doi: 10.1080/00324728.1979.10410449.

Romanello, M., McGushin, A., Di Napoli, C., Drummond, P., Hughes, N., Jamart, L., Kennard, H., Lampard, P., Solano Rodriguez, B., Arnell, N., et al. (2021). The 2021 report of the Lancet Countdown on health and climate change: code red for a healthy future. *Lancet*, 398 (10311): 1619-1662. doi: 10.1016/S0140-6736(21)01787-6.

Ruuhela, R., Hyvärinen, O. & Jylhä, K. (2018). Regional Assessment of Temperature-Related Mortality in Finland. *Int J Environ Res Public Health*, 15 (3). doi: 10.3390/ijerph15030406.

Saucy, A., Ragettli, M. S., Vienneau, D., de Hoogh, K., Tangermann, L., Schaffer, B., Wunderli, J. M., Probst-Hensch, N. & Roosli, M. (2021). The role of extreme temperature in cause-specific acute cardiovascular mortality in Switzerland: A case-crossover study. *Sci Total Environ*, 790: 147958. doi: 10.1016/j.scitotenv.2021.147958.

Son, J.-Y., Liu, J. C. & Bell, M. L. (2019). Temperature-related mortality: a systematic review and investigation of effect modifiers. *Environmental Research Letters*, 14 (7). doi: 10.1088/1748-9326/ab1cdb.

Statistisk sentralbyrå. (2022a). *Befolkning*. Available at: https://www.ssb.no/befolkning/folketall/statistikk/befolkning (accessed: 11.10.2022).

Statistisk sentralbyrå. (2022b). *Befolkningens utdanningsnivå* Available at: https://www.ssb.no/utdanning/utdanningsniva/statistikk/befolkningens-utdanningsniva (accessed: 05.05.2023).

Statistisk sentralbyrå. (2022c). *Tettsteders befolkning og areal*. Available at: https://www.ssb.no/befolkning/folketall/statistikk/tettsteders-befolkning-og-areal (accessed: 12.02.2023).

Statistisk sentralbyrå. (s.a). *Variabeldefinisjon Inntekt etter skatt*. Available at: https://www.ssb.no/a/metadata/conceptvariable/vardok/2/nb (accessed: 14.03.23).

Syse, A., Grøholt, E.-K., Madsen, C., Aarø, L. E., Strand, B. H. & Næss, Ø. E. (2022). *Sosiale helseforskjeller i Norge*. Available at: https://www.fhi.no/en/op/hin/population/social-inequalities/

The Lancet. (2018). Heatwaves and health. *The Lancet*, 392 (10145): 359. doi: https://doi.org/10.1016/S0140-6736(18)30434-3.

Torp, I. S. (2019). *Declaration of Helsinki* Available at: https://www.forskningsetikk.no/en/guidelines/medical-and-health-research/declaration-of-helsinki/ (accessed: 17.04.2023).

United Nations. (s.a.-a). *The 17 goals* Available at: https://sdgs.un.org/goals#goals (accessed: 25.09.2022).

United Nations. (s.a.-b). *Leave No One Behind* Available at: https://unsdg.un.org/2030-agenda/universal-values/leave-no-one-behind (accessed: 07.12.2022).

United Nations. (s.a.-c). *The Paris Agreement* Available at: https://www.un.org/en/climatechange/paris-agreement (accessed: 12.10.2022).

Universitetet I Oslo. (s.a.). *Tjenester for Sensitive Data (TSD)*. Available at: https://www.uio.no/tjenester/it/forskning/sensitiv/ (accessed: 03.05.2023).

Urban, A., Burkart, K., Kyselý, J., Schuster, C., Plavcová, E., Hanzlíková, H., Štěpánek, P. & Lakes, T. (2016). Spatial Patterns of Heat-Related Cardiovascular Mortality in the Czech Republic. *Int J Environ Res Public Health*, 13 (3). doi: 10.3390/ijerph13030284.

Venter, Z. S., Figari, H., Krange, O. & Gundersen, V. (2023). Environmental justice in a very green city: Spatial inequality in exposure to urban nature, air pollution and heat in Oslo, Norway. *Science of The Total Environment*, 858: 160193. doi: https://doi.org/10.1016/j.scitotenv.2022.160193.

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Belesova, K., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Capstick, S., et al. (2019). The 2019 report of The Lancet Countdown on health and climate change: ensuring that the health of a child born today is not defined by a changing climate. *The Lancet*, 394 (10211): 1836-1878. doi: https://doi.org/10.1016/S0140-6736(19)32596-6.

Webb, P., Bain, C. & Page, A. (2017). *Essential Epidemiology. An Introduction for Students and Health Professionals.* 3. ed. Cambridge, United Kingdom: Cambrigde University Pres.

Wichmann, J., Andersen, Z. J., Ketzel, M., Ellermann, T. & Loft, S. (2011). Apparent Temperature and Cause-Specific Mortality in Copenhagen, Denmark: A Case-Crossover Analysis. *International Journal of Environmental Research and Public Health*, 8 (9): 3712-3727.

Wilkinson, R. & Pickett, K. (2011). Greater equality. The hidden key to better healt and higher scores. . *American Educator*, 35 (1): 5-9.

Witt, C., Schubert, A. J., Jehn, M., Holzgreve, A., Liebers, U., Endlicher, W. & Scherer, D. (2015). The Effects of Climate Change on Patients With Chronic Lung Disease. A Systematic Literature Review. *Dtsch Arztebl Int*, 112 (51-52): 878-83. doi: 10.3238/arztebl.2015.0878.

World Health Organization. (2019). *Environmental health inequalities in Europe. Second assessment report.* Copenhagen: WHO Regional Office for Europe.

World Health Organization. (2021a). *Cardiovascular diseases (CVDs)*. Available at: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed: 29/08/2022).

World Health Organization. (2021b). *Climate change and health* Available at: https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health (accessed: 03.11.2022).

World Health Organization. (2021c). *Urban health*. Available at: https://www.who.int/health-topics/urban-health#tab=tab_1 (accessed: 11.01.2023).

World Health Organization. (s.a.-a). *Chronic respiratory diseases*. Available at: https://www.who.int/health-topics/chronic-respiratory-diseases#tab=tab_1 (accessed: 26.10.2022).

World Health Organization. (s.a.-b). *Social determinants of health*: World Health Organization. Available at: https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1 (accessed: 09.11.2022).

Yu, W., Vaneckova, P., Mengersen, K., Pan, X. & Tong, S. (2010). Is the association between temperature and mortality modified by age, gender and socio-economic status? *Science of The Total Environment*, 408 (17): 3513-3518. doi: https://doi.org/10.1016/j.scitotenv.2010.04.058.

Zanobetti, A., O'Neill, M. S., Gronlund, C. J. & Schwartz, J. D. (2013). Susceptibility to mortality in weather extremes: effect modification by personal and small-area

characteristics. *Epidemiology*, 24 (6): 809-19. doi: 10.1097/01.ede.0000434432.06765.91.

Zhang, S., Rai, M., Breitner, S., Aunan, K. & Schneider, A. (2020). Climate change and the projected burden of future health impacts – The Project EXHAUSTION. *Public Health Forum*, 28 (1): 17-20. doi: 10.1515/pubhef-2019-0105.

Region:	Saksbehandler:	Telefon:	Vår dato:	Vår referanse:
REK sør-øst C	Øyvind Grønlie Olsen	22857547	25.07.2022	10347

Shilpa Rao-Skirbekk

Prosjektsøknad: Kardiopulmonære effekter av eksponering for ekstreme temperaturer og luftforurensning i Europa
Søknadsnummer: 2019/297
Forskningsansvarlig institusjon: Folkehelseinstituttet

Prosjektsøknad: Endring godkjennes

Søkers beskrivelse

Prosjektet vil bruke epidemiologiske data og modellering for å avdekke sårbarhetsfaktorer av betydning for å anslå framtidig forekomst av CPD og å identifisere kostnadseffektive tiltak i Europa. Mens det er påvist en statistisk sammenheng mellom klimaparametere og helseproblemer, er de underliggende årsaksmekanismer og faktorer som øker eller reduserer sårbarheten fortsatt dårlig forstått. I dette prosjektet undersøker vi hvordan et bredt spekter av sårbarhetsfaktorer endrer virkningen av ekstrem varme og luftforurensning på CPD-dødelighet og sykelighet og i hvilken grad og hvordan tilpasningsstrategier er i stand til å speile og redusere disse sårbarhetsfaktorene. Prosjektet vil bruke registerdata, data fra helseundersøkelser og andre relevante datakilder til å undersøke kardiopulmonære effekter av ekstern temperatur og luftforurensning i den norske voksen populasjonen.

Vi viser til søknad om prosjektendring mottatt 08.06.2022 for ovennevnte forskningsprosjekt. Søknaden er behandlet av sekretariatet i Regional komité for medisinsk og helsefaglig forskningsetikk (REK) på delegert fullmakt fra komiteen, med hjemmel i forskningsetikkforskriften § 7, første ledd, tredje punktum. Søknaden er vurdert med hjemmel i helseforskningsloven § 11.

REKs vurdering

Søknaden om godkjenning av prosjektendringer omfatter at Ashley Ahimbisibwe, Kicki Svensson og Camilla Nyland inngår som prosjektmedarbeidere i forskningsprosjektet.

REK har ingen innvendinger til innlemmelsen av nye prosjektmedarbeisdere.

Vedtak

Med hjemmel i helseforskningsloven § 11godkjenner REK prosjektendringen.

Sluttmelding

Prosjektleder skal sende sluttmelding til REK på eget skjema via REK-portalen senest 6 måneder etter sluttdato, jf. helseforskningsloven § 12. Dersom prosjektet ikke starter opp

eller gjennomføres meldes dette også via skjemaet for sluttmelding.

Søknad om endring
Dersom man ønsker å foreta vesentlige endringer i formål, metode, tidsløp eller organisering må prosjektleder sende søknad om endring via portalen på eget skjema til REK, jf. helseforskningsloven § 11.

Klageadgang
Du kan klage på REKs vedtak, jf. forvaltningsloven § 28 flg. Klagen sendes på eget skjema via REK portalen. Klagefristen er tre uker fra du mottar dette brevet. Dersom REK opprettholder vedtaket, sender REK klagen videre til Den nasjonale forskningsetiske komité for medisin og helsefag (NEM) for endelig vurdering, jf. forskningsetikkloven § 10 og helseforskningsloven § 10.

Med vennlig hilsen

Jacob Hølen
Sekretariatsleder
REK sør-øst

Øyvind Grønlie Olsen
Seniorrådgiver
REK sør-øst

Kopi til:

Folkehelseinstituttet
Terese Bekkevold

R script main interaction

Camilla/Kicki

5/14/2023

```
#-----------------------------------------------------#
# Name: R code_short-term temperature and mortality #
# Adapted from EXHAUSTION script                      #
# Version: 02.03.2023                                 #
#-----------------------------------------------------#

rm(list=ls())

## Load packages to be used; Author's R version 4.1.0
library(dlnm)
library(dplyr)
library(survival)
library(splines)
library(miceadds)

## Set the working directory
path <- "N:/durable/LEVEL3_MORTALITY/"
setwd(path)

## load the functions
source("Kicki/R code_adapted/00_FUNCTION.R")


## 1. Import data sets ####

## Note: The data set should include the following variables
##       (1) status: indicator of the case (=1) or control (=0)
##       (2) temp_s0 - temp_s10: single-day lags of air temperature
##       (3) id (ID for participants): only participants who experienced events during the follow up (ca
load.RDS(filename="Kicki/R code_adapted/DATA_EXHAUSTION_HUBRO_mort.RDS","mort") #for the first dataset

mort<-DATA_EXHAUSTION_HUBRO_mort

## create data sets for each mortality outcome
mort_natural <- mort ## natural cause mortality
mort_cpd <- subset(mort,cause_death%in%c(sprintf("I%02d", 0:99),sprintf("J%02d", 0:99),as.character(390
mort_cvd <- subset(mort,cause_death%in%c(sprintf("I%02d", 0:99),as.character(390:459))) ## cardiovascul
mort_ihd <- subset(mort,cause_death%in%c(sprintf("I%02d", 20:25),as.character(410:414))) ## ischemic he
mort_cerebr <- subset(mort,cause_death%in%c(sprintf("I%02d", 60:69),as.character(430:438))) ## cerebrova
mort_resp <- subset(mort,cause_death%in%c(sprintf("J%02d", 0:99),as.character(460:519))) ## respiratory
mort_copd <- subset(mort,cause_death%in%c(sprintf("J%02d", 40:44),"J47",as.character(c(490:492,494:496))
```

```r
## create a list containing all datasets
dlist <- list(mort_natural,mort_cpd,mort_cvd,mort_ihd,mort_cerebr,mort_resp,mort_copd)
## names of the outcomes in the order of that in the dlist
name_outcome <- c("mort_natural","mort_cpd","mort_cvd","mort_ihd","mort_cerebr","mort_resp","mort_copd"


## 2. Effect estimation (refer to the function "casecrs" in "00_FUNCTION.R") ####

## output with elements for creating exposure-response curves ####
result1 <- list()

## output to be saved, not including "mat_temp" and "cb_temp" to reduce the size
out1 <- list()

for (i in 1:length(dlist)){
  mort_hubro <- dlist[[i]]
  result1[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(min=TRU
                          estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  out1[[i]] <- within(result1[[i]],rm(mat_temp,cb_temp))
}
names(result1) <- names(out1) <- name_outcome
saveRDS(out1,file = "Kicki/short-term_temp_mort_HUBRO.RDS")



## 3. Exposure-response curves ####

## e.g. natural-cause mortality
tiff("Kicki/short-term_temp_mort_natural_HUBRO.tiff",width=600,height=500) #change dataset here, exampl
layout(mat = matrix(c(1,2),2,1, byrow=TRUE),  height = c(7,3))
par(mar=c(2, 4, 1, 1))

# get the output for figures
mat_temp <- result1$mort_natural$mat_temp; cb_temp <- result1$mort_natural$cb_temp
coef <- result1$mort_natural$model_coef; vcov <- result1$mort_natural$model_vcov

# centering temperature MMT, if MMT < 5%, use 5%, if MMT > 95%, use 95%
mmt <- findmin(cb_temp,coef=coef,vcov=vcov);temp5 <- quantile(mat_temp,0.05);temp95 <- quantile(mat_temp
cen_plot <- ifelse(mmt<temp5,temp5,ifelse(mmt>temp95,temp95,mmt))

# exposure-response function
cp <- crosspred(cb_temp, coef=coef, vcov=vcov, model.link="logit", cen=cen_plot, by=0.1, cumul=T)
plot(cp, "overall", ylab="OR", xlab="Temperature", lwd=1.5, xlim=range(mat_temp),ylim=c(0,6),col=2) ## a
mtext("HUBRO Cohort",cex=1.2,line=-0.5,font=2) ## add cohort name

abline(v=quantile(mat_temp,c(1,99)/100),lty=2,lwd=1)
axis(1,at=-8:8*5)

# histogram
breaks <- c(min(mat_temp,na.rm=T)-1,seq(cp$predvar[1],cp$predvar[length(cp$predvar)],length=30),max(mat_
hist <- hist(mat_temp,breaks=breaks,plot=FALSE)
hist$density = hist$counts/sum(hist$counts)*100
plot(hist,freq=FALSE,col=grey(0.95),main=NULL,xaxt="n",xlim=range(mat_temp),ylab="Density (%)")
```

```
abline(v=quantile(mat_temp,c(1,99)/100),lty=2,lwd=1)
layout(mat = matrix(c(1,2),2,1, byrow=TRUE),  height = c(5,5))
dev.off()


## 4. Effect modification ####

## 4.1 Greenness ####
result1 <- list()
## output to be saved, not including "mat_temp" and "b1temp" to reduce the size
out1 <- list()

for (i in 1:length(dlist)){
  data <- dlist[[i]]
  data$ndvi_cat <- ifelse(data$NDVI>=median(data$NDVI,na.rm=T),1,0)
  result1[[i]] <- casecrs_int_2cats(status="status",id="LOPENR",modif="ndvi_cat",lag=10,varper=c(10,75,9
                                    estpct=c(1,2.5,5,10,90,95,97.5,99),data=data)
  out1[[i]] <- within(result1[[i]],rm(mat_temp,b1temp))
}
names(result1) <- names(out1) <- name_outcome
saveRDS(out1,file = "Kicki/short-term_temp_mort_greennes_HUBRO.RDS")


## 4.2 Socio-economic: income ####
result2 <- list()
## output to be saved, not including "mat_temp" and "b1temp" to reduce the size
out2 <- list()

for (i in 1:length(dlist)){
  data <- dlist[[i]]
  data$income_cat <- ifelse(data$income_h>=median(data$income_h,na.rm=T),1,0)
  result2[[i]] <- casecrs_int_2cats(status="status",id="LOPENR",modif="income_cat",lag=10,varper=c(10,75
                                    estpct=c(1,2.5,5,10,90,95,97.5,99),data=data)
  out2[[i]] <- within(result2[[i]],rm(mat_temp,b1temp))
}
names(result2) <- names(out2) <- name_outcome
saveRDS(out2,file = "Kicki/short-term_temp_mort_income_HUBRO.RDS")


## 4.3. Education  ####
result3 <- list()
## output to be saved, not including "mat_temp" and "b1temp" to reduce the size
out3 <- list()

for (i in 1:length(dlist)){
  data <- dlist[[i]]
  result3[[i]] <- casecrs_int_3cats(status="status",id="LOPENR",modif="edulev",lag=10,varper=c(10,75,90)
                                    estpct=c(1,2.5,5,10,90,95,97.5,99),data=data)
  out3[[i]] <- within(result3[[i]],rm(mat_temp,b1temp))
}
names(result3) <- names(out3) <- name_outcome
saveRDS(out3,file = "Kicki/short-term_temp_mort_education_HUBRO.RDS")
```

HUBRO+Sensitivity analysis

Camilla/Kicki

5/14/2023

```
#---------------------------------------#
# Name:5.1_sensitivity_case-crossover   #
# Project: EXHAUSTION                   #
# Version: 01.12.2021                   #
#---------------------------------------#
rm(list=ls())

##_____________________________________________________________________________________####
##                          Sensitivity analysis: case-crossover                        ####
##-------------------------------------------------------------------------------------###

## Load packages to be used; Author's R version 3.6.3
library(dlnm)
library(dplyr)
library(survival)
library(splines)
library(miceadds)
library(lubridate)

## Set the working directory
path <- "N:/durable/LEVEL3_MORTALITY/"
setwd(path)

## load the functions
source("Kicki/R code_adapted/00_FUNCTION.R")

## Import data sets ####

## Note: The data set should include the following variables
##       (1) status: indicator of the case (=1) or control (=0)
##       (2) temp_s0 - temp_s10: single-day lags of air temperature
##       (3) id (ID for participants): only participants who experienced events during the follow up (c
mort <- readRDS("N:/durable/LEVEL3_MORTALITY/Kicki/R code_adapted/DATA_EXHAUSTION_HUBRO_mort.RDS") ## w

## create data sets for each mortality outcome
mort_natural <- mort ## natural cause mortality
mort_cpd <- subset(mort_natural,cause_death%in%c(sprintf("I%02d", 0:99),sprintf("J%02d", 0:99),as.chara
mort_cvd <- subset(mort_natural,cause_death%in%c(sprintf("I%02d", 0:99),as.character(390:459))) ## cardi
mort_ihd <- subset(mort_natural,cause_death%in%c(sprintf("I%02d", 20:25),as.character(410:414))) ## isc
mort_cerebr <- subset(mort_natural,cause_death%in%c(sprintf("I%02d", 60:69),as.character(430:438))) ##
mort_resp <- subset(mort_natural,cause_death%in%c(sprintf("J%02d", 0:99),as.character(460:519))) ## resp
mort_copd <- subset(mort_natural,cause_death%in%c(sprintf("J%02d", 40:44),"J47",as.character(c(490:492,4
```

```r
## create a list containing all data sets
dlist <- list(mort_natural,mort_cpd,mort_cvd,mort_ihd,mort_cerebr,mort_resp,mort_copd)
## names of the outcomes in the order of that in the dlist
name_outcome <- c("mort_natural","mort_cpd","mort_cvd","mort_ihd","mort_cerebr","mort_resp","mort_copd"


## 1. Lag 0-3 of temperature ####
sens1 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  sens1[[i]] <- casecrs(status="status",id="LOPENR",lag=3,varper=c(10,75,90),lagnk=1,cen=list(pct=c(25,
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens1[[i]] <- within(sens1[[i]],rm(mat_temp,cb_temp))
}
names(sens1) <- name_outcome
saveRDS(sens1,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_lag_03.RDS")


## 2. Lag 0-21 of temperature
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens2 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  sens2[[i]] <- casecrs(status="status",id="LOPENR",lag=21,varper=c(10,75,90),lagnk=3,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens2[[i]] <- within(sens2[[i]],rm(mat_temp,cb_temp))
}
names(sens2) <- name_outcome
saveRDS(sens2,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_lag_21.RDS")



## 3. age 65+ ####
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens3 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort <- subset(mort,age_death>=65)
  sens3[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens3[[i]] <- within(sens3[[i]],rm(mat_temp,cb_temp))
}
names(sens3) <- name_outcome
saveRDS(sens3,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_age65plus.RDS")


## 4. age 75+ ####
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens4 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort <- subset(mort,age_death>=75)
  sens4[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
```

```
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens4[[i]] <- within(sens4[[i]],rm(mat_temp,cb_temp))
}
names(sens4) <- name_outcome
saveRDS(sens4,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_age75plus.RDS")


## 5. male ####
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens5 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort <- subset(mort,sex==1)
  sens5[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens5[[i]] <- within(sens5[[i]],rm(mat_temp,cb_temp))
}
names(sens5) <- name_outcome
saveRDS(sens5,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_male.RDS")


## 6. female ####
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens6 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort <- subset(mort,sex==0)
  sens6[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens6[[i]] <- within(sens6[[i]],rm(mat_temp,cb_temp))
}
names(sens6) <- name_outcome
saveRDS(sens6,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_female.RDS")


## 7. cold season (Nov-March) ####
##lag10
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens7 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort$month <- month(mort$d_death)
  sens7[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens7[[i]] <- within(sens7[[i]],rm(mat_temp,cb_temp))
}
names(sens7) <- name_outcome
saveRDS(sens7,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_coldseason_lag10.RDS")


## 8. warm season (May-Sep) ####
## lag 10
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
```

```r
sens8 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort$month <- month(mort$d_death)
  mort <- subset(mort,month%in%5:9)
  sens8[[i]] <- casecrs(status="status",id="LOPENR",lag=10,varper=c(10,75,90),lagnk=2,cen=list(pct=c(25
                        estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens8[[i]] <- within(sens8[[i]],rm(mat_temp,cb_temp))
}
names(sens8) <- name_outcome
saveRDS(sens8,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_warmseason_lag10.RDS")


## lag 1
mort<- mort_natural #to get a big dataset again, it was only 619 obs from the last loop
sens8.2 <- list()
for (i in 1:length(dlist)){
  mort <- dlist[[i]]
  mort$month <- month(mort$d_death)
  mort <- subset(mort,month%in%5:9)
  sens8.2[[i]] <- casecrs_lag01(status="status",id="LOPENR",lag=1,varper=c(10,75,90),cen=list(pct=c(25,
                                estpct=c(1,2.5,5,10,90,95,97.5,99),data=dat)
  sens8.2[[i]] <- within(sens8.2[[i]],rm(mat_temp,cb_temp))
}
names(sens8.2) <- name_outcome
saveRDS(sens8.2,file = "Camilla/Results-HUBRO/sensitivity analysis/sensitivity_warmseason_lag1.RDS")
```

00_Function.R

Camilla/Kicki

5/14/2023

```
#----------------------------------------#
# Name:00_FUNCTION                       #
# Project: EXHAUSTION                    #
# Version: 27.04.2021                    #
#----------------------------------------#

##-------------------------------------------------------------------------------------####
## 1. FUNCTION TO ESTIMATE MINIMUM OF A EXPOSURE-RESPONSE FUNCTION FROM A FITTED MODEL ####
##-------------------------------------------------------------------------------------####
findmin <- function(basis,model=NULL,coef=NULL,vcov=NULL,at=NULL,from=NULL,
                    to=NULL,by=NULL,sim=FALSE,nsim=5000) {
  #
  ##############################################################################
  # R code from https://github.com/gasparrini/2017_tobias_Epidem_Rcodedata/blob/master/findmin.R
  #
  #   ARGUMENTS:
  #   - basis: A SPLINE OR OTHER BASIS FOR AN EXPOSURE x CREATED BY DLNM FUNCTION
  #            CROSSBASIS OR ONEBASIS
  #   - model: THE FITTED MODEL
  #   - coef AND vcov: COEF AND VCOV FOR basis IF model IS NOT PROVIDED
  #
  #   - at: A NUMERIC VECTOR OF x VALUES OVER WHICH THE MINIMUM IS SOUGHT
  #   OR
  #   - from, to: RANGE OF x VALUES OVER WHICH THE MINIMUM IS SOUGHT.
  #   - by: INCREMENT OF THE SEQUENCES x VALUES OVER WHICH THE MINIMUM IS SOUGHT
  #
  #   - sim: IF BOOTSTRAP SIMULATION SAMPLES SHOULD BE RETURNED
  #   - nsim: NUMBER OF SIMULATION SAMPLES
  ##############################################################################



  ##############################################################################
  # CREATE THE BASIS AND EXTRACT COEF-VCOV
  #
  # CHECK AND DEFINE BASIS
  if(!any(class(basis)%in%c("crossbasis","onebasis")))
    stop("the first argument must be an object of class 'crossbasis' or 'onebasis'")
  #
  # INFO
  one <- any(class(basis)%in%c("onebasis"))
  attr <- attributes(basis)
```

```
range <- attr(basis,"range")
if(is.null(by)) by <- 0.1
lag <- if(one) c(0,0) else cb=attr(basis,"lag")
if(is.null(model)&&(is.null(coef)||is.null(vcov)))
  stop("At least 'model' or 'coef'-'vcov' must be provided")
name <- deparse(substitute(basis))
cond <- if(one) paste(name,"[[:print:]]*b[0-9]{1,2}",sep="") else
  paste(name,"[[:print:]]*v[0-9]{1,2}\\.1[0-9]{1,2}",sep="")
#
# SET COEF, VCOV CLASS AND LINK
if(!is.null(model)) {
  model.class <- class(model)
  coef <- dlnm:::getcoef(model,model.class)
  ind <- grep(cond,names(coef))
  coef <- coef[ind]
  vcov <- dlnm:::getvcov(model,model.class)[ind,ind,drop=FALSE]
  model.link <- dlnm:::getlink(model,model.class)
} else model.class <- NA
#
# CHECK
if(length(coef)!=ncol(basis) || length(coef)!=dim(vcov)[1] ||
   any(is.na(coef)) || any(is.na(vcov)))
  stop("model or coef/vcov not consistent with basis")
#
# DEFINE at
at <- dlnm:::mkat(at,from,to,by,range,lag,bylag=1)
predvar <- if(is.matrix(at)) rownames(at) else at
predlag <- dlnm:::seqlag(lag,by=1)
#
# CREATE THE MATRIX OF TRANSFORMED CENTRED VARIABLES (DEPENDENT ON TYPE)
type <- if(one) "one" else "cb"
Xpred <- dlnm:::mkXpred(type,basis,at,predvar,predlag,cen=NULL)
Xpredall <- 0
for(i in seq(length(predlag))) {
  ind <- seq(length(predvar))+length(predvar)*(i-1)
  Xpredall <- Xpredall + Xpred[ind,,drop=FALSE]
}
#
################################################################################
# FIND THE MINIMUM
#
pred <- drop(Xpredall%*%coef)
ind <- which.min(pred)
min <- predvar[ind]
#
################################################################################
# SIMULATIONS
#
if(sim) {
  # SIMULATE COEFFICIENTS
  k <- length(coef)
  eigen <- eigen(vcov)
  X <- matrix(rnorm(length(coef)*nsim),nsim)
```

```
    coefsim <- coef + eigen$vectors %*% diag(sqrt(eigen$values),k) %*% t(X)
    # COMPUTE MINIMUM
    minsim <- apply(coefsim,2,function(coefi) {
      pred <- drop(Xpredall%*%coefi)
      ind <- which.min(pred)
      return(predvar[ind])
    })
  }
  #
  ################################################################################
  #
  res <- if(sim) minsim else min
  #
  return(res)
}


##----------------------------------------------------------------------------------------####
## 2. FUNCTION TO ESTIMATE MAXIMUM OF A EXPOSURE-RESPONSE FUNCTION FROM A FITTED MODEL ####
##----------------------------------------------------------------------------------------####

findmax <- function(basis,model=NULL,coef=NULL,vcov=NULL,at=NULL,from=NULL,
                    to=NULL,by=NULL,sim=FALSE,nsim=5000) {
  #
  ################################################################################
  #   Adapted from R code findmin()
  #   ARGUMENTS:
  #   - basis: A SPLINE OR OTHER BASIS FOR AN EXPOSURE x CREATED BY DLNM FUNCTION
  #            CROSSBASIS OR ONEBASIS
  #   - model: THE FITTED MODEL
  #   - coef AND vcov: COEF AND VCOV FOR basis IF model IS NOT PROVIDED
  #
  #   - at: A NUMERIC VECTOR OF x VALUES OVER WHICH THE MINIMUM IS SOUGHT
  #   OR

  #   - from, to: RANGE OF x VALUES OVER WHICH THE MINIMUM IS SOUGHT.
  #   - by: INCREMENT OF THE SEQUENCES x VALUES OVER WHICH THE MINIMUM IS SOUGHT
  #
  #   - sim: IF BOOTSTRAP SIMULATION SAMPLES SHOULD BE RETURNED
  #   - nsim: NUMBER OF SIMULATION SAMPLES
  ################################################################################


  ################################################################################
  # CREATE THE BASIS AND EXTRACT COEF-VCOV
  #
  # CHECK AND DEFINE BASIS
  if(!any(class(basis)%in%c("crossbasis","onebasis")))
    stop("the first argument must be an object of class 'crossbasis' or 'onebasis'")
  #
  # INFO
  one <- any(class(basis)%in%c("onebasis"))
  attr <- attributes(basis)
```

```
range <- attr(basis,"range")
if(is.null(by)) by <- 0.1
lag <- if(one) c(0,0) else cb=attr(basis,"lag")
if(is.null(model)&&(is.null(coef)||is.null(vcov)))
  stop("At least 'model' or 'coef'-'vcov' must be provided")
name <- deparse(substitute(basis))
cond <- if(one) paste(name,"[[:print:]]*b[0-9]{1,2}",sep="") else
  paste(name,"[[:print:]]*v[0-9]{1,2}\\.1[0-9]{1,2}",sep="")
#
# SET COEF, VCOV CLASS AND LINK
if(!is.null(model)) {
  model.class <- class(model)
  coef <- dlnm:::getcoef(model,model.class)
  ind <- grep(cond,names(coef))
  coef <- coef[ind]
  vcov <- dlnm:::getvcov(model,model.class)[ind,ind,drop=FALSE]
  model.link <- dlnm:::getlink(model,model.class)
} else model.class <- NA
#
# CHECK
if(length(coef)!=ncol(basis) || length(coef)!=dim(vcov)[1] ||
   any(is.na(coef)) || any(is.na(vcov)))
  stop("model or coef/vcov not consistent with basis")
#
# DEFINE at
at <- dlnm:::mkat(at,from,to,by,range,lag,bylag=1)
predvar <- if(is.matrix(at)) rownames(at) else at
predlag <- dlnm:::seqlag(lag,by=1)
#
# CREATE THE MATRIX OF TRANSFORMED CENTRED VARIABLES (DEPENDENT ON TYPE)
type <- if(one) "one" else "cb"
Xpred <- dlnm:::mkXpred(type,basis,at,predvar,predlag,cen=NULL)
Xpredall <- 0
for(i in seq(length(predlag))) {
  ind <- seq(length(predvar))+length(predvar)*(i-1)
  Xpredall <- Xpredall + Xpred[ind,,drop=FALSE]
}
#
################################################################################
# FIND THE MINIMUM
#
pred <- drop(Xpredall%*%coef)
ind <- which.max(pred)
min <- predvar[ind]
#
################################################################################
# SIMULATIONS
#
if(sim) {
  # SIMULATE COEFFICIENTS
  k <- length(coef)
  eigen <- eigen(vcov)
  X <- matrix(rnorm(length(coef)*nsim),nsim)
```

```
    coefsim <- coef + eigen$vectors %*% diag(sqrt(eigen$values),k) %*% t(X)
    # COMPUTE MINIMUM
    minsim <- apply(coefsim,2,function(coefi) {
      pred <- drop(Xpredall%*%coefi)
      ind <- which.max(pred)
      return(predvar[ind])
    })
  }
  #
  ################################################################################
  #
  res <- if(sim) minsim else min
  #
  return(res)
}


##-------------------------------------------------------------------------------------------####
## 3. FUNCTION OF CASE-CROSSOVER ANALYSIS                                                     ####
##-------------------------------------------------------------------------------------------####

casecrs <- function (status, id, confounder=NULL, lag, varper, lagnk, cen=list(min=NULL,max=NULL,degree=
                     estpct, data){

  ## Input: status, id, confounder: variables used to define the formula applied to clogit() in the for
  ##                                case.status~exposure+confounder+strata(matched.set)
  ##                                status: case status, 1=case, 0=control
  ##                                id: ID for participants
  ##                                confounder: optional, vector of covariates to be included in the mo
  ##        lag: the maximum lag in the cross basis
  ##        varper: numeric vector of percentiles of the distribution of temperature for internal knots
  ##        lagnk: the number of internal knots in the lag-response dimension
  ##        cen: a list to define the centering temperature
  ##            - "min" and "max": optional, "TRUE" if the minimum or maximum mortality temperature to
  ##            - "degree": optional, numeric vector of temperature (?C)
  ##        estpct: numeric vector of percentiles of temperature distributions for effect estimate comp


  ## Build cross-basis function of temperature and lags
  ## Note: (1) exposure-response: natural cubic spline with internal knots placed at percentile of the
  ##                              distribution as defined by "varper"
  ##       (2) lag-response: natural cubic spline with an intercept and n="lag" internal knots placed a
  ##                         equally spaced values on the log scale

  ## 1. delete observations with NA in temperature
  dat <- subset(mort,rowSums(is.na(mort[which(names(mort)%in%paste0("temp_s",0:lag))]))==0)
  ## 2. extract matrix of temperature at lag0 to lag="lag"
  mat_temp <- as.matrix(dplyr::select(dat,paste0("temp_s",0:lag)))
  ## 3. define basis for temperature
  argvar <- list(fun="ns",knots=quantile(mat_temp,varper/100,na.rm=T))
  ## 4. define basis for lag
  arglag <- list(fun="ns",knots=logknots(lag,lagnk))
  ## 5. build the cross-basis function
```

```r
cb_temp <- crossbasis(mat_temp,lag=c(0,lag),argvar=argvar,arglag = arglag)

## Different percentile of the temperature matrix
tper <- quantile(mat_temp,seq(0,100,1)/100)

## Temperature summary for case days
tsum_case <- summary(subset(dat,status==1)$temp_s0)
tsum_case["SD"] <- sd(subset(dat,status==1)$temp_s0)

## Temperature summary for control days
tsum_control <- summary(subset(dat,status==0)$temp_s0)
tsum_control["SD"] <- sd(subset(dat,status==0)$temp_s0)

## Conditional logistic regression ####
## Note: "clogit" function in the "survival" package (same output as "clogistic" in "Epi" package)

if (is.null(confounder)==F){
  fml <- as.formula(paste0(status,"~cb_temp+strata(",id,")+",paste0(confounder,collapse = "+")))
} else {
  fml <- as.formula(paste0(status,"~cb_temp+strata(",id,")"))
}
mod <- clogit(fml,data=dat)

# Reduction to overall cumulative (it is irrelevant the cen value)
red <- crossreduce(cb_temp, mod, cen = 20)
# Store reduced coefs
coef <- coef(red)
vcov <- vcov(red)

## centering temperature
cen_temp <- NULL;cen_name <- NULL

if (is.null(cen$min)==F){
  cen_temp <- c(cen_temp,findmin(cb_temp,mod))
  cen_name <- c(cen_name,"min")
}
if (is.null(cen$max)==F){
  cen_temp <- c(cen_temp,findmax(cb_temp,mod))
  cen_name <- c(cen_name,"max")
}
if (is.null(cen$degree)==F){
  cen_temp <- c(cen_temp,cen$degree)
  cen_name <- c(cen_name,paste0(cen$degree," degree"))
}
if (is.null(cen$pct)==F){
  cen_temp <- c(cen_temp,quantile(mat_temp,cen$pct/100))
  cen_name <- c(cen_name,paste0(cen$pct,"th"))
}

## Predict ORs from each cen_temp to each estpct
estimate <- list()
for (k in 1:length(cen_temp)){
  pred <- crosspred(cb_temp, mod, model.link="logit", cen=cen_temp[k], at=quantile(mat_temp,estpct/10
```

```
    estimate[[k]] <-  round(data.frame(OR=pred$allRRfit,CIlow=pred$allRRlow,CIhigh=pred$allRRhigh),3)
    estimate[[k]]$temp <- as.numeric(rownames(estimate[[k]]))
    estimate[[k]]$perc <- paste0(estpct,"th")
    estimate[[k]]$cen <- cen_name[k]
    estimate[[k]] <- dplyr::select(estimate[[k]],c(cen,perc,temp,everything()))
  }
  estimate_all <- do.call(rbind,estimate)
  rownames(estimate_all) <- NULL

  ## output:result, a list containing the following elements
  ##        - n_case: number of cases
  ##        - n_control: number of controls
  ##        - tper: temperature distribution (percentiles)
  ##        - tsum_case: summary of temperature on case days
  ##        - tsum_control: summary of temperature on case days
  ##        - coef: coefficients for the overall association
  ##        - vcov: variance-covariance of coefs for overall association
  ##        - estimate: OR and CI at the "estpct" percentile of temperature distribution compared to ea
  ##        output for plots
  ##        - mat_temp: matrix of temperature
  ##        - cb_temp: cross-basis of temperature
  ##        - model_coef: coefficients of conditional logistic regression model
  ##        - model_vcov: variance matrix of conditional logistic regression model

  result <- NULL
  result$n_case <- nrow(subset(dat,status==1))
  result$n_control <- nrow(subset(dat,status==0))
  result$tper <- tper
  result$tsum_case <- tsum_case
  result$tsum_control <- tsum_control
  result$coef <- coef
  result$vcov <- vcov
  result$estimate <- estimate_all
  result$mat_temp <- mat_temp
  result$cb_temp <- cb_temp
  result$model_coef <- mod$coefficients
  result$model_vcov <- mod$var
  return(result)
}




##-------------------------------------------------------------------------------------####
## 4. FUNCTION OF CASE-CROSSOVER ANALYSIS FOR SUMMER MONTHS (LAG01)                    ####
##-------------------------------------------------------------------------------------####

casecrs_lag01 <- function (status, id, confounder=NULL, lag, varper, cen=list(min=NULL,max=NULL,degree=
                           estpct, data){

  ## Input: status, id, confounder: variables used to define the formula applied to clogit() in the for
  ##                                case.status~exposure+confounder+strata(matched.set)
  ##                                status: case status, 1=case, 0=control
```

```
##                                id: ID for participants
##                                confounder: optional, vector of covariates to be included in the mo
##        lag: the maximum lag in the cross basis
##        varper: numeric vector of percentiles of the distribution of temperature for internal knots
##        cen: a list to define the centering temperature
##            - "min" and "max": optional, "TRUE" if the minimum or maximum mortality temperature to
##            - "degree": optional, numeric vector of temperature (?C)
##            - "pct": optional, numeric vector of the percentiles of temperature distribution
##        estpct: numeric vector of percentiles of temperature distributions for effect estimate comp


## Build cross-basis function of temperature and lags
## Note: (1) exposure-response: natural cubic spline with internal knots placed at percentile of the
##                              distribution as defined by "varper"
##       (2) lag-response: natural cubic spline

## 1. delete observations with NA in temperature
dat <- subset(data,rowSums(is.na(data[which(names(data)%in%paste0("temp_s",0:lag))]))==0)
## 2. extract matrix of temperature at lag0 to lag="lag"
mat_temp <- as.matrix(dplyr::select(dat,paste0("temp_s",0:lag)))
## 3. define basis for temperature
argvar <- list(fun="ns",knots=quantile(mat_temp,varper/100,na.rm=T))
## 4. define basis for lag
arglag <- list(fun="ns")
## 5. build the cross-basis function
cb_temp <- crossbasis(mat_temp,lag=c(0,lag),argvar=argvar,arglag = arglag)

## Different percentile of the temperature matrix
tper <- quantile(mat_temp,seq(0,100,1)/100)

## Temperature summary for case days
tsum_case <- summary(subset(dat,status==1)$temp_s0)
tsum_case["SD"] <- sd(subset(dat,status==1)$temp_s0)

## Temperature summary for control days
tsum_control <- summary(subset(dat,status==0)$temp_s0)
tsum_control["SD"] <- sd(subset(dat,status==0)$temp_s0)

## Conditional logistic regression ####
## Note: "clogit" function in the "survival" package (same output as "clogistic" in "Epi" package)

if (is.null(confounder)==F){
  fml <- as.formula(paste0(status,"~cb_temp+strata(",id,")+",paste0(confounder,collapse = "+")))
} else {
  fml <- as.formula(paste0(status,"~cb_temp+strata(",id,")"))
}
mod <- clogit(fml,data=dat)

# Reduction to overall cumulative (it is irrelevant the cen value)
red <- crossreduce(cb_temp, mod, cen = 20)
# Store reduced coefs
coef <- coef(red)
vcov <- vcov(red)
```

```r
## centering temperature
cen_temp <- NULL;cen_name <- NULL

if (is.null(cen$min)==F){
  cen_temp <- c(cen_temp,findmin(cb_temp,mod))
  cen_name <- c(cen_name,"min")
}
if (is.null(cen$max)==F){
  cen_temp <- c(cen_temp,findmax(cb_temp,mod))
  cen_name <- c(cen_name,"max")
}
if (is.null(cen$degree)==F){
  cen_temp <- c(cen_temp,cen$degree)
  cen_name <- c(cen_name,paste0(cen$degree," degree"))
}
if (is.null(cen$pct)==F){
  cen_temp <- c(cen_temp,quantile(mat_temp,cen$pct/100))
  cen_name <- c(cen_name,paste0(cen$pct,"th"))
}

## Predict ORs from each cen_temp to each estpct
estimate <- list()
for (k in 1:length(cen_temp)){
  pred <- crosspred(cb_temp, mod, model.link="logit", cen=cen_temp[k], at=quantile(mat_temp,estpct/100
  estimate[[k]] <-  round(data.frame(OR=pred$allRRfit,CIlow=pred$allRRlow,CIhigh=pred$allRRhigh),3)
  estimate[[k]]$temp <- as.numeric(rownames(estimate[[k]]))
  estimate[[k]]$perc <- paste0(estpct,"th")
  estimate[[k]]$cen <- cen_name[k]
  estimate[[k]] <- dplyr::select(estimate[[k]],c(cen,perc,temp,everything()))
}
estimate_all <- do.call(rbind,estimate)
rownames(estimate_all) <- NULL

## output:result, a list containing the following elements
##        - n_case: number of cases
##        - n_control: number of controls
##        - tper: temperature distribution (percentiles)
##        - tsum_case: summary of temperature on case days
##        - tsum_control: summary of temperature on case days
##        - coef: coefficients for the overall association
##        - vcov: variance-covariance of coefs for overall association
##        - estimate: OR and CI at the "estpct" percentile of temperature distribution compared to ea
##        output for plots
##        - mat_temp: matrix of temperature
##        - cb_temp: cross-basis of temperature
##        - model_coef: coefficients of conditional logistic regression model
##        - model_vcov: variance matrix of conditional logistic regression model

result <- NULL
result$n_case <- nrow(subset(dat,status==1))
result$n_control <- nrow(subset(dat,status==0))
result$tper <- tper
result$tsum_case <- tsum_case
```

```
  result$tsum_control <- tsum_control
  result$coef <- coef
  result$vcov <- vcov
  result$estimate <- estimate_all
  result$mat_temp <- mat_temp
  result$cb_temp <- cb_temp
  result$model_coef <- mod$coefficients
  result$model_vcov <- mod$var
  return(result)
}



##-------------------------------------------------------------------------------------------####
## 5. crossreduce_int FUNCTION FOR INTERACTION ANALYSIS: 2 CATEGORIES                        ####
## Adapted from crossreduce_int_2APcats by Kai
## https://github.com/CHENlab-Yale/Two-way_effect_modifications/blob/master/crossreduce_int_2APcats.R
##-------------------------------------------------------------------------------------------####
crossreduce_int_2cats <- function (basis, model = NULL, type = "overall", value = NULL,
                                   coef = NULL, vcov = NULL, model.link = NULL, at = NULL, from = NULL,
                                   to = NULL, by = NULL, lag, bylag = 1, cen = NULL, ci.level = 0.95)
{
  if (all(class(basis) != "crossbasis")) {
    stop("the first argument must be an object of class 'crossbasis'")
  }
  name <- deparse(substitute(basis))
  attr <- attributes(basis)
  if (ncol(basis) == 1)
    cond <- name
  if (is.null(model) && (is.null(coef) || is.null(vcov))) {
    stop("At least 'model' or 'coef'-'vcov' must be provided")
  }
  type <- match.arg(type, c("overall", "var", "lag"))
  if (type != "overall") {
    if (is.null(value))
      stop("'value' must be provided for type 'var' or 'lag'")
    else if (!is.numeric(value) || length(value) > 1) {
      stop("'value' must be a numeric scalar")
    }
    if (type == "lag" && (any(value < attr$lag[1]) || any(value >
                                                          attr$lag[2]))) {
      stop("'value' of lag-specific effects must be within the lag range")
    }
  } else value <- NULL
  lag <- attr$lag
  if (lag != attr$lag && attr$arglag$fun == "integer")
    stop("prediction for lag sub-period not allowed for type 'integer'")
  if (!is.numeric(ci.level) || ci.level >= 1 || ci.level <=
      0) {
    stop("'ci.level' must be numeric and between 0 and 1")
  }
  cond <- if (ncol(basis) == 1L) name else
    paste(name, "[[:print:]]*v[0-9]{1,2}\\.l[0-9]{1,2}", sep = "")
```

```
cond.modif_cat2 <-  paste(name, "[[:print:]]*v[0-9]{1,2}\\.l[0-9]{1,2}\\:modif_cat2",
                          sep = "")

if (!is.null(model)) {
  model.class <- class(model)
  coef <- dlnm:::getcoef(model, model.class)
  ind.all <- grep(cond, names(coef))
  ind.modif_cat2 <- grep(cond.modif_cat2, names(coef))
  ind.main <- ind.all[ind.all != ind.modif_cat2]

  ### Extract the coef and vcov from the interaction model for modif categories
  coef.main <- coef[ind.main]
  coef.int_cat2 <- coef[ind.modif_cat2]
  ##vcov for modif categories
  vcov.all <- dlnm:::getvcov(model, model.class)
  vcov.main <- dlnm:::getvcov(model, model.class)[ind.main, ind.main, drop = FALSE]
  vcov.int_cat2 <- dlnm:::getvcov(model, model.class)[ind.modif_cat2, ind.modif_cat2, drop = FALSE]

  #cat=1
  coef_modifcat1 <- coef.main
  vcov_modifcat1 <- vcov.main
  #cat=2
  coef_modifcat2 <- coef_modifcat1+coef.int_cat2
  ####Important!! note that for interaction analysis, vcov(b1*b2)=var(b1)+var(b2)+2cov(b1,b2)
  ####This is only correct for cov(b1, b2) == cov(b2, b1); otherwise(like here), must using cov(b1, b
  vcov_modifcat2 <- vcov_modifcat1+vcov.int_cat2+dlnm:::getvcov(model, model.class)[ind.main, ind.modi
    dlnm:::getvcov(model, model.class)[ind.modif_cat2, ind.main, drop=FALSE]

  #model.link
  model.link <- dlnm:::getlink(model, model.class)
}
else model.class <- NA
npar <- ncol(basis)
range <- attr$range
at <- dlnm:::mkat(at, from, to, by, range, lag, bylag)
cen <- dlnm:::mkcen(cen, type = "cb", basis, range)
attributes(basis)$argvar$cen <- attr$argvar$cen <- NULL
if (type == "overall") {
  lagbasis <- do.call("onebasis", c(list(x = dlnm:::seqlag(lag)),
                                    attr$arglag))
  M <- diag(ncol(basis)/ncol(lagbasis)) %x% (t(rep(1, diff(lag) +
                                                     1)) %*% lagbasis)
  newbasis <- do.call("onebasis", c(list(x = at), attr$argvar))
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
                                      attr$argvar))
    newbasis <- scale(newbasis, center = basiscen, scale = FALSE)
  }
}
else if (type == "lag") {
  lagbasis <- do.call("onebasis", c(list(x = value), attr$arglag))
  M <- diag(ncol(basis)/ncol(lagbasis)) %x% lagbasis
  newbasis <- do.call("onebasis", c(list(x = at), attr$argvar))
```

```
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
                                      attr$argvar))
    newbasis <- scale(newbasis, center = basiscen, scale = FALSE)
  }
}
else if (type == "var") {
  varbasis <- do.call("onebasis", c(list(x = value), attr$argvar))
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
                                      attr$argvar))
    varbasis <- scale(varbasis, center = basiscen, scale = FALSE)
  }
  M <- varbasis %x% diag(ncol(basis)/ncol(varbasis))
  newbasis <- do.call("onebasis", c(list(x = seqlag(lag,
                                                    bylag)), attr$arglag))
}
dimnames(newbasis) <- list(seq(nrow(newbasis)), paste0("b",
                                                       seq(ncol(newbasis))))
##cat=1
newcoef_modifcat1 <- as.vector(M %*% coef_modifcat1)
names(newcoef_modifcat1) <- colnames(newbasis)
newvcov_modifcat1 <- M %*% vcov_modifcat1 %*% t(M)
dimnames(newvcov_modifcat1) <- list(colnames(newbasis), colnames(newbasis))
fit_modifcat1 <- as.vector(newbasis %*% newcoef_modifcat1)
se_modifcat1 <- sqrt(pmax(0, rowSums((newbasis %*% newvcov_modifcat1) * newbasis)))

if (type == "var") {
  names(fit_modifcat1) <- names(se_modifcat1) <- outer("lag", seqlag(lag, bylag),
                                                       paste, sep = "")
}
else names(fit_modifcat1) <- names(se_modifcat1) <- at

##cat=2
newcoef_modifcat2 <- as.vector(M %*% coef_modifcat2)
names(newcoef_modifcat2) <- colnames(newbasis)
newvcov_modifcat2 <- M %*% vcov_modifcat2 %*% t(M)
dimnames(newvcov_modifcat2) <- list(colnames(newbasis), colnames(newbasis))
fit_modifcat2 <- as.vector(newbasis %*% newcoef_modifcat2)
se_modifcat2 <- sqrt(pmax(0, rowSums((newbasis %*% newvcov_modifcat2) * newbasis)))

if (type == "var") {
  names(fit_modifcat2) <- names(se_modifcat2) <- outer("lag", seqlag(lag, bylag),
                                                       paste, sep = "")
}
else names(fit_modifcat2) <- names(se_modifcat2) <- at


##result list
list <- list(coef_modifcat1 = newcoef_modifcat1, vcov_modifcat1 = newvcov_modifcat1,
             coef_modifcat2 = newcoef_modifcat2, vcov_modifcat2 = newvcov_modifcat2,
             basis = newbasis, type = type, value = value)
if (type != "var")
```

```
    list$predvar <- at
  if (!is.null(cen))
    list$cen <- cen
  list <- c(list, list(lag = lag, bylag = bylag, fit_modifcat1 = fit_modifcat1, se_modifcat1 = se_modif
                       fit_modifcat2 = fit_modifcat2, se_modifcat2 = se_modifcat2))

  z <- qnorm(1 - (1 - ci.level)/2)
  if (model.link %in% c("log", "logit")) {
    #cat=1
    list$RRfit_modifcat1 <- exp(fit_modifcat1)
    list$RRlow_modifcat1 <- exp(fit_modifcat1 - z * se_modifcat1)
    names(list$RRlow_modifcat1) <- names(fit_modifcat1)
    list$RRhigh_modifcat1 <- exp(fit_modifcat1 + z * se_modifcat1)
    names(list$RRhigh_modifcat1) <- names(fit_modifcat1)
    #cat=2
    list$RRfit_modifcat2 <- exp(fit_modifcat2)
    list$RRlow_modifcat2 <- exp(fit_modifcat2 - z * se_modifcat2)
    names(list$RRlow_modifcat2) <- names(fit_modifcat2)
    list$RRhigh_modifcat2 <- exp(fit_modifcat2 + z * se_modifcat2)
    names(list$RRhigh_modifcat2) <- names(fit_modifcat2)
  }
  else {
    #cat1
    list$low_modifcat1 <- fit_modifcat1 - z * se_modifcat1
    names(list$low_modifcat1) <- names(fit_modifcat1)
    list$high_modifcat1 <- fit_modifcat1 + z * se_modifcat1
    names(list$high_modifcat1) <- names(fit_modifcat1)
    #cat2
    list$low_modifcat2 <- fit_modifcat2 - z * se_modifcat2
    names(list$low_modifcat2) <- names(fit_modifcat2)
    list$high_modifcat2 <- fit_modifcat2 + z * se_modifcat2
    names(list$high_modifcat2) <- names(fit_modifcat2)
  }
  list$ci.level <- ci.level
  list$model.class <- model.class
  list$model.link <- model.link
  class(list) <- "crossreduce"
  return(list)
}




##-------------------------------------------------------------------------------------------####
## 6. FUNCTION OF CASE-CROSSOVER EFFECT MODIFICATION: 2 CATEGORIES                           ####
##-------------------------------------------------------------------------------------------####

casecrs_int_2cats <- function (status, id, modif, confounder=NULL, lag, varper, lagnk, cen=list(degree=
                               estpct, data){

  ## Input: status, id, modif_cat, confounder: variables used to define the formula applied to clogit()
  ##                                case.status~exposure*modif+confounder+strata(matched.set)
  ##                                status: case status, 1=case, 0=control
```

```
##                                id: ID for participants
##                                modif: effect modifier with two categories
##                                confounder: optional, vector of covariates to be included in the mo
##        lag: the maximum lag in the cross basis
##        varper: numeric vector of percentiles of the distribution of temperature for internal knots
##        lagnk: the number of internal knots in the lag-response dimension
##        cen: a list to define the centering temperature
##            - "degree": optional, numeric vector of temperature (?C)
##            - "pct": optional, numeric vector of the percentiles of temperature distribution
##        estpct: numeric vector of percentiles of temperature distributions for effect estimate comp


## Build cross-basis function of temperature and lags
## Note: (1) exposure-response: natural cubic spline with internal knots placed at percentile of the
##                              distribution as defined by "varper"
##       (2) lag-response: natural cubic spline with an intercept and n="lag" internal knots placed a
##                         equally spaced values on the log scale

## 1. delete observations with NA in temperature and effect modifier
dat <- subset(data,rowSums(is.na(data[which(names(data)%in%c(paste0("temp_s",0:lag),modif))]))==0)
## 2. define the effect modifier
dat$modif_cat <- dat[,modif]
dat$modif_cat <- as.factor(dat$modif_cat)
dat$modif_cat <- ifelse(dat$modif_cat==levels(dat$modif_cat)[1],1,2)
dat$modif_cat <- as.factor(dat$modif_cat)
## 3. extract matrix of temperature at lag0 to lag="lag"
mat_temp <- as.matrix(dplyr::select(dat,paste0("temp_s",0:lag)))
## 4. define basis for temperature
argvar <- list(fun="ns",knots=quantile(mat_temp,varper/100,na.rm=T))
## 5. define basis for lag
arglag <- list(fun="ns",knots=logknots(lag,lagnk))
## 6. build the cross-basis function
cb_temp <- crossbasis(mat_temp,lag=c(0,lag),argvar=argvar,arglag = arglag)
## 7. build the one-basis function for temperature
b1temp <- onebasis(mat_temp, fun="ns", knots=quantile(mat_temp,varper/100,na.rm=T))

## Different percentile of the temperature matrix
tper <- quantile(mat_temp,seq(0,100,1)/100)

## Temperature summary for case days in the first subgroup
tsum_cat1 <- summary(subset(dat,status==1&modif_cat==1)$temp_s0)
tsum_cat1["SD"] <- sd(subset(dat,status==1&modif_cat==1)$temp_s0)

## Temperature summary for case days in the second subgroup
tsum_cat2 <- summary(subset(dat,status==1&modif_cat==2)$temp_s0)
tsum_cat2["SD"] <- sd(subset(dat,status==1&modif_cat==2)$temp_s0)

## Conditional logistic regression ####
## Note: "clogit" function in the "survival" package (same output as "clogistic" in "Epi" package)

if (is.null(confounder)==F){
  fml <- as.formula(paste0(status,"~cb_temp*modif_cat+strata(",id,")+",paste0(confounder,collapse = "+
} else {
```

```
  fml <- as.formula(paste0(status,"~cb_temp*modif_cat+strata(",id,")"))
}

mod <- try(clogit(fml,data=dat), silent=TRUE)

if (class(mod)[1]!="try-error"){
  # Reduction to overall cumulative (it is irrelevant the cen value)
  red <- crossreduce_int_2cats(cb_temp, mod, cen = 20)
  # Store reduced coefs
  #cat1
  coef.cat1 <- red$coef_modifcat1
  vcov.cat1 <- red$vcov_modifcat1
  #cat2
  coef.cat2 <- red$coef_modifcat2
  vcov.cat2 <- red$vcov_modifcat2

  ## centering temperature
  cen_temp <- NULL;cen_name <- NULL

  if (is.null(cen$degree)==F){
    cen_temp <- c(cen_temp,cen$degree)
    cen_name <- c(cen_name,paste0(cen$degree," degree"))
  }
  if (is.null(cen$pct)==F){
    cen_temp <- c(cen_temp,quantile(mat_temp,cen$pct/100))
    cen_name <- c(cen_name,paste0(cen$pct,"th"))
  }

  ## Predict ORs from each cen_temp to each estpct for each subgroup
  estimate_cat1 <- list()
  for (k in 1:length(cen_temp)){
    pred_cat1 <- crosspred(b1temp, coef=coef.cat1, vcov=vcov.cat1, model.link="logit",cen=cen_temp[k]
    estimate_cat1[[k]] <-  round(data.frame(OR=pred_cat1$allRRfit,CIlow=pred_cat1$allRRlow,CIhigh=pre
    estimate_cat1[[k]]$temp <- as.numeric(rownames(estimate_cat1[[k]]))
    estimate_cat1[[k]]$perc <- paste0(estpct,"th")
    estimate_cat1[[k]]$cen <- cen_name[k]
    estimate_cat1[[k]] <- dplyr::select(estimate_cat1[[k]],c(cen,perc,temp,everything()))
  }
  estimate_cat1_all <- do.call(rbind,estimate_cat1)
  rownames(estimate_cat1_all) <- NULL

  estimate_cat2 <- list()
  for (k in 1:length(cen_temp)){
    pred_cat2 <- crosspred(b1temp, coef=coef.cat2, vcov=vcov.cat2, model.link="logit", cen=cen_temp[k]
    estimate_cat2[[k]] <-  round(data.frame(OR=pred_cat2$allRRfit,CIlow=pred_cat2$allRRlow,CIhigh=pre
    estimate_cat2[[k]]$temp <- as.numeric(rownames(estimate_cat2[[k]]))
    estimate_cat2[[k]]$perc <- paste0(estpct,"th")
    estimate_cat2[[k]]$cen <- cen_name[k]
    estimate_cat2[[k]] <- dplyr::select(estimate_cat2[[k]],c(cen,perc,temp,everything()))
  }
  estimate_cat2_all <- do.call(rbind,estimate_cat2)
  rownames(estimate_cat2_all) <- NULL
```

```
    ## output:result, a list containing the following elements
    ##        - n_cat1: number of participants in the 1st subgroup
    ##        - n_cat2: number of participants in the 2nd subgroup
    ##        - tper: temperature distribution (percentiles)
    ##        - tsum_cat1: summary of temperature on case days for the 1st subgroup
    ##        - tsum_cat2: summary of temperature on case days for the 2nd subgroup
    ##        - coef_cat1: coefficients for the overall association for the 1st subgroup
    ##        - coef_cat2: coefficients for the overall association for the 2nd subgroup
    ##        - vcov_cat1: variance-covariance of coefs for overall association for the 1st subgroup
    ##        - vcov_cat2: variance-covariance of coefs for overall association for the 2nd subgroup
    ##        - estimate_cat1: OR and CI at the "estpct" percentile of temperature distribution compare
    ##        - estimate_cat2: OR and CI at the "estpct" percentile of temperature distribution compare
    ##        output for plots
    ##        - mat_temp: matrix of temperature
    ##        - b1temp: one-basis of temperature

    result <- NULL
    result$n_cat1 <- nrow(subset(dat,modif_cat==1))
    result$n_cat2 <- nrow(subset(dat,modif_cat==2))
    result$tper <- tper
    result$tsum_cat1 <- tsum_cat1
    result$tsum_cat2 <- tsum_cat2
    result$coef_cat1 <- coef.cat1
    result$vcov_cat1 <- vcov.cat1
    result$coef_cat2 <- coef.cat2
    result$vcov_cat2 <- vcov.cat2
    result$estimate_cat1 <- estimate_cat1_all
    result$estimate_cat2 <- estimate_cat2_all
    result$mat_temp <- mat_temp
    result$b1temp <- b1temp
    return(result)
  } else {
    result <- NULL
    result$n_cat1 <- nrow(subset(dat,modif_cat==1))
    result$n_cat2 <- nrow(subset(dat,modif_cat==2))
    result$tper <- tper
    result$tsum_cat1 <- tsum_cat1
    result$tsum_cat2 <- tsum_cat2
    result$coef_cat1 <- NA
    result$vcov_cat1 <- NA
    result$coef_cat2 <- NA
    result$vcov_cat2 <- NA
    result$estimate_cat1 <- NA
    result$estimate_cat2 <- NA
    result$mat_temp <- NA
    result$b1temp <- NA
    return(result)
  }
}



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

```r
## 7. crossreduce_int FUNCTION FOR INTERACTION ANALYSIS: 3 CATEGORIES                    ####
## Adapted from crossreduce_int_2APcats by Kai
## https://github.com/CHENlab-Yale/Two-way_effect_modifications/blob/master/crossreduce_int_2APcats.R
##-------------------------------------------------------------------------------------------####
crossreduce_int_3cats <- function (basis, model = NULL, type = "overall", value = NULL,
                                   coef = NULL, vcov = NULL, model.link = NULL, at = NULL, from = NULL,
                                   to = NULL, by = NULL, lag, bylag = 1, cen = NULL, ci.level = 0.95)
{
  if (all(class(basis) != "crossbasis")) {
    stop("the first argument must be an object of class 'crossbasis'")
  }
  name <- deparse(substitute(basis))
  attr <- attributes(basis)
  if (ncol(basis) == 1)
    cond <- name
  if (is.null(model) && (is.null(coef) || is.null(vcov))) {
    stop("At least 'model' or 'coef'-'vcov' must be provided")
  }
  type <- match.arg(type, c("overall", "var", "lag"))
  if (type != "overall") {
    if (is.null(value))
      stop("'value' must be provided for type 'var' or 'lag'")
    else if (!is.numeric(value) || length(value) > 1) {
      stop("'value' must be a numeric scalar")
    }
    if (type == "lag" && (any(value < attr$lag[1]) || any(value >
                                                          attr$lag[2]))) {
      stop("'value' of lag-specific effects must be within the lag range")
    }
  } else value <- NULL
  lag <- attr$lag
  if (lag != attr$lag && attr$arglag$fun == "integer")
    stop("prediction for lag sub-period not allowed for type 'integer'")
  if (!is.numeric(ci.level) || ci.level >= 1 || ci.level <=
      0) {
    stop("'ci.level' must be numeric and between 0 and 1")
  }
  cond <- if (ncol(basis) == 1L) name else
    paste(name, "[[:print:]]*v[0-9]{1,2}\\.l[0-9]{1,2}", sep = "")

  cond.modif_cat2 <-  paste(name, "[[:print:]]*v[0-9]{1,2}\\.l[0-9]{1,2}\\:modif_cat2",
                            sep = "")

  cond.modif_cat3 <-  paste(name, "[[:print:]]*v[0-9]{1,2}\\.l[0-9]{1,2}\\:modif_cat3",
                            sep = "")

  if (!is.null(model)) {
    model.class <- class(model)
    coef <- dlnm:::getcoef(model, model.class)
    ind.all <- grep(cond, names(coef))
    ind.modif_cat2 <- grep(cond.modif_cat2, names(coef))
    ind.modif_cat3 <- grep(cond.modif_cat3, names(coef))
    ind.main <- ind.all[ind.all != ind.modif_cat2&ind.all != ind.modif_cat3]
```

```
  ### Extract the coef and vcov from the interaction model for modif categories
  coef.main <- coef[ind.main]
  coef.int_cat2 <- coef[ind.modif_cat2]
  coef.int_cat3 <- coef[ind.modif_cat3]

  ##vcov for modif categories
  vcov.all <- dlnm:::getvcov(model, model.class)
  vcov.main <- dlnm:::getvcov(model, model.class)[ind.main, ind.main, drop = FALSE]
  vcov.int_cat2 <- dlnm:::getvcov(model, model.class)[ind.modif_cat2, ind.modif_cat2, drop = FALSE]
  vcov.int_cat3 <- dlnm:::getvcov(model, model.class)[ind.modif_cat3, ind.modif_cat3, drop = FALSE]

  #cat=1
  coef_modifcat1 <- coef.main
  vcov_modifcat1 <- vcov.main
  #cat=2
  coef_modifcat2 <- coef_modifcat1+coef.int_cat2
  ####Important!! note that for interaction analysis, vcov(b1*b2)=var(b1)+var(b2)+2cov(b1,b2)
  ####This is only correct for cov(b1, b2) == cov(b2, b1); otherwise(like here), must using cov(b1, b
  vcov_modifcat2 <- vcov_modifcat1+vcov.int_cat2+dlnm:::getvcov(model, model.class)[ind.main, ind.modi
    dlnm:::getvcov(model, model.class)[ind.modif_cat2, ind.main, drop=FALSE]
  #cat=3
  coef_modifcat3 <- coef_modifcat1+coef.int_cat3
  ####Important!! note that for interaction analysis, vcov(b1*b2)=var(b1)+var(b2)+2cov(b1,b2)
  ####This is only correct for cov(b1, b2) == cov(b2, b1); otherwise(like here), must using cov(b1, b
  vcov_modifcat3 <- vcov_modifcat1+vcov.int_cat3+dlnm:::getvcov(model, model.class)[ind.main, ind.modi
    dlnm:::getvcov(model, model.class)[ind.modif_cat3, ind.main, drop=FALSE]
  #model.link
  model.link <- dlnm:::getlink(model, model.class)
}
else model.class <- NA
npar <- ncol(basis)
range <- attr$range
at <- dlnm:::mkat(at, from, to, by, range, lag, bylag)
cen <- dlnm:::mkcen(cen, type = "cb", basis, range)
attributes(basis)$argvar$cen <- attr$argvar$cen <- NULL
if (type == "overall") {
  lagbasis <- do.call("onebasis", c(list(x = dlnm:::seqlag(lag)),
                                    attr$arglag))
  M <- diag(ncol(basis)/ncol(lagbasis)) %x% (t(rep(1, diff(lag) +
                                                     1)) %*% lagbasis)
  newbasis <- do.call("onebasis", c(list(x = at), attr$argvar))
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
                                      attr$argvar))
    newbasis <- scale(newbasis, center = basiscen, scale = FALSE)
  }
}
else if (type == "lag") {
  lagbasis <- do.call("onebasis", c(list(x = value), attr$arglag))
  M <- diag(ncol(basis)/ncol(lagbasis)) %x% lagbasis
  newbasis <- do.call("onebasis", c(list(x = at), attr$argvar))
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
```

```
                                        attr$argvar))
    newbasis <- scale(newbasis, center = basiscen, scale = FALSE)
  }
}
else if (type == "var") {
  varbasis <- do.call("onebasis", c(list(x = value), attr$argvar))
  if (!is.null(cen)) {
    basiscen <- do.call("onebasis", c(list(x = cen),
                                        attr$argvar))
    varbasis <- scale(varbasis, center = basiscen, scale = FALSE)
  }
  M <- varbasis %x% diag(ncol(basis)/ncol(varbasis))
  newbasis <- do.call("onebasis", c(list(x = seqlag(lag,
                                                    bylag)), attr$arglag))
}
dimnames(newbasis) <- list(seq(nrow(newbasis)), paste0("b",
                                                       seq(ncol(newbasis))))
##cat=1
newcoef_modifcat1 <- as.vector(M %*% coef_modifcat1)
names(newcoef_modifcat1) <- colnames(newbasis)
newvcov_modifcat1 <- M %*% vcov_modifcat1 %*% t(M)
dimnames(newvcov_modifcat1) <- list(colnames(newbasis), colnames(newbasis))
fit_modifcat1 <- as.vector(newbasis %*% newcoef_modifcat1)
se_modifcat1 <- sqrt(pmax(0, rowSums((newbasis %*% newvcov_modifcat1) * newbasis)))

if (type == "var") {
  names(fit_modifcat1) <- names(se_modifcat1) <- outer("lag", seqlag(lag, bylag),
                                                       paste, sep = "")
}
else names(fit_modifcat1) <- names(se_modifcat1) <- at

##cat=2
newcoef_modifcat2 <- as.vector(M %*% coef_modifcat2)
names(newcoef_modifcat2) <- colnames(newbasis)
newvcov_modifcat2 <- M %*% vcov_modifcat2 %*% t(M)
dimnames(newvcov_modifcat2) <- list(colnames(newbasis), colnames(newbasis))
fit_modifcat2 <- as.vector(newbasis %*% newcoef_modifcat2)
se_modifcat2 <- sqrt(pmax(0, rowSums((newbasis %*% newvcov_modifcat2) * newbasis)))

if (type == "var") {
  names(fit_modifcat2) <- names(se_modifcat2) <- outer("lag", seqlag(lag, bylag),
                                                       paste, sep = "")
}
else names(fit_modifcat2) <- names(se_modifcat2) <- at

##cat=3
newcoef_modifcat3 <- as.vector(M %*% coef_modifcat3)
names(newcoef_modifcat3) <- colnames(newbasis)
newvcov_modifcat3 <- M %*% vcov_modifcat3 %*% t(M)
dimnames(newvcov_modifcat3) <- list(colnames(newbasis), colnames(newbasis))
fit_modifcat3 <- as.vector(newbasis %*% newcoef_modifcat3)
se_modifcat3 <- sqrt(pmax(0, rowSums((newbasis %*% newvcov_modifcat3) * newbasis)))
```

```
if (type == "var") {
  names(fit_modifcat3) <- names(se_modifcat3) <- outer("lag", seqlag(lag, bylag),
                                                       paste, sep = "")
}
else names(fit_modifcat3) <- names(se_modifcat3) <- at

##result list
list <- list(coef_modifcat1 = newcoef_modifcat1, vcov_modifcat1 = newvcov_modifcat1,
             coef_modifcat2 = newcoef_modifcat2, vcov_modifcat2 = newvcov_modifcat2,
             coef_modifcat3 = newcoef_modifcat3, vcov_modifcat3 = newvcov_modifcat3,
             basis = newbasis, type = type, value = value)
if (type != "var")
  list$predvar <- at
if (!is.null(cen))
  list$cen <- cen
list <- c(list, list(lag = lag, bylag = bylag, fit_modifcat1 = fit_modifcat1, se_modifcat1 = se_modif
                     fit_modifcat2 = fit_modifcat2, se_modifcat2 = se_modifcat2,
                     fit_modifcat3 = fit_modifcat3, se_modifcat3 = se_modifcat3))

z <- qnorm(1 - (1 - ci.level)/2)
if (model.link %in% c("log", "logit")) {
  #cat=1
  list$RRfit_modifcat1 <- exp(fit_modifcat1)
  list$RRlow_modifcat1 <- exp(fit_modifcat1 - z * se_modifcat1)
  names(list$RRlow_modifcat1) <- names(fit_modifcat1)
  list$RRhigh_modifcat1 <- exp(fit_modifcat1 + z * se_modifcat1)
  names(list$RRhigh_modifcat1) <- names(fit_modifcat1)
  #cat=2
  list$RRfit_modifcat2 <- exp(fit_modifcat2)
  list$RRlow_modifcat2 <- exp(fit_modifcat2 - z * se_modifcat2)
  names(list$RRlow_modifcat2) <- names(fit_modifcat2)
  list$RRhigh_modifcat2 <- exp(fit_modifcat2 + z * se_modifcat2)
  names(list$RRhigh_modifcat2) <- names(fit_modifcat2)
  #cat=3
  list$RRfit_modifcat3 <- exp(fit_modifcat3)
  list$RRlow_modifcat3 <- exp(fit_modifcat3 - z * se_modifcat3)
  names(list$RRlow_modifcat3) <- names(fit_modifcat3)
  list$RRhigh_modifcat3 <- exp(fit_modifcat3 + z * se_modifcat3)
  names(list$RRhigh_modifcat3) <- names(fit_modifcat3)
}
else {
  #cat1
  list$low_modifcat1 <- fit_modifcat1 - z * se_modifcat1
  names(list$low_modifcat1) <- names(fit_modifcat1)
  list$high_modifcat1 <- fit_modifcat1 + z * se_modifcat1
  names(list$high_modifcat1) <- names(fit_modifcat1)
  #cat2
  list$low_modifcat2 <- fit_modifcat2 - z * se_modifcat2
  names(list$low_modifcat2) <- names(fit_modifcat2)
  list$high_modifcat2 <- fit_modifcat2 + z * se_modifcat2
  names(list$high_modifcat2) <- names(fit_modifcat2)
  #cat3
  list$low_modifcat3 <- fit_modifcat3 - z * se_modifcat3
```

```
    names(list$low_modifcat3) <- names(fit_modifcat3)
    list$high_modifcat3 <- fit_modifcat3 + z * se_modifcat3
    names(list$high_modifcat3) <- names(fit_modifcat3)
  }
  list$ci.level <- ci.level
  list$model.class <- model.class
  list$model.link <- model.link
  class(list) <- "crossreduce"
  return(list)
}


##-------------------------------------------------------------------------------------####
## 8. FUNCTION OF CASE-CROSSOVER EFFECT MODIFICATION: 3 CATEGORIES                      ####
##-------------------------------------------------------------------------------------####

casecrs_int_3cats <- function (status, id, modif, confounder=NULL, lag, varper, lagnk, cen=list(degree=
                               estpct, data){

  ## Input: status, id, modif_cat, confounder: variables used to define the formula applied to clogit()
  ##                                case.status~exposure*modif+confounder+strata(matched.set)
  ##                                status: case status, 1=case, 0=control
  ##                                id: ID for participants
  ##                                modif: effect modifier with three categories
  ##                                confounder: optional, vector of covariates to be included in the mo
  ##        lag: the maximum lag in the cross basis
  ##        varper: numeric vector of percentiles of the distribution of temperature for internal knots
  ##        lagnk: the number of internal knots in the lag-response dimension
  ##        cen: a list to define the centering temperature
  ##            - "degree": optional, numeric vector of temperature (?C)
  ##            - "pct": optional, numeric vector of the percentiles of temperature distribution
  ##        estpct: numeric vector of percentiles of temperature distributions for effect estimate comp


  ## Build cross-basis function of temperature and lags
  ## Note: (1) exposure-response: natural cubic spline with internal knots placed at percentile of the
  ##                              distribution as defined by "varper"
  ##       (2) lag-response: natural cubic spline with an intercept and n="lag" internal knots placed a
  ##                         equally spaced values on the log scale

  ## 1. delete observations with NA in temperature and effect modifier
  dat <- subset(data,rowSums(is.na(data[which(names(data)%in%c(paste0("temp_s",0:lag),modif))]))==0)
  ## 2. define the effect modifier
  dat$modif_cat <- dat[,modif]
  dat$modif_cat <- as.factor(dat$modif_cat)
  dat$modif_cat <- ifelse(dat$modif_cat==levels(dat$modif_cat)[1],1,ifelse(dat$modif_cat==levels(dat$mo
  dat$modif_cat <- as.factor(dat$modif_cat)
  ## 3. extract matrix of temperature at lag0 to lag="lag"
  mat_temp <- as.matrix(dplyr::select(dat,paste0("temp_s",0:lag)))
  ## 4. define basis for temperature
  argvar <- list(fun="ns",knots=quantile(mat_temp,varper/100,na.rm=T))
  ## 5. define basis for lag
  arglag <- list(fun="ns",knots=logknots(lag,lagnk))
```

```r
## 6. build the cross-basis function
cb_temp <- crossbasis(mat_temp,lag=c(0,lag),argvar=argvar,arglag = arglag)
## 7. build the one-basis function for temperature
b1temp <- onebasis(mat_temp, fun="ns", knots=quantile(mat_temp,varper/100,na.rm=T))

## Different percentile of the temperature matrix
tper <- quantile(mat_temp,seq(0,100,1)/100)

## Temperature summary for case days in the first subgroup
tsum_cat1 <- summary(subset(dat,status==1&modif_cat==1)$temp_s0)
tsum_cat1["SD"] <- sd(subset(dat,status==1&modif_cat==1)$temp_s0)

## Temperature summary for case days in the second subgroup
tsum_cat2 <- summary(subset(dat,status==1&modif_cat==2)$temp_s0)
tsum_cat2["SD"] <- sd(subset(dat,status==1&modif_cat==2)$temp_s0)

## Temperature summary for case days in the third subgroup
tsum_cat3 <- summary(subset(dat,status==1&modif_cat==3)$temp_s0)
tsum_cat3["SD"] <- sd(subset(dat,status==1&modif_cat==3)$temp_s0)

## Conditional logistic regression ####
## Note: "clogit" function in the "survival" package (same output as "clogistic" in "Epi" package)

if (is.null(confounder)==F){
  fml <- as.formula(paste0(status,"~cb_temp*modif_cat+strata(",id,")+",paste0(confounder,collapse = "+
} else {
  fml <- as.formula(paste0(status,"~cb_temp*modif_cat+strata(",id,")"))
}

mod <- try(clogit(fml,data=dat), silent=TRUE)

if (class(mod)[1]!="try-error"){
  # Reduction to overall cumulative (it is irrelevant the cen value)
  red <- crossreduce_int_3cats(cb_temp, mod, cen = 20)
  # Store reduced coefs
  #cat1
  coef.cat1 <- red$coef_modifcat1
  vcov.cat1 <- red$vcov_modifcat1
  #cat2
  coef.cat2 <- red$coef_modifcat2
  vcov.cat2 <- red$vcov_modifcat2
  #cat3
  coef.cat3 <- red$coef_modifcat3
  vcov.cat3 <- red$vcov_modifcat3

  ## centering temperature
  cen_temp <- NULL;cen_name <- NULL

  if (is.null(cen$degree)==F){
    cen_temp <- c(cen_temp,cen$degree)
    cen_name <- c(cen_name,paste0(cen$degree," degree"))
  }
  if (is.null(cen$pct)==F){
```

```
  cen_temp <- c(cen_temp,quantile(mat_temp,cen$pct/100))
  cen_name <- c(cen_name,paste0(cen$pct,"th"))
}

## Predict ORs from each cen_temp to each estpct for each subgroup
estimate_cat1 <- list()
for (k in 1:length(cen_temp)){
  pred_cat1 <- crosspred(b1temp, coef=coef.cat1, vcov=vcov.cat1, model.link="logit",cen=cen_temp[k]
  estimate_cat1[[k]] <-  round(data.frame(OR=pred_cat1$allRRfit,CIlow=pred_cat1$allRRlow,CIhigh=pred
  estimate_cat1[[k]]$temp <- as.numeric(rownames(estimate_cat1[[k]]))
  estimate_cat1[[k]]$perc <- paste0(estpct,"th")
  estimate_cat1[[k]]$cen <- cen_name[k]
  estimate_cat1[[k]] <- dplyr::select(estimate_cat1[[k]],c(cen,perc,temp,everything()))
}
estimate_cat1_all <- do.call(rbind,estimate_cat1)
rownames(estimate_cat1_all) <- NULL

estimate_cat2 <- list()
for (k in 1:length(cen_temp)){
  pred_cat2 <- crosspred(b1temp, coef=coef.cat2, vcov=vcov.cat2, model.link="logit", cen=cen_temp[k
  estimate_cat2[[k]] <-  round(data.frame(OR=pred_cat2$allRRfit,CIlow=pred_cat2$allRRlow,CIhigh=pred
  estimate_cat2[[k]]$temp <- as.numeric(rownames(estimate_cat2[[k]]))
  estimate_cat2[[k]]$perc <- paste0(estpct,"th")
  estimate_cat2[[k]]$cen <- cen_name[k]
  estimate_cat2[[k]] <- dplyr::select(estimate_cat2[[k]],c(cen,perc,temp,everything()))
}
estimate_cat2_all <- do.call(rbind,estimate_cat2)
rownames(estimate_cat2_all) <- NULL

estimate_cat3 <- list()
for (k in 1:length(cen_temp)){
  pred_cat3 <- crosspred(b1temp, coef=coef.cat3, vcov=vcov.cat3, model.link="logit", cen=cen_temp[k
  estimate_cat3[[k]] <-  round(data.frame(OR=pred_cat3$allRRfit,CIlow=pred_cat3$allRRlow,CIhigh=pred
  estimate_cat3[[k]]$temp <- as.numeric(rownames(estimate_cat3[[k]]))
  estimate_cat3[[k]]$perc <- paste0(estpct,"th")
  estimate_cat3[[k]]$cen <- cen_name[k]
  estimate_cat3[[k]] <- dplyr::select(estimate_cat3[[k]],c(cen,perc,temp,everything()))
}
estimate_cat3_all <- do.call(rbind,estimate_cat3)
rownames(estimate_cat3_all) <- NULL

## output:result, a list containing the following elements
##        - n_cat1: number of participants in the 1st subgroup
##        - n_cat2: number of participants in the 2nd subgroup
##        - tper: temperature distribution (percentiles)
##        - tsum_cat1: summary of temperature on case days for the 1st subgroup
##        - tsum_cat2: summary of temperature on case days for the 2nd subgroup
##        - coef_cat1: coefficients for the overall association for the 1st subgroup
##        - coef_cat2: coefficients for the overall association for the 2nd subgroup
##        - vcov_cat1: variance-covariance of coefs for overall association for the 1st subgroup
##        - vcov_cat2: variance-covariance of coefs for overall association for the 2nd subgroup
##        - estimate_cat1: OR and CI at the "estpct" percentile of temperature distribution compare
##        - estimate_cat2: OR and CI at the "estpct" percentile of temperature distribution compare
```

```
    ##        output for plots
    ##        - mat_temp: matrix of temperature
    ##        - b1temp: one-basis of temperature

    result <- NULL
    result$n_cat1 <- nrow(subset(dat,modif_cat==1))
    result$n_cat2 <- nrow(subset(dat,modif_cat==2))
    result$n_cat3 <- nrow(subset(dat,modif_cat==3))
    result$tper <- tper
    result$tsum_cat1 <- tsum_cat1
    result$tsum_cat2 <- tsum_cat2
    result$tsum_cat3 <- tsum_cat3
    result$coef_cat1 <- coef.cat1
    result$vcov_cat1 <- vcov.cat1
    result$coef_cat2 <- coef.cat2
    result$vcov_cat2 <- vcov.cat2
    result$coef_cat3 <- coef.cat3
    result$vcov_cat3 <- vcov.cat3
    result$estimate_cat1 <- estimate_cat1_all
    result$estimate_cat2 <- estimate_cat2_all
    result$estimate_cat3 <- estimate_cat3_all
    result$mat_temp <- mat_temp
    result$b1temp <- b1temp
    return(result)
  } else {
    result <- NULL
    result$n_cat1 <- nrow(subset(dat,modif_cat==1))
    result$n_cat2 <- nrow(subset(dat,modif_cat==2))
    result$n_cat3 <- nrow(subset(dat,modif_cat==3))
    result$tper <- tper
    result$tsum_cat1 <- tsum_cat1
    result$tsum_cat2 <- tsum_cat2
    result$tsum_cat3 <- tsum_cat3
    result$coef_cat1 <- NA
    result$vcov_cat1 <- NA
    result$coef_cat2 <- NA
    result$vcov_cat2 <- NA
    result$coef_cat3 <- NA
    result$vcov_cat3 <- NA
    result$estimate_cat1 <- NA
    result$estimate_cat2 <- NA
    result$estimate_cat3 <- NA
    result$mat_temp <- NA
    result$b1temp <- NA
    return(result)
  }
}


## 9. Description of continuous variables ####
desc_con <- function(x0,digit){
  x <- x0[!is.na(x0)]
  nmin <- length(x)
```

```
  nall<-length(as.vector(x0))
  percent<-round((nall-nmin)/nall*100,digits=1)
  mean<-round(mean(x),digits=digit)
  sd <- round(sd(x),digits=digit)
  p25 <- round(quantile(x,probs=0.25),digits=digit)
  p50 <- round(quantile(x,probs=0.50),digits=digit)
  p75 <- round(quantile(x,probs=0.75),digits=digit)
  iqr <- round(p75-p25,digits=digit)
  min <- round(min(x),digits=digit)
  max <- round(max(x),digits=digit)
  out <- data.frame(N=nmin,Missing=percent,mean=mean,SD= sd,min=min,p25=p25,median=p50,p75=p75,max=max,
  out
}

## 10. Description of categorical variables ####
desc_cat <- function(x0) {
  x <- x0[!is.na(x0)]
  nmin <- length(x)
  nall<-length(as.vector(x0))
  nmiss <- nall - nmin
  percentmiss<-round(nmiss/nall*100,digits=1)

  n <- c(table(x),"NA"=nmiss)
  percent <- round(prop.table(table(x))*100,1) # ATTENTION: Percentages are calculated for reduced N!
  perc<-c(percent,percentmiss)
  out <- cbind(n,perc)
  out
}
```

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway