Heatwaves and clinical vulnerability in England; development of a risk stratification tool for use in primary care

Ross Thompson

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy
of the
University of London

December 2024

Department of Public Health, Environments and Society
Faculty of Public Health and Policy

London School Of Hygiene & Tropical Medicine, University of London

Funded by National Institute for Health Research (NIHR) Health Protection Research Unit in Environmental Change and Health (NIHR200909)

Research group affiliation(s): Environmental Health Research Group

Declaration of own work

I, Ross Thompson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Ross Thompson

Date:19 December 2024

Abstract

Background: A key action outlined within Heat-Health Action Plans is to identify at-risk individuals for targeted interventions, but this is rarely done, partially due to limited understanding of specific individual-level risk factors. Identifying those most vulnerable to heat is challenging due to the complex nature of the risks involved. Previous efforts to map population vulnerability have been of limited use for public health. In England, electronic health records have been successfully used in other health areas to develop risk prediction tools. Aim of this thesis was to assess the feasibility of developing a risk stratification tool to identify those at high risk that could be used within primary care settings.

Method: A time-stratified case-crossover analysis using conditional logistic regression was performed on 37 clinical risk factors and 9 socio-environmental factors using data from the Clinical Practice Research Datalink. Two Random Forest (RF) models were then developed in an attempt to identify individuals at risk.

Results: Results indicate that heat mortality risk is significantly affected by various chronic conditions and medications. A range of socio-environmental factors further influenced risk. The RF models, however, performed poorly overall, though they identified age and circulatory diseases as the most important predictors.

Conclusion: The study demonstrates that clinical records alone are insufficient for accurately predicting individuals at risk of death during heatwaves. The poor performance of the RF models reflects the limitations of existing tools. Despite this, the study is significant as it is the first to comprehensively explore individual-level clinical and socio-environmental factors in heat risk using primary care records in England, highlight the importance of specific conditions and medications which need to be considered in patient management, suggest heat-risk should be considered in broader health policies, and highlight the role of socio-economic disadvantage in the unequal distribution of climate impacts.

.

Acknowledgements

The work presented in this thesis was made possible by the invaluable input and support of various individuals to whom I am immensely grateful.

First and foremost, I'd like to thank my supervision team; Sari Kovats, Emer O'Connell and Shakoor Hajat, for sharing their expertise and vast knowledge on this subject over the last three years. Their collective and individual guidance, advice and input into the development and delivery of this project and its outputs have helped ensure that this work is robust, relevant and impactful. I would also like to thank Luis Baptista Mieiro, Dagmar Zeuner and Krishnan Bhaskaran who have all provided expert input over the life course of this project at different points which has been invaluable and greatly appreciated. I would also like to thank colleagues at UKHSA in the Extreme Events and Health Protection team specifically for their support while undertaking this work and for providing additional advice and guidance.

Several people have helped me with gaining access to the required data along the way, and I'd like to pay special thanks to the Electronic Health Records Group at LSHTM, specifically Helena Carreira and Patrick Bidulka for their support and guidance across the whole data management element of the project. I'd also like to thank Helen Macintyre for providing the essential temperature exposure data and Stephen O'Neil for providing a sounding board and offering valued advice and guidance on machine learning used in this project.

I am indebted to all the members of the PLANET group and members of the Greener Practise Network who gave their time and contributed to discussions both formally and informally. Their insights were hugely beneficial.

I would also like to thank all of my family and friends for their patience and support over the last few years. Specifically, to my parents who offered encouragement and supported us financially when needed. This would not have been possible without them. Zoe, my wife – thank you for picking up my slack and for your continued support, without which I wouldn't have been able to undertake this challenge. Finally, I'd like to dedicate this thesis to the memory of my mother who sadly passed away in December 2023, and did not get to see me complete this project. Her support and encouragement were ever-present throughout my life, and I know that she would have been extremely proud that I have completed this challenge and achieved what we had talked about on many occasions over the years. I miss you.

Table of contents

List of tables

List of Figures

Abbreviations

AWHP	Adverse Weather and Health Plan
BMI	Body Mass Index
CCRA3	Third Climate Change Risk Assessment
CI	Confidence interval
CO_2	Carbon Dioxide
CPRD	Clinical Practise Research Datalink
CVD	Cardiovascular disease
DBP	Diastolic Blood Pressure
eFI	Electronic Frailty Index
EHR	Electronic Health Records
EPC	Energy Performance Certificate
GPs	General Practitioners
HES	Hospital Episode Statistics
HHA	Heat-Health Alerts
HHAP	Heat Health Action Plans

HVI	Heat Vulnerability Index
HWP	Heatwave Plan for England
ICD-10	International Classification of Diseases tenth revision
ICD-9	International Classification of Diseases ninth revision
IMD	Index of Multiple Deprivation
ML	Machine Learning
NHS	National Health Service
NO_x	Nitrous oxides
NSAIDs	Non-steroidal anti-inflammatory drugs
O_3	Ozone
OR	Odds ratio
PLANET	Public Led and Knowledge Engagement Team
PM_{10}	Particulate matter with diameter of 10microns or less
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REM	Relative effect modification index
RF	Random Forest
SBP	Systolic Blood Pressure

UK	United Kingdom
UKHSA	UK Health Security Agency
WHO	World Health Organization

Chapter 1 – Introduction

1.0 Heat, health and responding to heatwaves

High ambient temperatures, particularly heatwaves, pose a significant risk to health[1-8]. In response to increased ambient air temperatures, the body will attempt to maintain its core temperature via several physiological mechanisms which allow the body to thermoregulate. High temperatures can interfere with these mechanisms and/or push the body beyond its capacity which can lead to adverse health outcomes, including death.[9] Previous epidemiological analysis suggests that individuals at risk of heat-related mortality include older people, those with chronic medical conditions, those with alcohol or drug dependence, or other mental health issues that affect behaviour[10] and the homeless[11 12].

Heatwaves have been forecast to increase in frequency, duration, and intensity due to over a century of unabated anthropogenic emissions.[13] Recent observations of weather conditions indicate we have already begun to experience these changes to our climate with all regions seeing extremes, including the UK.[14] Associated with episodes of heat are measurable acute increases in mortality across the population, which has been observed in the UK and across the globe.[15-19]

The United Kingdom's (UK) Third Climate Change Risk Assessment (CCRA3) published in 2021 confirmed hot weather remains a risk to health for the UK population, and that current action is insufficient to address both current risk and the increased risk in the future. [13 20]

Summer 2022 saw the first ever Level 4 Heat-Health Alert/red extreme heat warning issued as temperatures breached 40°C for the first time.[21] This record-breaking summer was also associated with the highest heat-related excess mortality ever observed in England.[21] Indeed, heat-associated death estimates in England have remained elevated since summer 2020, with a clear increasing trend.[22]

Recognising the potential threat to public health, national and local governments have developed and implemented Heat Health Action Plans (HHAP) and associated Heat-Health Alerting (HHA) systems.[23] These outline a range of recommended actions that should be taken across the health and care sector, which includes informal care within the community, to prevent harm to health and wider society as a result of high temperatures. One of the core recommendations across these plans is to identify individuals who are at high risk of ill health as a result of heatwave conditions and to implement targeted public health interventions.[23]

In April 2022, the UK Health Security Agency (UKHSA) introduced the Adverse Weather and Health Plan (AWHP), which replaced the Heatwave Plan for England (HWP) and is one of the most mature systems of its kind globally. The AWHP recommends a range of actions for commissioners and providers of health and social care, national and local government agencies and for individuals, local communities and voluntary groups.[24] The recommended actions outlined are aligned with the different alert types, which signify the potential severity of the episode and the impacts that are likely to occur. “Establish methods to identify, alert and monitor individuals most at risk of heat-related illnesses on your caseload", is a key recommendation within the AWHP for commissioners and providers of health and social care services, including primary care.[24]

Evidence from the independent evaluation of the HWP[25] and stakeholder engagement undertaken by UKHSA suggests that the identification of vulnerable individuals is not widely undertaken by providers of health and social care.[26 27] There is a common view that those who are within care are by definition vulnerable and therefore this particular recommendation is redundant.[25] However, analysis of heat-associated mortality in England shows that all-cause mortality does increase during heatwaves across all settings (at home, in hospitals and care homes).[28] Evidence from Italy also suggests that the risk of death occurring within hospitals on hot days is highest for those in general medicine wards compared to those in more intensive care units.[29] This suggests that risk is not necessarily evenly distributed across all in-care settings and that there are individuals at risk in the community. There are perhaps many other factors that contribute to why this particular recommendation is not widely acted upon, such as the complex nature of the risk itself, including individual susceptibility, physiological reserve, environmental factors, such as the natural and built environment, and social and contextual factors, and a lack of evidence upon which to develop robust methodologies.

1.0.1 – Identification of vulnerable individuals

The recent publication of the World Health Organization (WHO) heat and health evidence book provides an example of the coordination of national, regional, and local heat-health action in Italy.[23] Specifically, how regional and local plans should identify vulnerable subgroups within their population and activate active monitoring on those identified by local health services and GPs, and tailored response mechanisms for the different alert levels (e.g. establishing heat-health call centres).

In Italy, epidemiological studies of heatwave mortality[29 30] are used to develop lists of vulnerable individuals combined with local data on social markers of vulnerability (e.g. isolation, income etc) which are used to direct actions in the event of a heatwave as previously described. The WHO document suggests Italy is one of only a few countries to take this approach. In fact, no evidence of other examples have been identified from other countries that outline specific methodologies for identifying vulnerable populations.

Linked electronic health records have been used in England to develop and validate risk prediction tools for use by healthcare providers to identify vulnerable individuals.[31-33] Risk stratification has become an important tool in population health management. Stratifying patients by their level of risk allows practitioners to deliver the appropriate level of care and services to specific sub-groups. With the increasing availability of electronic health records, researchers and practitioners are developing tools which allow the identification of individuals considered at high risk for specific outcomes. Methodologies used range from simple ratios of risk factors present to machine learning.

While the specific methodologies used previously differ, they provide insight into the potential benefits of using primary care data to develop and validate a risk stratification tool for identifying those at high risk of death on hot days. Table 1.1 below briefly describes six specific tools developed and validated for use in primary care.

Table 1. 1 Examples of where primary care data was used to develop and validate risk stratification tools in England.

Name of risk stratification tool	**Methods used in development**	**Brief description and use in primary care**
Electronic Frailty Index (eFI)[32]	Retrospective cohort study using electronic health records. Hazard ratios were estimated using bivariate and multivariate Cox regression analysis. Discrimination of outcomes assessed using receiver operating characteristic curves and calibration using pseudo-R^2 estimates. eFI calculated as ratio of the present “deficits” against the total 36 deficits used in the assessment.	The tool involves a list of 36 deficits (or conditions, clinical measurements, social markers etc) which are all associated with frailty and is calculated by dividing the number of deficits present within an individual’s clinical records by the total number of deficits in the index (36). This ratio is then categorised in to Fit, Mildly Frail, Moderately Frail and Severely Frail. eFI is used within primary care to guide clinical care and is stipulated within with GP contract requirements.
QRISK3[31]	Cox proportional hazards models to derive risk equations for males and females, using a range of risk factors to predict incidents of cardiovascular disease. Validation measured via calibration and discrimination measures.	QRISK3 is the evolution of QRISK2, a widely used risk stratification tool that estimates risk of getting cardiovascular disease over a lifetime compared to an individual’s risk controlling for a range of risk factors. QRISK3 is used operationally during the

Name of risk stratification tool	Methods used in development	Brief description and use in primary care
		NHS Health Check which is aimed at adults aged 40 to 74.
Q-COVID[33]	The use of time-to-event models developed to derive risk equations for adults with the primary outcome being death or hospitalisation due to COVID-19. Risk model performance measured via prediction errors, calibration, and discrimination measures.	The COVID assessment tool (or screening algorithm) used observational primary care data (linked to ONS mortality registry, and COVID-19 test results) to derive which personal risk factors were associated with the two outcomes of interest and the strength of those associations with those outcomes. This was then used to establish a method by which each potential risk factor was ranked in relation to the weight of that factors influence on the outcomes of interest. This was then validated with further observational data to assess the predictive power of the approach. Based on the output of the analysis and validation process, clinicians used this approach to screen their patients and deliver specific individualised advice and guidance to those flagged as high risk from Autumn 2020.
Prediction of risk and risk factors of type 2 diabetes[34]	Machine learning, specifically Random Forest (RF) - versatile and useable algorithm that allows the user to generate accurate predictions of an outcome based on a	The use of RF was investigated as an appropriate method for predicting type 2 diabetes using 2 years follow-up electronic health data. Predictions were compared to multiple logistic

Name of risk stratification tool	Methods used in development	Brief description and use in primary care
	number of predictor variables used to train the model	regressions models, with RF more accurate at identifying the outcome of interest based on selected risk factors than traditional approaches. However, it is unclear how the output of this approach has been implemented in practice.
Cancer symptom screening tools[35]	Five studies assessing risk factors for different types of cancer, all using similar methods. Individual risk factors identified using univariable conditional logistic regression, with all variables found to be associated with the outcome of interest entered into multivariable analysis, likelihood ratio and positive predictive value calculated	The five studies used in this main study all used primary care data (either manual searching of primary care practice data or electronic health data). The output of the analysis is a form of risk quantification based on symptoms reported to primary care for specific types of cancer. Primary care practitioners can then use this output to guide their decision making on when they should refer their patients based on their potential quantified risk. It is unclear how this is used in practise.
Falls assessments[36]	Systematic review of literature to derive risk factors; expert consensus and ranking of individual factors based on their odds ratios and 95% Cis as reported in the included studies.	This tool has been widely used in the UK as a means to identify individuals which require assistance or require services to reduce risk of falling in their own homes. However, this was first published in 2004, and therefore is likely to be quite out of date today, and surpassed using eFI which is an assessment of general frailty that includes risk of falls.

1.0.2 – Why primary care is important for heat-health

Evidence in England suggests that those who are dying during heatwave episodes are largely not presenting to health care services. When examined in detail, heat morbidity outcomes across syndromic surveillance systems, suggest the impact on health service demand is minimal during the acute response phase of a heatwave.[37] This is potentially due to the fact the window of opportunity to act is minimal, with spikes in daily mortality observed within 0-24 hours of high temperatures occurring.[15-19] Additional deaths are consistently observed to occur across heatwaves episodes in hospitals, care homes and for those at home.[28 30] This raises several questions about the level of care being provided. Those within health and social care facilities will generally be in receipt of some level of care, while those at home may not.

Primary care is generally the first entry point into, and contact point with, the health and care system for many patients. Currently, the NHS report that 50% of GP consultations are related to existing conditions and evidence suggests that about 40% are amongst frequent attenders.[38] Meanwhile, the workload of primary care is also increasing as changes to the health care provision landscape continue to evolve as the UK population is both growing and aging, in addition to a shift to increasing, and complex care being provided at home and in the community. This last factor potentially increasing the number of patients at risk in their own homes and in care homes.

The core attributes of primary care suggest that primary care professionals are perhaps best placed to implement targeted and preventative interventions when episodes of heat do occur, given their role and knowledge of the local area, community and their patient's needs. However, due to the already high level of objectives set on primary care, effective strategies would need to be developed that would allow primary care professionals to easily identify those most at risk so that appropriate preventative actions could then be taken to reduce risk for those individuals.

1.0.3 – Lack of information on individual-level risk factors

There is a plethora of literature on heat-related mortality globally.[39-41] The majority of these are population-level studies that focus on factors which are well recorded on death certificates, such as age, sex and cause of death. However, there is significantly less high-quality epidemiological evidence on a range of other sub-groups such as those with chronic medical conditions, those with alcohol or drug dependence, or other mental health issues that affect behaviour[10] and the homeless[11 12]. Where evidence is available these are mostly based on ecological study designs. A limitation of these types of study are that the resultant associations observed may not translate to the individual level, as an individual's risk will be dependent on a range of interconnected factors that are specific to them, data on which is not available within this type of study. Where epidemiological studies are exploring individual heat risk factors, they have generally used routine mortality and emergency hospitalisation data to characterise important chronic conditions.[29 30 42-44] Results from these studies are not specific to the English population leading to uncertainty about the generalisability of the findings. The use of emergency admissions data may not account for individuals receiving treatment or care within the community who do not enter the hospital system prior to death, so these studies are likely biased towards individuals with more severe disease. This is especially relevant as a large proportion of deaths during heatwaves occur in the home.[45] Nor do they include information about ongoing treatment, such as prescribed medication. All of which may play a significant role in an individual's overall risk and may underestimate the effects of these risk factors. Where other factors such as wider determinants of health and heat risk have been investigated, these data have generally been linked to restricted registries[30] or used a small number of proxy measures where individuals are assigned a relevant category based solely on their geographic location.[40] While this does provide some evidence on area-level risk factors, such as the level of vegetation cover in London[46], there are a number of assumptions made which may miss some of the individual-level context of heat risk.

Previous attempts to map vulnerability to heat have relied on the use of routinely available population-level data sources combining environmental, social and demographic factors into one overarching risk index.[47-49] Results then have been mapped to reveal the geographic areas that may be particularly at risk. However, there are several limitations to these approaches. The use of routine data sources such as census data may not be reflective of the dynamic nature of the vulnerable population. The geographic resolution of the outputs limit the end-user's ability to identify individuals, as the aggregate data used to develop these maps will hide the distribution of risk and capacity to respond across the population. This is particularly important where the risk is likely to be inequitably distributed, even within a small spatial zone. Where these methods have been evaluated against health data, the accuracy of identifying where the health impacts occur was little over 50%.[47 50]

While these approaches may have been implemented from a town planning perspective in addressing heatwaves[51], they are of minimal use from a public health perspective when targeting interventions at vulnerable individuals, as recommended in HHAPs. The lack of individual-level information therefore is a limiting factor and raises the question: *what are the key individual-level risk factors associated with mortality during heatwave events; what are the relative importance of those factors; and can we use this information to develop a risk stratification tool?*

1.0.4 – About this Thesis

This thesis aims to address the broad questions presented above and provide new evidence on individual-level risk factors associated with the risk of death during periods of heat in England, to explore the utility of using primary care data to characterise individual-level heat risk and using that information to predict who is at increased risk during heatwaves. Subsequent sections of this chapter explore the existing literature on individual-level risk factors associated with increased risk of death, presentation of the specific research questions to be addressed within this thesis, and finally considerations from members of the public and primary care professionals about how the use of individual-level data might be perceived and how any risk stratification tool might be implemented.

Chapters two and three then explore the utility of primary care records to gain insight into individual-level factors associated with increased risk of death on hot days by presenting results from a time-stratified case-crossover epidemiological analysis. Chapter two focuses on clinical factors such as pre-existing conditions, prescribed medications, and clinical measurements, while chapter three then focuses on wider socio-environmental factors such as age, sex, ethnicity, frailty etc., all as captured within primary care records. Building upon chapters two and three, chapter four then explores the feasibility of using machine learning to identify individuals at risk of death on hot days based on their clinical records as captured within primary care patient management systems.

Finally, chapter five then presents a summary of the results of each chapter, discusses the limitations, strengths and novel contributions of this work to the wider evidence base, the policy relevance of the findings and potential themes for future research.

1.1 Review of the literature

Due to the overwhelming number of studies in the field of heat and health, a search was undertaken to identify any relevant, recently published reviews to assess if other researchers were asking similar questions. After completing a comprehensive search for previous reviews on the subject of individual-level risk factors associated with increased risk of ill health during heat episodes, one review was identified which addressed part of the above question, Son et al.[40]

The aim of the Son et al[40] review was to identify individual and community-level risk factors which are associated with hot and cold temperatures and assess the strength of evidence for those factors. The authors comment on the strength of evidence, simply based on the number of studies in which evidence of an association is found. There is no discussion of the effect size of each factor (or ranges of effect sizes reported) or the generalisability of the results. The individual-level factors reported are:

- Age
- Gender
- Education
- Place of death
- Occupation
- Ethnicity
- Marital status
- Chronic conditions

Given that the Son et al review was published relatively recently (2019), and followed a systematic search approach for the identification, inclusion and exclusion of relevant studies (PRISMA) and the sheer volume of heat-health studies, a pragmatic approach was taken to re-review the relevant studies included in Son et al to answer our specific questions. The rationale for this was that it was unlikely that Son et al missed many relevant published studies. Upon completion of this re-review, a further search of the literature was undertaken to identify any studies published since the Son et al review was undertaken, and where any factors may have been missed.

1.1.1 Individual-level factors

Of the 207 studies included in the Son et al review, the focus of this re-review was on studies which assessed individual-level factors and their potential modifying effect on mortality at high temperatures. Studies which focused on specific heatwave events were excluded.

The re-review focused primarily on European studies, however, where a study investigated chronic conditions, these were included regardless of location. The rationale for this was that there were very few studies included by Son et al that investigated chronic conditions.

Twenty-one of the 207 studies included in the Son et al review, were included in this re-review. The studies varied in design, populations and individual factors, country of origin and exposure metrics used. See Appendix 1 for full details of the extracted data from the included studies.

Measures of effect reported varied across studies. None of the included studies attempted to rank the individual-level factors based on their relative importance or effect size associated with the outcome of interest. However, three studies from Italy provide Relative Effect Modification indices as a means to interpret reported odds ratios. Below is a brief description of the general findings by category of individual-level factor, and an exploration of the plausibility for these to be associated with increased risk during periods of heat.

1.1.2 Age

All 21 studies assessed the effect modification of age on mortality, all finding risk increased with age.[29 30 42-44 52-66] However, one study also assessed the risk associated with heatwave duration and found that risk is higher among the younger age groups when heatwaves are longer.[62] Age ranges also differed across studies, with some focusing only on those 65+, while others assessed risk more widely across the population in smaller age bands (e.g. 45-64, 65-74, 75+).

Evidence suggests that impaired thermoregulation and haemodynamic stability are key factors in the increased vulnerability with increasing age.[9] The risk of suffering from multiple chronic conditions also increase with age.[67] Where studies explored factors beyond age and gender alone, age was not the strongest effect modifying factor. In fact, age alone as a modifying factor is generally low on the list of relative strengths of modifying factors within studies which assess more than just age.

Recent analysis of age-specific death rates suggests that those diagnosed with chronic conditions have death rates equivalent to older age groups when compared to those of the same age without the chronic condition.[68] In other words, while age in general may be a good indicator of mortality risk, the presence of other factors, for example, chronic conditions, may increase an individual's mortality risk beyond the level of risk that their age alone may infer. Therefore, any assessments of vulnerability would need to consider more than just age, as increasing age can also be considered an indicator for general population level risk, but individual-level risk will also be driven by other, additional individual factors.

Age will be an important factor to consider however in the development of any risk stratification tool as the risk of mortality during heatwaves appears the biggest for the older age groups in general and the ageing population of the UK.[69] There is some evidence to suggest that as life expectancy has increased, the age-specific trajectory of health has also shifted. For example, one study found that the rate of change in reduced functional ability was faster for an older cohort of patients compared to a younger cohort, at the same age.[70] In other words, the older cohorts declined in functional ability at a faster rate compared to the younger cohort at the same age.

There is a considerable amount of focus on healthy ageing and in taking a more holistic view of individuals' intrinsic capacity to adapt to external stressors such as heat, and to identify precursors for serious ill health so that early interventions can be implemented, such as electronic frailty index (eFI). There are likely to be synergies between this work and addressing heat-health risk for older adults.

1.1.3 Sex

Nine studies provided an assessment of effect modification by sex; six of which identified a higher risk of mortality amongst females[29 30 43 56 60 62]; one identified a higher risk amongst males[65] with two studies observing no significant difference.[42 44]

The general finding that older females are perhaps at higher risk than older males is a common one. A recent literature review which aimed to investigate the difference in risk by gender concluded that more research is needed to fully understand any physiological mechanism that may be at play and the role of social factors that may increase the risk of older females.[71]

There are perhaps some plausible reasons for these observed differences, such as the mean age of women in the older age groups (65+) are likely to be higher than the mean age of males in the same group, as women have a longer life expectancy. Living longer may also lead to more ongoing health issues as females age which can contribute to increased vulnerability. From a physiological perspective, there is some evidence of differences in thermoregulatory responses by sex from both endogenous and exogenous heat loads.[72] Post-menopausal women are at increased risk of cardiovascular disease (CVD) due to the reduction of oestrogen hormone in the body and the protective effect the hormone has on CVD risk.[73] However, this shift in CVD risk likely only increases female risk in line with male risk, so is unlikely to significantly contribute to any differences observed. There are other factors which may also contribute to this apparent difference in risk, such as social and cultural influences which may play a significant role in mediating exposures experienced.[74] Therefore, age and gender together are likely to be important factors to consider, and how they interact to increase the likelihood of other potential heat risk factors.

1.1.4 Ethnicity, wealth and place of birth

Only one study investigated the effect of ethnicity, based in the United States and found that the risk of mortality increased for the African American population compared to white Americans.[44] Wealth was also found to be a significant modifying factor in Sweden.[42] Ethnicity and income (and other socioeconomic factors such as education attainment, AC prevalence etc) are correlated in the United States, with census data going back to the 1960s showing this, so perhaps this particular finding is more of a reflection of the associated social and economic inequalities that are observed within the United States amongst different ethnicities and the structural racism that contribute to inequitable exposures to environmental stressors and the resultant poor health outcomes.[75] There is limited evidence in England to suggest heat risk differs by ethnicity,[56] however, it is well-recognised that ethnicity in England is poorly recorded and where it is, it is often incomplete.[76] Therefore, this lack of evidence does not necessarily equate to a lack of an association.

One study also investigated the association between place of birth and risk of mortality at higher temperatures between those born in Nordic countries and those not, with no significant association observed.[42] This might suggest that being born in a country with a different climate than one might be exposed to in later life, will not predispose them to change in risk. In other words, people acclimatise to their surroundings. This aligns with experimental studies in which results suggest there was no difference in thermoregulatory response to extreme heat conditions by men who were of African descent in Canada and white males of European descent[77,78] suggesting the physiological response to heat is not driven by genetic or physiological differences in ethnic groups, but by other environmental adaptations.

Ethnicity may represent a proxy for other societal factors that are associated with increased inequalities and deprivation, such as structural racism experienced by ethnic minority groups leading to increased health inequalities.[79] Similarly, a recent study demonstrated that deprivation and access to greenspace may also impact the inequitable distribution of heat risk in London.[46] This all suggests that social and economic factors are potentially vital to consider when investigating heat risk, to ensure that any interventions that are deployed are equitably targeted and distributed across the population.

1.1.5 Marital status

Three studies included marital status as an individual-level factor, all based in Italy.[29 30 43] All three found a significant association of mortality among those not married, widowed, or divorced (i.e. not married at the time of death). Marital status is perhaps a proxy for living alone and social isolation. However, this factor is likely to be very context-specific as different cultures may have different norms. In addition, research has suggested that living alone may increase the risk of premature mortality in general[80], not associated with heatwaves. Loneliness and social isolation have been found to be associated with an increase in risk of coronary heart disease and stroke; both potential risk factors associated with heatwave mortality.[81] Therefore, loneliness and social isolation are likely to be important risk factors to consider in the current study.

1.1.6 Place of death

Five studies investigated the place of death and risk of mortality on hot days, [29 30 44 60 65] with a consistent finding that risk increased for deaths occurring out of hospital (at home), in hospital and in nursing homes, with nursing homes observing the highest risk of death on hot days.

One Italian study found that there was a significant risk of mortality for those who were already admitted to hospital when the high temperatures occurred, but not for those entering the hospital on the day on which the high temperatures occurred.[29] The same study further investigated which type of hospital ward might have a higher risk of death on hot days for those already admitted into hospital and found that risk was highest for those in general medicine wards compared to other wards where patients receive higher levels of care.[29]

Increased risk of death on hot days for those in care homes is a consistent finding. There may be several reasons for this increased risk such as a high concentration of older individuals who are suffering from a range of chronic conditions and taking prescribed medications which may also add to underlying vulnerability during heatwaves. According to the British Geriatric Society, average life expectancy in UK care homes is 24 months for those who do not require nursing, and 12 months for those who do. This is reduced further for those who enter the home with one or more deteriorating conditions.[82] Therefore this population is likely to be towards the end of their lives, but with differing levels of overall risk. In these settings, patients will be receiving a level of care depending on the individual's needs. Therefore, it's unclear if the increased risk is due to a potential deficit of care as a result of a lack of understanding of the hazard by care staff or if the concentration of potentially vulnerable individuals are the main drivers of this observed impact.

The risk of mortality was also consistently found to increase for those at home on hot days. This is likely due to a range of factors such as the individual susceptibility, their dwelling's propensity to overheat, the occupants' behaviours, cultural and social norms and/or an inability to either sense they are overheating or lack the ability to adapt their own environments or behaviours. Unlike the above two settings, (hospital and care/nursing homes) there may be little to no care provided. These heatwave deaths are likely to represent the population who are not accessing health care during a heatwave and dying before they enter the system. This is consistent with observations in England during heatwaves in that there are very low numbers observed for heat-associated emergency attendances, GP consultations and ambulance calls compared to all-cause mortality values.[19 37] These results suggest that where people are when heatwaves occur may be a significant risk factor.

The individual-level risk factors identified above represent a group of socio-environmental factors which may increase an individual's overall risk status. Inclusion of such factors in any heat risk stratification tool for use in clinical settings would be vital should the data allow, given that heat risk is driven by many domains, including social and environmental determinants of health.

1.1.7 Pre-existing clinical/medical risk factors

One of the key domains of heat risk is individual-level susceptibility, and an individual's health status can play a large part in determining their overall risk during periods of heat. Therefore, information about chronic conditions and other medical risk factors are likely to play a significant role in an individual's heat risk.

Six of the studies included in this re-review had an assessment of the risk of mortality on hot days for a range of chronic and acute conditions. [29 30 42-44 66] Precise conditions investigated vary across the studies but are generally categorised as circulatory diseases, respiratory diseases, and other diseases as can be seen in Table 1.2. Each group is then further discussed in the following sub-sections.

Table 1. 2 Specific conditions associated with increased risk of heat-related mortality.

Chronic diseases investigated	**Study 1**	**Study 2**	**Study 3**	**Study 4**	**Study 5**	**Study 6**
Circulatory system diseases						
Heart failure (ICD-9: 428)	*X*	***X****			*X*	
Other ischemic heart diseases (ICD-9: 411, 413–414)	*X*	*X*			*X*	
Conduction disorders (ICD-9: 426)	***X****	*X*			*X*	
Cardiac dysrhythmias (ICD-9: 427)	*X*	*X*			*X*	
Cerebrovascular diseases (ICD-9: 430–438)	***X****	***X****		*X*	*X*	
Previous acute myocardial infarction (ICD-9: 410, 412)	*X*			***X****		
Diseases of arteries, arterioles, and capillaries (ICD-9: 440–448)	*X*	*X*				
Diseases of pulmonary circulation (ICD-9: 415–417)	*X*	*X*				
Hypertensive disease (ICD-9: 401–405)	*X*	*X*				

Chronic diseases investigated	**Study 1**	**Study 2**	**Study 3**	**Study 4**	**Study 5**	**Study 6**
Diseases of valves (ICD-9: 394.0–397.1, 424, 746.3–746.6, 093.2)	*X*					
Cardiovascular diseases (ICD-9: 390-459; ICD-10: I)				*X*	*X*	
All circulatory system diseases (ICD codes not given)						***X****
Respiratory system diseases						
Pneumonia (ICD-9: 480–486)	*X*	*X*				
Chronic pulmonary diseases (ICD-9: 490–505)	*X*	***X****	*X*	***X****	***X****	*X*
All respiratory diseases (ICD-9: 460-519; ICD-10: J)				***X****		
Malignant neoplasms (ICD-9: 140–208)	*X*	*X*			*X*	
Diabetes mellitus (ICD-9: 250)	*X*	*X*	***X****	*X*	*X*	*X*
Psychosis (ICD-9: 290–299)	***X****	***X****		***X****	*X*	
Depression (ICD-9: 300.4, 301.1, 309.0, 309.1, 311)	***X****					
Substance abuse (ICD-9: 303-305; ICD-10: F10-19)				*X*		
Other disorders of the central nervous system (ICD-9: 330–341, 345–349)	*X*	*X*			*X*	

Chronic diseases investigated	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Renal failure (ICD-9: 584–588)	*X*	*X*			*X*	
Acute and chronic liver diseases (ICD-9: 570–572)	*X*				*X*	
Anaemias (ICD-9: 280–285)	*X*	*X*				
Diseases of the osteo-muscular system (ICD-9: 710–739)	*X*	*X*				
Fracture of femur (ICD-9: 820–821) or hip (ICD-9: 820–821)	*X*	*X*				
AIDS (ICD-9: 042)	*X*					
Disorders of thyroid gland (ICD-9: 240–246)	*X*					
Disorders of fluid, electrolyte, and acid-base balance (ICD-9: 276)	*X*					
Obesity and other hyperalimentation (ICD-9: 278)	*X*					
Coagulation defects (ICD-9: 286–287)	*X*					
Paralysis (ICD-9: 342–344)	*X*					

The six chronic disease studies are referred to as follows in the table: Study 1 = Stafoggia et al 2006[30]; Study 2 = Stafoggia et al 2008[29]; Study 3 = Medina-Ramón et al 2006[44]; Study 4 = Rocklöv et al 2014[42]; Study 5 = Schifano et al 2009[43]; Study 6 = Sun et al 2016[66]

*Each condition investigated per study indicated with X, ICD9/10 codes used to define the various conditions are provided, and an * indicating conditions found to significantly increase risk of mortality on hot days*

Study designs differed across the six studies, with two Italian studies[29 30] and one Swedish study,[42] using case-crossover, one cohort study,[43] one nested case-control study[66] and one case-only study.[44] Each study design used has advantages and disadvantages associated with them for investigating the modifying effect of chronic disease on heat-associated mortality and for addressing potential confounding factors.

Cohort studies are used for rare exposures meaning studies can focus on subjects who have been exposed to a certain factor, here high temperatures. However, for rare diseases or outcomes, a large study population is required for sufficient statistical power. Furthermore, cohort studies are better suited when following patients over time following an exposure and the evolution of disease, rather than the relatively sudden onset of death following heat exposure.[83]

The nested case-control study design used by Sun et al allows for the collection of precise individual-level data from the cohort of patients included in the study which may otherwise have been hard to obtain; the study design is suitable for rare diseases; by using available data, fewer resources were needed to conduct the study. However, bias in selecting controls is a potential source of error in this study design, with the most prominent limitation of such an approach being that only the relative risks can be estimated, as opposed to the absolute risks.[83]

Case-only studies allow the investigation of how individual-level factors which do not vary over time (or at least very little) modify the effect of time-varying exposures on outcomes. However, as the analysis is focused on cases only (i.e. no controls or control periods) the associations derived are relative only to the cases included in the study. For example, a negative association which may be derived does not necessarily mean an overall negative effect.[84]

Case-crossover studies have the advantage that by design, common individual-level confounding factors are controlled for as the cases act as their own controls at differing time intervals when selected appropriately. In addition, the case-crossover study design is especially suited to investigating acute onset of disease as a result of a sudden exposure, in this case, death and high temperatures. Another advantage of this study design is in the ability to assign precise exposure metrics to each individual, should that data be available.[85]

Definitions used for chronic conditions also differed, with Rocklov et al[42], both Stafoggia et al studies[29 30] and Schifano et al[43] all using hospital episode data to define chronic conditions as either the primary or secondary contributing factor for admission between 28 days and two years before death. The second Stafoggia et al study[29] uses the same definition for chronic conditions but also investigated acute admissions (using the same ICD-10 codes) which are defined as admissions for either the primary or secondary reasons within two days of death. Medina-Ramón et al[44] used information related to the presence of chronic diseases contained within death certificates, while chronic disease was defined by self-reporting by each case by Sun et al.[66]

Exposure measures used differ across studies, with mean apparent temperature, maximum apparent temperature, maximum absolute temperature, and minimum temperature all used across the studies. In terms of confounding, all studies attempt to control for the potential effects of a range of factors including seasonal trends (day of week, month, public holidays, changes in population characteristics in summer), air quality (PM_{10}, O_3 and NO_x), humidity and influenza activity.

It is well understood that heatwaves and poor air quality episodes are correlated, with the weather conditions which lead to heatwaves also contributing to conditions which lead to poor air quality. In addition, impacts on health associated with heat are also strongly associated with poor air quality, therefore controlling for air pollutants is appropriate. However, due to the dynamic nature of air pollutant concentrations, accurate measures of exposure at the individual level is extremely difficult, meaning that proxy aggregate measures are used, which may result in mischaracterisation of air pollutant exposures. There is limited epidemiological evidence that humidity plays a major role in heat-related mortality,[86] however, it does likely affect temperature perception and thermal comfort. Other potential confounding factors, such as influenza activity, is minimal in the summer months in the UK, with UKHSA syndromic surveillance systems (amongst others) demonstrating the seasonal nature of this activity.[87]

Two of the Italian studies that used the same methods focused on slightly different sub-groups of the same populations – i.e. one focused on the general population while the other focused solely on those dying in hospitals.[29 30] This represents two distinct populations with the potential to have different risk profiles. For example, those dying in hospital were admitted for a range of different factors and are likely to represent the severe end of the disease state spectrum, while those dying at home are likely to represent individuals who were not at the severe end prior to the heat arriving.

One of the six studies assessed individual-level factors in two ways; first, as the risk of mortality for an individual per 1C increase in temperature and second, assessed the risk of individual-level factors based on the duration of a heatwave event.[42] This study suggests that the population profile of those at risk as temperature increases may differ for different durations of heatwave. Reported heatwave excess mortality in England in 2020 supports this suggestion.[19] In August 2020 there was a prolonged heatwave which lasted for 15 days and coincided with a number of tropical nights and resulted in significant all-cause heat-associated mortality observed in the 0-64 years group of 247 (95% CI 113-446) additional deaths. Significant heat-associated mortality among this group had not been observed at a national level in previous summers.[19]

Due to the heterogeneous nature of the studies included here, it was not possible to compare or perform a meta-analysis. However as previously stated, there were some common themes identified in terms of the relative strengths of association of individual-level factors. In summary, across the six studies, age and sex did not demonstrate the strongest association with heat-related mortality. This was observed for either circulatory system or respiratory system diseases, with other chronic diseases identified as significant modifying factors demonstrating more of an effect than both age and sex. While this does not provide a definitive view into what conditions are the most important to consider when attempting to identify vulnerable individuals and the effect size they may have on the risk of mortality in England, it does allow us to view which individual-level factors are consistently identified as significant modifiers and inform what factors may be useful to consider in future analysis. These chronic disease groups are described in the following sub-sections.

1.1.7.1 – Respiratory system diseases

Chronic pulmonary disease represented either the highest or second highest estimated effect size of individual-level risk factors associated with high temperatures in four of the studies, however, definitions differed.[29 42 43 66] In addition, one study also identified hospital admission due to any respiratory disease in the previous two years as a significant risk factor for those over 65 years of age.[42]

From a physiological perspective, the likely mechanism for ill health in relation to respiratory system diseases is the reduced ability to get sufficient oxygen to the cells when the body is overheating. However, any association of respiratory system diseases would need to be assessed considering air pollutant concentrations, as air quality is known to be potential confounding factor.[53]

1.1.7.2 – Circulatory system diseases

General circulatory system issues and specific conditions were identified as significant risk modifiers in four of the six studies.[29 30 42] This included conduction disorders, acute myocardial infarction, heart failure, cerebrovascular disease and cardiovascular disease. These specific circulatory system diseases were identified in more than one study, although perhaps defined in slightly different ways.

Physiologically, there are numerous mechanisms by which an individual with an underlying health issue associated with their circulatory system could be at increased risk. These include:

1. high temperatures leading to increased blood viscosity and potentially increased strain on the heart while trying to maintain blood pressure[88] (myocardial infarction, cardiomyopathy, cardiac arrest and heart failure)
2. flow is redirected to areas of the body and surface where heat can escape the body which in turn increases the potential for burst vessels and bleeding in and around the brain[88] (haemorrhage)
3. due to thermoregulatory response, disruption in the supply of blood and oxygen to the brain[88] (stroke and other cerebrovascular diseases)

4. additional strain on the heart is too much for a heart that is already beating irregularly[88] (arrhythmia)

1.1.7.3 – Diabetes

All six studies investigated the association between heatwave deaths and diabetes. Two of those studies identified diabetes as a significant risk factor, however, neither were based in Europe.[44 66] The use of hospital admission data alone to define chronic disease may not adequately identify all chronic conditions which lead to ill health during a heatwave such as diabetes. For example, the rate of hospitalisation in England for those with type 2 diabetes for causes other than diabetes is much higher than for those admitted directly for diabetes.[89] This suggests that the use of hospitalisation data alone may underestimate the modifying effect of certain chronic diseases which may not lead directly to hospitalisation, but rather contribute to the risk of hospitalisation for other primary reasons. In addition, there is evidence in England that those with type-2 diabetes are at increased risk of requiring medical consultation during days of temperature extremes, especially during hot weather.[90]

The physiological pathways for increased risk for those with diabetes is not entirely clear, however, some studies suggest that blood skin flow may be altered for diabetes patients, therefore potentially reducing the individual's thermoregulation efficiency.[91] In addition, diabetes is a recognised risk factor for a range of other conditions which are also known to increase the risk of mortality during heatwaves, e.g. cardiovascular diseases. Therefore, diabetes may both directly and indirectly increase an individual's risk.

1.1.7.4 – Mental health conditions

Four of the included studies also explored the relationship between mental health disorders and the risk of mortality during heatwaves[29 30 42 43], with three of those studies identifying at least one mental health disorder as being a significant effect modifier.[29 30 42] Mental health conditions considered significant include psychosis and depression. There is strong evidence in England and around the world that the risk of suicide as the cause of death increases with increasing temperature.[10] Previous studies have also identified that those suffering from depression and psychosis are at increased risk of suicide.[92] None of the included studies which identified these mental health conditions as risk modifying factors considered cause of death, therefore there is no way of exploring that thread further, however, it is a plausible mechanism.

More generally, the physiological pathways for those with mental health conditions and heat-associated mortality is unclear, however, there may be several elements to it. These may include medication prescribed to control symptoms interfering with the thermoregulatory response or suppressing thirst; the inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive the risk; or a combination of all three. In addition, the way in which mental health conditions were defined within the included studies are very broad, and it is plausible that different conditions may have different ways in which that risk manifests, and the role of medication used to control symptoms may also differ across conditions.

1.1.8 Other individual-level factors not identified by Son et al

While Son et al identified a range of chronic conditions which have been found to be associated with increased risk during a heatwave, there are some potentially important individual risk factors that could be of interest from a physiological perspective. In addition, new evidence may have been published since the publication of Son et al. For example, a recent review not included in the Son et al review also found evidence for strong associations of hospital admissions for those diagnosed with a range of mental health conditions, including bipolar disorder, schizophrenia and other cognitive disorders such as Alzheimer's and dementia.[10] Therefore, a further search of the literature was carried out, with either additional evidence or the plausible physiological mechanisms leading to ill health outlined below.

1.1.8.1 - Obesity

Obesity has been identified as a significant risk factor for older adults (65+) who died in Paris during the 2003 European heatwave[93] yet was not identified as a significant risk factor in the included studies. This is perhaps linked to the data sources used in the studies and the fact that obesity is generally diagnosed and treated within primary care[94] and is unlikely to be well recorded within hospitalisation data. For example, NHS England data suggests that in 2018/19 there were 11,117 hospital admissions with a primary diagnosis of obesity. This compares to 876,000 hospital admissions over the same period where obesity was only a contributing factor.[95] Therefore, this suggests that the use of hospital records to define some chronic conditions, as used in the included studies, is perhaps not the best approach, as there is potential for underestimation due to the way in which they are diagnosed and treated.

There is potential for both direct and indirect risks for those who are obese. One potential mechanism leading to increased risk may be that due to higher body mass, the individuals' organs are likely to need to work harder (higher strain on the heart for example) than for those with lower body mass to lose heat from the body. In addition, high BMI or obesity is a known risk factor for a range of other heat-sensitive conditions, and thus may indirectly increase an individual's risk.[96]

1.1.8.2 – Parkinson's Diseases

Analysis of 2020 heatwave mortality in England by place and cause of death identified a small peak in deaths with the underlying cause of death recorded as Parkinson's disease during the August heatwave[39]. Evidence also exists from Spain suggesting that increased risk of mortality is observed for those diagnosed with Parkinson's disease.[97] However, this was not investigated in the studies included by Son et al. The potential physiological pathways are unclear, but potentially linked to dehydration as a side effect of medication taken to control symptoms. Some anti-Parkinson's medications are known to have this side effect.[98]

1.1.8.3 – Chronic kidney disease

Chronic kidney and other urinary diseases have been linked with occupational heat exposures and increasing extremes in temperatures[99 100] and were not included in the Son et al studies. There is strong evidence suggesting an association between prolonged and continuous heat exposure and chronic kidney disease in occupational settings globally.[100 101] In addition, recent analysis in the UK suggests that heat-related acute kidney injury is becoming a public health challenge.[102] Dehydration is one of the mechanisms of serious kidney injury and reduced kidney function due to reduced water content in blood.[103] Dehydration is one of the mechanisms which can lead to ill health during periods of extreme heat.

1.1.8.4 – Alzheimer's and dementia

As briefly mentioned above, individuals with a diagnosis of Alzheimer's and dementia have been shown to be at increased risk during heatwaves.[104 105] The mechanism by which an individual with a diagnosis of Alzheimer's or dementia's risk is elevated is unclear, however, it may be partly due to medication prescribed to control symptoms interfering with either thermoregulation or hydration; the inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive risk on hot days; or a combination of all the above.

1.1.8.5 – Endocrine system

No evidence of the effect of high temperatures on individuals with thyroid diseases were found either in the Son et al review or in other relevant literature. However, the thyroid system is intrinsically linked with thermoregulation as it regulates the body's metabolism; body temperature rises because the basal metabolic rate is raised as there is increased oxygen consumption and the patient's hyperactive adrenal function is globally enhanced.[106] The body may therefore have to work harder in someone with thyroid or adrenal diseases to lose excess heat, potentially increasing strain on other organs sensitive to high temperatures as described above.

1.1.8.6 – Frailty

Frailty is defined "as a clinically recognisable state of increased vulnerability resulting from an ageing-associated decline in reserve and function across multiple physiologic systems such that the ability to cope with every day or acute stressors is compromised".[107] In England, frailty is clinically assessed annually for all individuals over the age of 65 years as part of the contract between the National Health Service (NHS) and general practice via the electronic frailty index (eFI) approach. eFI is calculated as a ratio of the number of specified markers of frailty, termed deficits, an individual has within their primary care records out of a total of 36.[32]

There has been no study to date that has explored frailty and heat risk, however, a number of the specified clinical factors which feed into the eFI assessment are aligned with those identified within the literature as potentially heat sensitive, such as a number of circulatory and respiratory system conditions, diabetes etc.[32] In addition, eFI also considers social vulnerability as one of the potential “deficits”.[32] Therefore, it is plausible that as an individual’s frailty increases, their heat risk may also increase.

1.1.8.7 The homeless and those sleeping rough

Evidence is emerging globally that the homeless population and those sleeping rough are at increased risk of adverse health effects during periods of heat.[108 109] Reasons for this are likely to be complex, but potentially include a combination of the fact that this population simply do not have access to cooler spaces when it’s hot, and may not have the autonomy or capacity to adapt their environment or behaviours. Furthermore, evidence suggests that there is a high prevalence of a range of mental health conditions such as depressive and anxiety disorders, schizophrenia spectrum and psychotic disorders, substance use disorders, suicidal behaviour, bipolar and mood disorders, neurocognitive disorders and other mental disorders amongst this population group.[110] As highlighted previously, this may increase the risk of this population group further. Therefore, attempting to understand an individual’s living arrangement may be an important factor to investigate in terms of individual-level heat risk.

1.1.8.8 Medication use and heat risk

Guidance associated with the AWHP[24] and other international Heat-Health Action Plans[23], lists a range of medications which potentially alter the body’s ability to thermoregulate or alter the mineral-water balance and ultimately lead to dehydration. There is evidence to suggest that those who suffer from chronic conditions and receive medications to control their illness may be at higher risk of mortality and hospitalisation due to the potential modifying effect of the prescribed medications.[111] This potential risk has also been observed at more moderate temperatures during the summer months.[111]

A recent review investigated the evidence base for the physiological changes, focusing on the thermolytic processes associated with medication use during heat stress.[112] The review specifically focused on diabetes and antidiabetic drugs, cardiovascular diseases and cardiovascular disease drugs, neuropsychiatric conditions and drugs used to treat them, and cancer and cancer treatment drugs. The study concluded that while there is evidence to suggest that the conditions and drugs used to treat them may impair thermoregulation, many evidence gaps remain on the interaction of chronic conditions and medications and their effect on thermoregulation in older adults, and ultimately risk during heatwaves. The review also highlighted that more epidemiological evidence is needed to allow the formulation of clinically relevant guidance for practitioners to refer to when considering patient medicine management during periods of extreme temperature.

There is also evidence that suggests that while some conditions may affect thermoregulation, the medications used to treat them may also lead to dehydration, further increasing that individual's risk.[113] For example, heat-associated mortality amongst those with Parkinson's disease listed as the underlying cause of death has been observed in England[39] and Spain.[97] While there is evidence that the physiological changes in thermoregulation of individuals diagnosed with Parkinson's (such as sweat volume for example), it is perhaps the effect of anti-Parkinson's disease medication (such as levodopa) has on risk of dehydration that drives risk during a period of heat, as thirst suppression is a known side effect of these drugs.[114]

One of the major barriers to investigating the effect of medication use on heat risk is the complexity of the pharmacodynamic mechanisms by which a drug molecule interacts with biological systems to elicit the desired effect. Drugs used to treat one condition may fall into different groups of drugs and the side effects may differ, with one group increasing risk of either dehydration or altering thermoregulation. For example, the drugs commonly used to treat Alzheimer's and dementia are either an inhibitor or an antagonist type drug.[115] These different modes of action may manifest in different ways within the patient leading to different side effects and risk. In the example of Alzheimer's and dementia, vomiting, diarrhoea and dehydration are common side effects experienced by those prescribed acetylcholinesterase inhibitors, which potentially could lead to increased risk during periods of high temperature due to dehydration.[115] Whereas for the drug memantine (antagonist), these side effects are less likely.[115] The example demonstrated here is replicated across conditions which are known to be sensitive to high temperatures and demonstrates the complexity in this specific topic area, before even beginning to examine the potential interactions between condition, medication, heat exposure and health outcome.

As has been indicated above, the use of hospitalisation data as a proxy to define chronic disease, or indeed as a means to analyse potential individual clinical and medical factors associated with increased heat risk may underestimate some important relationships. The use of primary care records would allow the investigation of these more nuanced risk factors.

1.2 Research Question

As has been demonstrated, there is a lack of intelligence on a range of potential individual-level risk factors and their potential risk modification effect in relation to risk of mortality during heatwaves. And where there is evidence, studies have generally used hospitalisation statistics, which may only capture part of the relevant sub-population of interest, or the severe end of the population. These gaps in the knowledge present significant barriers to the development of effective and evidence-based approaches and widespread deployment of one key recommendation outlined in HHAPs such as the AWHP, namely identifying individuals at high risk and deploying targeted interventions.

The aim of this PhD was to explore the feasibility of developing a risk stratification tool that is capable of identifying individuals at risk of death during heatwaves. To do this, two key research questions are addressed. First, is the use of primary care records a viable source of data on individual-level heat-risk factors? And if so, what type of information contained within primary care records can be used to improve our understanding of individual-level risk. And second, can we use information on individual-level risk factors as recorded within primary care data to predict which individuals are at risk of death during periods of heat.

To achieve this, first, an epidemiological analysis of individual-level risk factors associated with heat risk using electronic health records (EHR) containing primary care data was undertaken, with the initial focus on clinical risk factors such as pre-existing conditions, prescribed medications and clinical measurements (objective 1). This was followed by exploring a range of wider determinants of health as recorded within primary care data where possible (objective 2). Following the initial epidemiological analysis, Random Forest, a machine learning approach, was used to determine the feasibility of using primary care records to identify those considered at high risk of death during heatwaves (objective 3).

To help ensure that the research is as useful from a policy and implementation perspective as possible, the research was informed by engagement and consultation with members of the general public (intended beneficiaries of any heat-risk stratification tool) and primary care professionals (intended users of any heat-risk stratification tool). Insights gained from both formal and informal engagement highlighted key factors considered within the project and areas of future research to ensure that the implementation of an evidenced-based heat-risk stratification tool is as impactful as possible. Figure 1.1 below summarises the three objectives and how they are interlinked, and the intended outputs of the project.

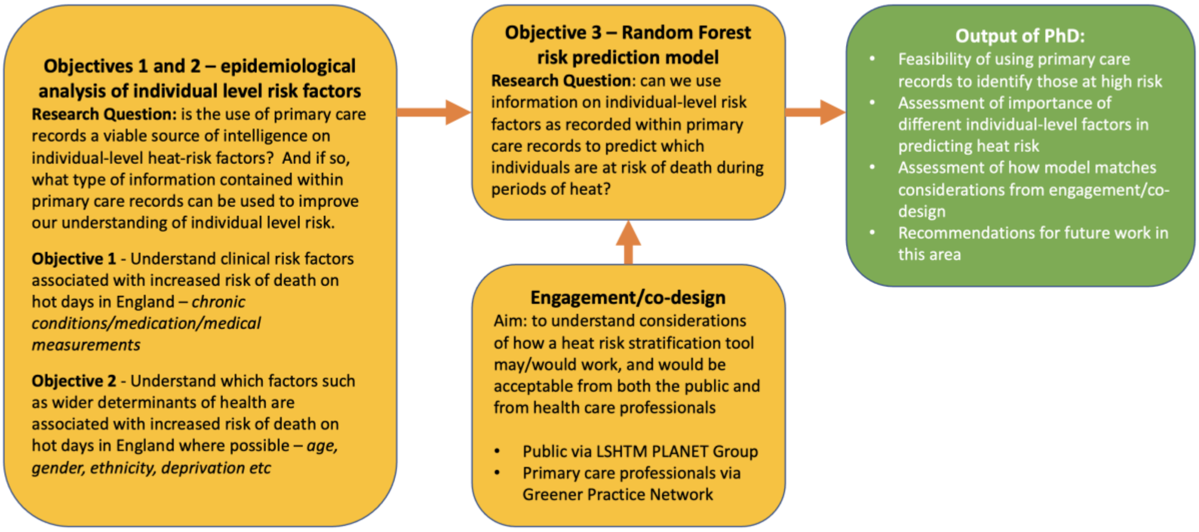


Fig 1. 1 Outline of Thesis

1.3 Co-design: considerations for any risk stratification tool from key users/recipients of output

The views of service providers and the public are key to ensuring that any risk stratification tool is acceptable, useable, and used. Therefore, the best approach to gaining valuable user insight is to engage directly to ensure that key views, concerns and suggestions are considered from project initiation. To address the need for the research to be relevant to both the intended user (i.e. clinicians) and recipients (i.e. patients), a series of engagement sessions were held. This was done both formally via The Health Protection Research Unit in Environmental Change and Health public engagement/involvement group called PLANET, and informally via the Greener Practice, an informal network of primary care professionals interested in making primary care more sustainable and increase awareness of the risk posed by a changing climate.

The PLANET group were asked three broad questions which were open to exploring themes that organically came through the discussion. The first question was related to how the public might react to being contacted by their primary care practice suggesting they were at high risk during a heat event. The second asked if primary care was the right place to receive this type of information from, or if it should come from other sources. And finally, what type of information they might want to receive if they were contacted to be told they were at risk. With the primary care professionals' group, the questions were less formulaic and started with a discussion about the role of primary care in addressing heat risk more generally and then moved on to explore some of the practical aspects of how a risk stratification tool might be incorporated into their systems and some of the key concerns they might have about risk stratification tools more generally.

The outcome of these sessions helped to stimulate thinking about the resultant tool being developed and some further thinking about where such a tool might sit, and what would primary care professionals actually be expected to do with the output of a risk stratification tool on top of all other responsibilities they have. The subsequent sections of this chapter outline the key themes of discussion from the two groups with some context on how the conversations evolved, including some of the concerns raised.

1.3.1 The general public

The NIHR funded Health Protection Research Unit in Environmental Change and Health has a public engagement/involvement group called PLANET, which was established in Autumn of 2020. PLANET stands for Public Led and Knowledge Engagement Team. The group has 30 members and approximately 20 attend the regular meetings every 3-4 months to discuss research projects. A session exploring heat risk and the use of primary care records to identify those at the highest risk was carried out in May 2022. The aim of this engagement was to gain insight from a sample of the general public on core themes that should be considered when developing such a tool following an open deliberative discussion approach.

Five key themes were identified within the session. First, **there was broad agreement that primary care was considered to be the right source of personal health information at the individual level**. There was broad agreement across participants on this. Primary care providers are a trusted source of information on individual health issues, and therefore if someone is contacted by their GP/practice nurse, they are more likely to take note and act. Participants also agreed that other organisations (such as UKHSA) do not have the level of understanding to be trusted on individual-level health issues. Primary care contact would also give reassurance that personal and clinical data is not being provided to third parties.

Second, the group suggested that **any heat-health risk stratification tool needs to have a high degree of accuracy with GPs having the final say on individual risk**. Participants raised the concern that if the tool is not accurate then trust in the output would be low and undermine the purpose of the tool. Participants also raised concerns that if the GP is not given a final decision on overall risk, there is a danger that people could be contacted by mistake leading to further mistrust.

Third, it was suggested that **how individuals are contacted will need to be sympathetic to the individual and their location**. When asked about how the tool might be used by primary care professionals, there were varying opinions on how people may want to be contacted by primary care with some suggesting that direct contact would be preferable, with others suggesting it may cause undue anxiety. It was also raised that where an individual is located may be relevant in terms of the information they are given, for example, information relevant to those in London may not be of relevance to those in the North West. In addition, the timing of contact was also a topic which had varying opinions. However, this last point would be determined by the purpose of the tool, i.e. if the purpose is to address the health burden during a heatwave, contact would only be where a heatwave is forecast. But a more general awareness-raising campaign amongst patients might also be beneficial at the start of the summer for example.

Fourth, the group largely agreed that **a range of information sources would need to be available for those identified as high-risk**. Again, there were differing opinions on the level of information that would be wanted, ranging from very specific details about individual risk factors, through to more general information, with the individual responsible for investigating the issue and making an informed decision. In addition, information about other hazards which could compound risk should also be made available where relevant.

And finally, it was flagged that **the language used in any communication with individuals identified as high-risk would need to be considered carefully.** There was agreement that any information provided should be clear and precise, with no jargon and easily understood. In addition, terms such as vulnerable should be thought through so as to not offend those identified. Any individual-level communications should be supplemented by a wider communications strategy so that if someone is contacted, there is a wealth of information for them to look to for more detail if they want it.

1.3.2 Primary care professionals

Primary care professionals were recruited via the Greener Practice, a network of primary care professionals encouraging action on sustainability within primary care. Due to capacity issues with those recruited, a group discussion was not possible. Therefore, a series of one-to-one discussions were held to gain insight into what type of considerations might be useful to consider before the development of a risk stratification tool for identifying individuals at risk during heatwaves.

Five key themes emerged from the discussions which align very much with the views and opinions of the PLANET Group. First, it was unanimously suggested **the tool must have a high degree of accuracy in identifying those actually at high risk,** even at the proof-of-concept stage**.** As discussed by the Planet Group, lack of accuracy could undermine trust in the primary practice when contacting those not at risk, and potentially lead to inaction. However, a point raised by one of the participants suggested that even if a heat-risk stratification tool had a high degree of accuracy, it is completely possible that an individual at risk, and accurately identified by the tool, may not accept that they are at risk, meaning the tool may still prove to be ineffective. Therefore, it's not just about the accuracy of a tool, but about how the risk is communicated in addition to the actions undertaken as a result of the identification of the risk that may influence the desired outcomes.

Second, it was strongly suggested that **clinicians would need to have the final say on an individual's overall risk.** There was a consensus that any risk stratification tool will never be

100% accurate, and that there will need to be a degree of sense checking by a healthcare professional. In addition, clinicians within primary care are likely to know their patients, particularly those who they think are likely to be at increased risk, therefore that additional knowledge needs to be allowed to factor into the equation. A flagging system may be acceptable from this point of view, in that a risk stratification tool simply flags those identified, and the clinician will then consider that when making the final decision on what action to take, if any, once all potential factors are considered.

Third, **a heat-risk stratification tool must be simple to use, with minimal effort.** Somewhat counter to the above, there was a clear steer that the use of a tool needs to be seamless and not add to the workload of primary care. With an ageing population, and staffing issues currently affecting the NHS and primary care in particular, which manifests in expected increases in workload, any additional ask to primary care needs to fit within current systems and processes. Therefore, it's important to think about the expected actions as a result of someone being identified as at risk, and if there are any current processes to which heat risk could be attached.

Fourth, a question about **primary care being the right place utilising such a tool** was raised. Reasons for this question relate to the capacity issue flagged above, but also relate to the data required to characterise heat risk. It was flagged that there are other factors which affect an individual's risk that clinicians would have no knowledge of, such as housing for example. It was suggested that it might be more appropriate for a tool like this to be used within Integrated Care Boards (ICBs) which will generally have more data than primary care practices, and perhaps a useful approach would be for ICBs to carry out the identification and liaise with primary care networks to implement targeted action on those identified.

Finally, **the output of a tool needs to be accompanied by clinical guidance to help clinicians know the most appropriate action to take**. The example issue raised several times included the lack of clinical guidance on how to manage patients on certain medications which evidence suggests may increase an individual's risk during heatwaves. Without having appropriate guidance on what interventions to implement, identifying an individual would only lead to general advice already provided by public health agencies, and therefore would increase the ask on primary care without providing the required tools to implement effective interventions.

1.3.3 – Key considerations for this research

The discussions with both the Planet Group and primary care professionals were striking in that the two groups identified almost the same themes and issues that need to be considered. Of prime importance was the accuracy of any tool. Both groups also highlighted that any tool would need to be accompanied by relevant information, either as guidance for professionals or information for the public about why they were identified as being at risk and what they can do to address their individual level risk. This indicates that the "so what" is just as important as the "how" to both groups. Therefore, before attempting to implement any tool, thought needs to go into what actions would be required when an individual at risk is identified. While this last point is beyond the scope of this PhD, it is an important point to consider. In general, however, there was broad support for a system by which individual risk is investigated and those with the highest risk identified so that targeted interventions could be deployed, even if questions were raised about primary care being the most appropriate place for a risk stratification tool to be implemented. Most of the themes and issues raised by the groups were perhaps beyond the scope of this PhD. Nevertheless, one of the target outcomes of this project is to provide some suggestions on the next steps for developing an evidence-based approach to identifying those most at risk during heatwaves, and insight gained here will be invaluable when combined with learning from the epidemiological analysis and machine learning objectives.

References

1. Armstrong B. Models for the Relationship Between Ambient Temperature and Daily Mortality: Epidemiology, 2006:624-31.
2. Armstrong BG, Chalabi Z, Fenn B, et al. Association of mortality with high temperatures in a temperate climate: England and Wales. Journal of Epidemiology and Community Health 2011;65(4):340-45. doi: 10.1136/jech.2009.093161
3. Gasparrini A, Guo Y, Sera F, et al. Projections of temperature-related excess mortality under climate change scenarios. The Lancet Planetary Health 2017 doi: 10.1016/S2542-5196(17)30156-0
4. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? Occupational and Environmental Medicine 2007;64(2):93.

5. Hajat S, Vardoulakis S, Heaviside C, et al. Climate change effects on human health: projections of temperature-related mortality for the UK during the 2020s, 2050s and 2080s. Journal of Epidemiology and Community Health 2014 doi: 10.1136/jech-2013-202449
6. Kovats RS, Hajat S, Wilkinson P. Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. Occupational and Environmental Medicine 2004;61(11):893.
7. Climate Change 2014: Synthesis Report. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. AR5. Geneva, Switzerland: Intergovernmental Panel on Climate Change, 2014:732.
8. Allen M, Mustafa B, Shukla P, et al. GLOBAL WARMING OF 1.5 °C - an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. 2018
9. Ebi KL, Capon A, Berry P, et al. Hot weather and heat extremes: health risks. The Lancet 2021;398(10301):698-708. doi: https://doi.org/10.1016/S0140-6736(21)01208-3
10. Thompson R, Hornigold R, Page L, et al. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. Public Health 2018;161:171-91. doi: https://doi.org/10.1016/j.puhe.2018.06.008
11. Heatwave Plan for England. In: England PH, ed. https://www.gov.uk/government/publications/heatwave-plan-for-england: Public Health England, 2015.

12. Heatwave plan for England Making the case: the impact of heat on health – now and in the future. In: England PH, ed.: Public Health England, 2016.
13. Arnell NW, Kay AL, Freeman A, et al. Changing climate risk in the UK: A multi-sectoral analysis using policy-relevant indicators. Climate Risk Management 2021;31:100265. doi: https://doi.org/10.1016/j.crm.2020.100265
14. IPCC. Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland: IPCC,, 2023.
15. PHE Heatwave Mortality Monitoring: Summer 2016. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2016.
16. PHE Heatwave Mortality Monitoring: Summer 2017. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2017.
17. PHE heatwave mortality monitoring; Summer 2018. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2018.
18. PHE Heatwave Mortality Monitoring: Summer 2019: Public Health England; 2019 [
19. Heatwave mortality monitoring report: 2020. PHE heatwave mortality monitoring. Gov.uk: Public Health England, 2020.
20. Kovats SaB, R. Health, communities and the built environment. In: The Third UK Climate Change Risk Assessment Technical Report The Third UK Climate Change Risk Assessment Technical Report Committee on Climate Change, 2021.
21. Heat mortality monitoring report: 2022. In: team EEaHP, ed. Heat mortality monitoring. gov.uk: UK Health Securety Agency, 2022.

22. Adverse Weather and Health Plan: Supporting evidence. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2023.

23. Heat and health in the WHO European Region: updated evidence for effective prevention. In: Europe WROf, ed. Copenhagen: WHO, 2021.

24. Adverse Weather and Health Plan; Protecting health from weather related harm. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2023.

25. Lorraine Williams BE, Stefanie Ettelt, Shakoor Hajat, Tommaso Manacorda, Nicholas Mays. Evaluation of the Heatwave Plan for England, FINAL REPORT. piru.ac.uk: London School of Hygiene and Tropical Medicine, 2019.

26. Roberts T, Seymour V, Brooks K, et al. Stakeholder perspectives on extreme hot and cold weather alerts in England and the proposed move towards an impact-based approach. Environmental Science & Policy 2022;136:467-75. doi: https://doi.org/10.1016/j.envsci.2022.07.012

27. Brooks K, Landeg O, Kovats S, et al. Heatwaves, hospitals and health system resilience in England: a qualitative assessment of frontline perspectives from the hot summer of 2019. BMJ Open 2023;13(3):e068298. doi: 10.1136/bmjopen-2022-068298

28. Kovats S, Johnson, H., and Griffiths, C. . Mortality in southern England during the 2003 heat wave by place of death. Health Statistics Quarterly 2006;29:6-8.

29. Stafoggia M, Forastiere F, Agostini D, et al. Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis. J Epidemiol Community Health 2008;62(3):209-15. doi: 10.1136/jech.2007.060715 [published Online First: 2008/02/15]

30. Stafoggia M, Forastiere F, Agostini D, et al. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. Epidemiology 2006;17(3):315-23. doi: 10.1097/01.ede.0000208477.36665.34 [published Online First: 2006/03/30]
31. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ 2017;357:j2099. doi: 10.1136/bmj.j2099
32. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and Ageing 2016;45(3):353-60. doi: 10.1093/ageing/afw039
33. Clift AK, Coupland CAC, Keogh RH, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. BMJ 2020;371:m3731. doi: 10.1136/bmj.m3731
34. Ooka T, Johno H, Nakamoto K, et al. Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. BMJ Nutrition, Prevention & Health 2021:bmjnph-2020-000200. doi: 10.1136/bmjnph-2020-000200
35. Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. Br J Cancer 2009;101 Suppl 2(Suppl 2):S80-S86. doi: 10.1038/sj.bjc.6605396
36. Nandy S, Parsons S, Cryer C, et al. Development and preliminary examination of the predictive validity of the Falls Risk Assessment Tool (FRAT) for use in primary care. Journal of Public Health 2004;26(2):138-43. doi: 10.1093/pubmed/fdh132
37. Syndromic surveillance: weekly summaries for 2020. GOV.UK: UKHSA, 2020.

38. Kontopantelis E, Panagioti M, Farragher T, et al. Consultation patterns and frequent attenders in UK primary care from 2000 to 2019: a retrospective cohort analysis of consultation events across 845 general practices. BMJ Open 2021;11(12):e054666. doi: 10.1136/bmjopen-2021-054666

39. Thompson R, Landeg O, Kar-Purkayastha I, et al. Heatwave Mortality in Summer 2020 in England: An Observational Study. Int J Environ Res Public Health 2022;19(10) doi: 10.3390/ijerph19106123 [published Online First: 20220518]

40. Son J-Y, Liu JC, Bell ML. Temperature-related mortality: a systematic review and investigation of effect modifiers. Environmental Research Letters 2019;14(7):073004. doi: 10.1088/1748-9326/ab1cdb

41. Murage P, Hajat S, Kovats RS. Effect of night-time temperatures on cause and age-specific mortality in London. Environmental Epidemiology 2017;1(2):e005. doi: 10.1097/ee9.0000000000000005

42. Rocklöv J, Forsberg B, Ebi K, et al. Susceptibility to mortality related to temperature and heat and cold wave duration in the population of Stockholm County, Sweden. Global health action 2014;7:22737-37. doi: 10.3402/gha.v7.22737

43. Schifano P, Cappai G, De Sario M, et al. Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome. Environmental health : a global access science source 2009;8:50-50. doi: 10.1186/1476-069X-8-50

44. Medina-Ramón M, Zanobetti A, Cavanagh DP, et al. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. Environmental health perspectives 2006;114(9):1331-36. doi: 10.1289/ehp.9074

45. Masi AT. POTENTIAL USES AND LIMITATIONS OF HOSPITAL DATA IN EPIDEMIOLOGIC RESEARCH. Am J Public Health Nations Health 1965;55(5):658-67. doi: 10.2105/ajph.55.5.658

46. Murage P, Kovats S, Sarran C, et al. What individual and neighbourhood-level factors increase the risk of heat-related mortality? A case-crossover study of over 185,000 deaths in London using high-resolution climate datasets. Environment International 2020;134:105292. doi: https://doi.org/10.1016/j.envint.2019.105292

47. Wolf T, McGregor G. The development of a heat wave vulnerability index for London, United Kingdom. Weather and Climate Extremes 2013;1:59-68. doi: https://doi.org/10.1016/j.wace.2013.07.004

48. Taylor J, Wilkinson P, Davies M, et al. Mapping the effects of urban heat island, housing, and age on excess heat-related mortality in London. Urban Climate 2015;14:517-28. doi: https://doi.org/10.1016/j.uclim.2015.08.001

49. Reid CE, O'Neill MS, Gronlund CJ, et al. Mapping community determinants of heat vulnerability. Environ Health Perspect 2009;117(11):1730-6. doi: 10.1289/ehp.0900683 [published Online First: 20090610]

50. Reid CE, Mann JK, Alfasso R, et al. Evaluation of a heat vulnerability index on abnormally hot days: an environmental public health tracking study. Environ Health Perspect 2012;120(5):715-20. doi: 10.1289/ehp.1103766 [published Online First: 20120131]

51. Nayak SG, Shrestha S, Kinney PL, et al. Development of a heat vulnerability index for New York State. Public Health 2018;161:127-37. doi: https://doi.org/10.1016/j.puhe.2017.09.006
52. Almeida SP, Casimiro E, Calheiros J. Effects of apparent temperature on daily mortality in Lisbon and Oporto, Portugal. Environmental Health 2010;9(1):12. doi: 10.1186/1476-069X-9-12
53. Breitner S, Wolf K, Devlin RB, et al. Short-term effects of air temperature on mortality and effect modification by air pollution in three cities of Bavaria, Germany: A time-series analysis. Science of The Total Environment 2014;485-486:49-61. doi: https://doi.org/10.1016/j.scitotenv.2014.03.048
54. Gasparrini A, Armstrong B, Kovats S, et al. The effect of high temperatures on cause-specific mortality in England and Wales. Occup Environ Med 2012;69(1):56-61. doi: 10.1136/oem.2010.059782 [published Online First: 2011/03/11]
55. Goodman PG, Dockery DW, Clancy L. Cause-specific mortality and the extended effects of particulate pollution and temperature exposure. Environmental health perspectives 2004;112(2):179-85. doi: 10.1289/ehp.6451
56. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? Occupational and environmental medicine 2007;64(2):93-100. doi: 10.1136/oem.2006.029017 [published Online First: 2006/09/21]
57. Huynen MM, Martens P, Schram D, et al. The impact of heat waves and cold spells on mortality rates in the Dutch population. Environmental health perspectives 2001;109(5):463-70. doi: 10.1289/ehp.01109463
58. Iñiguez C, Ballester F, Ferrandiz J, et al. Relation between temperature and mortality in thirteen Spanish cities. International journal of environmental research and public

health 2010;7(8):3196-210. doi: 10.3390/ijerph7083196 [published Online First: 2010/08/11]

59. Morabito M, Crisci A, Moriondo M, et al. Air temperature-related human health outcomes: Current impact and estimations of future risks in Central Italy. Science of The Total Environment 2012;441:28-40. doi: https://doi.org/10.1016/j.scitotenv.2012.09.056

60. Oudin Åström D, Åström C, Rekker K, et al. High Summer Temperatures and Mortality in Estonia. PloS one 2016;11(5):e0155045-e45. doi: 10.1371/journal.pone.0155045

61. Pattenden S, Armstrong B, Milojevic A, et al. Ozone, heat and mortality: acute effects in 15 British conurbations. Occupational and Environmental Medicine 2010;67(10):699. doi: 10.1136/oem.2009.051714

62. Ragettli MS, Vicedo-Cabrera AM, Schindler C, et al. Exploring the association between heat and mortality in Switzerland between 1995 and 2013. Environmental Research 2017;158:703-09. doi: https://doi.org/10.1016/j.envres.2017.07.021

63. Rocklöv J, Forsberg B. The effect of temperature on mortality in Stockholm 1998—2003: A study of lag structures and heatwave effects. Scandinavian Journal of Public Health 2008;36(5):516-23. doi: 10.1177/1403494807088458

64. Schaeffer L, de Crouy-Chanel P, Wagner V, et al. How to estimate exposure when studying the temperature-mortality relationship? A case study of the Paris area. Int J Biometeorol 2016;60(1):73-83. doi: 10.1007/s00484-015-1006-x [published Online First: 2015/05/15]

65. Wichmann J, Andersen ZJ, Ketzel M, et al. Apparent temperature and cause-specific mortality in Copenhagen, Denmark: a case-crossover analysis. International journal of

environmental research and public health 2011;8(9):3712-27. doi: 10.3390/ijerph8093712 [published Online First: 2011/09/16]

66. Sun S, Tian L, Qiu H, et al. The influence of pre-existing health conditions on short-term mortality risks of temperature: Evidence from a prospective Chinese elderly cohort in Hong Kong. Environmental Research 2016;148:7-14. doi: https://doi.org/10.1016/j.envres.2016.03.012

67. Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. European Respiratory Journal 2014;44(4):1055. doi: 10.1183/09031936.00059814

68. Ebeling M, Rau R, Malmström H, et al. The rate by which mortality increase with age is the same for those who experienced chronic disease as for the general population. Age and Ageing 2021;50(5):1633-40. doi: 10.1093/ageing/afab085

69. Khaw K-T. Healthy aging. BMJ 1997;315(7115):1090-96. doi: 10.1136/bmj.315.7115.1090

70. Liang J, Bennett JM, Shaw BA, et al. Gender Differences in Functional Status in Middle and Older Age: Are There Any Age Variations? The Journals of Gerontology: Series B 2008;63(5):S282-S92. doi: 10.1093/geronb/63.5.S282

71. van Steen Y, Ntarladima AM, Grobbee R, et al. Sex differences in mortality after heat waves: are elderly women at higher risk? Int Arch Occup Environ Health 2019;92(1):37-48. doi: 10.1007/s00420-018-1360-1 [published Online First: 2018/10/08]

72. Kaciuba-Uscilko H, Grucza R. Gender differences in thermoregulation. Curr Opin Clin Nutr Metab Care 2001;4(6):533-6. doi: 10.1097/00075197-200111000-00012

73. Dessapt AL, Gourdy P. [Menopause and cardiovascular risk]. J Gynecol Obstet Biol Reprod (Paris) 2012;41(7 Suppl):F13-9. doi: 10.1016/j.jgyn.2012.09.003

74. Kovats RS, Hajat S. Heat Stress and Public Health: A Critical Review. Annual Review of Public Health 2008;29(1):41-55. doi: 10.1146/annurev.publhealth.29.020907.090843

75. Downey L, Hawkins B. RACE, INCOME, AND ENVIRONMENTAL INEQUALITY IN THE UNITED STATES. Sociol Perspect 2008;51(4):759-81. doi: 10.1525/sop.2008.51.4.759

76. Iqbal G, Johnson MRD, Szczepura A, et al. UK ethnicity data collection for healthcare statistics: the South Asian perspective. BMC Public Health 2012;12(1):243. doi: 10.1186/1471-2458-12-243

77. Muia CM, Notley SR, Saci S, et al. Whole-body heat exchange in black-African and Caucasian men during exercise eliciting matched heat-loss requirements in dry heat. Exp Physiol 2020;105(1):7-12. doi: 10.1113/ep088091 [published Online First: 20191119]

78. Lang JA. Thermoregulation in the heat: Not so black and white. Experimental Physiology 2020;105(1):1-2. doi: https://doi.org/10.1113/EP088261

79. Gee GC, Ford CL. STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions. Du Bois Rev 2011;8(1):115-32. doi: 10.1017/s1742058x11000130

80. Holt-Lunstad J, Smith TB, Baker M, et al. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015;10(2):227-37. doi: 10.1177/1745691614568352

81. Valtorta NK, Kanaan M, Gilbody S, et al. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. Heart 2016;102(13):1009. doi: 10.1136/heartjnl-2015-308790

82. Society BG. End of Life Care in Frailty: Care homes British Geriatrics Society: British Geriatrics Society; 2020 [Available from: https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-care-homes.

83. Song JW, Chung KC. Observational studies: cohort and case-control studies. Plast Reconstr Surg 2010;126(6):2234-42. doi: 10.1097/PRS.0b013e3181f44abc

84. Moon J, Kim H-C. Case-only approach applied in environmental epidemiology: 2 examples of interaction effect using the US National Health and Nutrition Examination Survey (NHANES) datasets. BMC Medical Research Methodology 2022;22(1):254. doi: 10.1186/s12874-022-01706-6

85. Fung KY, Krewski D, Chen Y, et al. Comparison of time series and case-crossover analyses of air pollution and hospital admission data. International Journal of Epidemiology 2003;32(6):1064-70. doi: 10.1093/ije/dyg246

86. Baldwin JW, Benmarhnia T, Ebi KL, et al. Humidity's Role in Heat-Related Health Outcomes: A Heated Debate. Environ Health Perspect 2023;131(5):55001. doi: 10.1289/ehp11807 [published Online First: 20230531]

87. Nicoll A, Ciancio BC, Lopez Chavarrias V, et al. Influenza-related deaths - available methods for estimating numbers and detecting patterns for seasonal and pandemic influenza in Europe. Eurosurveillance 2012;17(18):20162. doi: doi:https://doi.org/10.2807/ese.17.18.20162-en

88. Osilla EV, Marsidi JL, Shumway KR, et al. Physiology, Temperature Regulation. StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2023, StatPearls Publishing LLC. 2023.

89. Khalid JM, Raluy-Callado M, Curtis BH, et al. Rates and risk of hospitalisation among patients with type 2 diabetes: retrospective cohort study using the UK General

Practice Research Database linked to English Hospital Episode Statistics. Int J Clin Pract 2014;68(1):40-48. doi: 10.1111/ijcp.12265 [published Online First: 2013/09/23]

90. Hajat S, Haines A, Sarran C, et al. The effect of ambient temperature on type-2-diabetes: case-crossover analysis of 4+ million GP consultations across England. Environmental Health 2017;16 doi: https://doi.org/10.1186/s12940-017-0284-7

91. Kenny GP, Sigal RJ, McGinn R. Body temperature regulation in diabetes. Temperature (Austin) 2016;3(1):119-45. doi: 10.1080/23328940.2015.1131506 [published Online First: 20160104]

92. Rihmer Z. Can better recognition and treatment of depression reduce suicide rates? A brief review. European Psychiatry 2001;16(7):406-09. doi: https://doi.org/10.1016/S0924-9338(01)00598-3

93. Vandentorren S, Bretin P, Zeghnoun A, et al. August 2003 Heat Wave in France: Risk Factors for Death of Elderly People Living at Home. European Journal of Public Health 2006;16(6):583-91. doi: 10.1093/eurpub/ckl063

94. Team TCP. Current approaches to obesity management in UK Primary Care: the Counterweight Programme. Journal of Human Nutrition and Dietetics 2004;17(3):183-90. doi: https://doi.org/10.1111/j.1365-277X.2004.00528.x

95. NHS Digital. Statistics on Obesity, Physical Activity and Diet, England, 2020. digital.nhs.uk: NHS England, 2020.

96. Dixon JB. The effect of obesity on health outcomes. Molecular and Cellular Endocrinology 2010;316(2):104-08. doi: https://doi.org/10.1016/j.mce.2009.07.008

97. Linares C, Martinez-Martin P, Rodríguez-Blázquez C, et al. Effect of heat waves on morbidity and mortality due to Parkinson's disease in Madrid: A time-series analysis.

Environ Int 2016;89-90:1-6. doi: 10.1016/j.envint.2016.01.017 [published Online First: 20160126]

98. Menezes-Rodrigues FS, Scorza CS, Fiorini AC, et al. Sudden unexpected death in Parkinson's disease: why is drinking water important? Neurodegener Dis Manag 2019;9(4):241-46. doi: 10.2217/nmt-2019-0010 [published Online First: 20190813]

99. Borg M, Nitschke M, Williams S, et al. Using the excess heat factor to indicate heatwave-related urinary disease: a case study in Adelaide, South Australia. International Journal of Biometeorology 2019;63(4):435-47. doi: 10.1007/s00484-019-01674-5

100. Kim E, Kim H, Kim YC, et al. Association between extreme temperature and kidney disease in South Korea, 2003–2013: Stratified by sex and age groups. Science of The Total Environment 2018;642:800-08. doi: https://doi.org/10.1016/j.scitotenv.2018.06.055

101. Alayyannur PA, Ramdhan DH. Relationship of heat stress with acute kidney disease and chronic kidney disease: A literature review. Journal of Public Health Research 2022;11(2):22799036221104149. doi: 10.1177/22799036221104149

102. Hajat S, Casula A, Murage P, et al. Ambient heat and acute kidney injury: case-crossover analysis of 1 354 675 automated e-alert episodes linked to high-resolution climate data. The Lancet Planetary Health 2024;8(3):e156-e62. doi: 10.1016/S2542-5196(24)00008-1

103. Roncal-Jimenez C, Lanaspa MA, Jensen T, et al. Mechanisms by Which Dehydration May Lead to Chronic Kidney Disease. Ann Nutr Metab 2015;66 Suppl 3:10-3. doi: 10.1159/000381239 [published Online First: 20150618]

104. Gong J, Part C, Hajat S. Current and future burdens of heat-related dementia hospital admissions in England. Environ Int 2022;159:107027. doi: 10.1016/j.envint.2021.107027 [published Online First: 20211207]

105. Xu Z, Tong S, Cheng J, et al. Heatwaves, hospitalizations for Alzheimer's disease, and postdischarge deaths: A population-based cohort study. Environ Res 2019;178:108714. doi: 10.1016/j.envres.2019.108714 [published Online First: 20190903]

106. Silva JE. Thermogenic mechanisms and their hormonal regulation. Physiol Rev 2006;86(2):435-64. doi: 10.1152/physrev.00009.2005

107. Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med 2011;27(1):1-15. doi: 10.1016/j.cger.2010.08.009

108. English T, Larkin M, Vasquez Hernandez A, et al. Heat Illness Requiring Emergency Care for People Experiencing Homelessness: A Case Study Series. Int J Environ Res Public Health 2022;19(24) doi: 10.3390/ijerph192416565 [published Online First: 20221209]

109. Hajat S, Sarran CE, Bezgrebelna M, et al. Ambient Temperature and Emergency Hospital Admissions in People Experiencing Homelessness: London, United Kingdom, 2011–2019. American Journal of Public Health 2023;113(9):981-84. doi: 10.2105/ajph.2023.307351

110. Hossain MM, Sultana A, Tasnim S, et al. Prevalence of mental disorders among people who are homeless: An umbrella review. International Journal of Social Psychiatry 2020;66(6):528-41. doi: 10.1177/0020764020924689

111. Layton JB LW, Yuan J, Gilman JP, Horton DB, Setoguchi S. Heatwaves, medications, and heat-related hospitalization in older Medicare beneficiaries with chronic conditions. PLoS ONE 2020;15(e0243665) doi: doi.org/10.1371/journal.pone.0243665

112. Wee J, Tan XR, Gunther SH, et al. <strong>Effects of medications on heat loss capacity in chronic disease patients: health implications amidst global warming. Pharmacological Reviews 2023:PHARMREV-AR-2022-000782. doi: 10.1124/pharmrev.122.000782

113. Coon EA, Low PA. Thermoregulation in Parkinson disease. Handb Clin Neurol 2018;157:715-25. doi: 10.1016/b978-0-444-64074-1.00043-4

114. Genever RW. Drugs for Parkinson's disease. BMJ 2009;339:b3145. doi: 10.1136/bmj.b3145

115. Hansen RA, Gartlehner G, Webb AP, et al. Efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer's disease: a systematic review and meta-analysis. Clin Interv Aging 2008;3(2):211-25.

Chapter 2 - Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records

2.1 Introduction

The first study forms the second chapter of the PhD and lays the foundations for subsequent analysis and exploration of the feasibility of using EHRs to identify individuals most at risk of death during heatwaves using their clinical information as recorded within primary care records. As previously highlighted, previous studies which have looked at individual-level risk factors have generally used hospitalisation data to characterise chronic conditions, often using differing approaches for defining the conditions of interest and representing the most severe end of the disease spectrum. In addition, gaining insight into medication used to control heat-sensitive conditions such as hypertension is challenging. Therefore, this initial study aimed to explore individual-level clinical risk factors associated with heat-related mortality in England by using primary care records and to estimate potential effect modification of a range of pre-existing conditions, clinical measurements, and prescribed medications.

To date, this is the first study to explore individual-level heat risk factors using primary care records in England, focusing on identifying pre-existing conditions, prescribed medications and clinical measurements. To do this a time-stratified case-crossover study design was employed to assess the association between temperature and mortality. The main relationship under investigation was the association between temperature and risk of death on days above specified temperature thresholds using conditional logistic regression. First, the association between temperature and mortality was modelled to assess the dose-response relationship between temperature and risk of death at temperature thresholds aligned with the new impact-based Heat-Health Alert system developed by UKHSA. Results were then stratified by population sub-groups with pre-existing conditions, medications, and clinical measurements to assess effect modification.

The study manuscript was submitted to the BMJ Public Health journal on 29 February 2024 and was published on 27 May 2024. Appendix 2 within this PhD outlines the data management approach taken when obtaining, extracting, and formatting data used in all three studies. Supplementary materials to accompany this first study is also available in Appendix 3.

2.2 Research Paper

Cover page and research paper on subsequent pages.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	2006076	**Title**	MR
First Name(s)	Ross		
Surname/Family Name	Thompson		
Thesis Title	Heatwaves and clinical vulnerability in England; development of a risk stratification tool for use in primary care		
Primary Supervisor	Sari Kovats		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?	BMJ Public Health (https://bmjpublichealth.bmj.com/content/2/1/e000927)		
When was the work published?	Ross Thompson, Sari Kovats, Shakoor Hajat, Helen Macintyre, Emer O'Connell		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	**Undertook all data formatting, coding and analysis, interpretation of results, manuscript drafting, submission and responding to reviewer comments etc. Co-authors advised throughout the process or supplied exposure data**

SECTION E

Student Signature	[redacted]
Date	28/09/2024

Supervisor Signature	[redacted]
Date	28/09/2024

BMJ Public Health

Identification of individual-level clinical factors associated with increased risk of death during heatwaves: a time-stratified case-crossover study using national primary care records in England

Ross Thompson,[1,2] Sari Kovats ,[1] Shakoor Hajat,[1] Helen Macintyre,[3,4] Emer O'Connell[1]

To cite: Thompson R, Kovats S, Hajat S, *et al*. Identification of individual-level clinical factors associated with increased risk of death during heatwaves: a time-stratified case-crossover study using national primary care records in England. *BMJ Public Health* 2024;**2**:e000927. doi:10.1136/bmjph-2024-000927

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjph-2024-000927).

Received 15 January 2024
Accepted 24 April 2024

Check for updates

For numbered affiliations see end of article.

Correspondence to
Mr Ross Thompson;
ross.thompson@lshtm.ac.uk

ABSTRACT

Background Despite an increase in heat-related deaths occurring in England in recent years, one of the key recommended actions of identifying individuals at risk and deploying targeted interventions is not routinely undertaken. A major contributing factor to this is a lack of understanding of the individual-level risk factors that would support an evidence-based approach to targeted prevention.

Objective To identify individual-level clinical risk factors for heat-related mortality in England by using primary care records and to estimate potential effect modification of a range of pre-existing conditions, clinical measurements and prescribed medications.

Methods A time-stratified case-crossover analysis was undertaken of 37 individual-level clinical risk factors. Patient's data were obtained from the Clinical Practice Research Datalink. Conditional logistic regression was used to characterise associations between temperature and the risk of death on hot days.

Results Heat mortality risk was modified by a large range of pre-existing conditions, with cardiorespiratory, mental health and cognitive function conditions, diabetes and Parkinson's, all increasing risk. The most striking increase was observed for depression with an OR of 1.25 (95% CI 1.09 to 1.44), the highest observed for pre-existing conditions. Individuals prescribed medications to treat heart failure and high blood pressure also have increased odds of death during heatwaves. There appears to be evidence of an increasing trend in ORs for diastolic blood pressure (DBP) categories, with ORs increasing from low DBP up to prehypertensive DBP group.

Conclusions This is the first study to explore a comprehensive set of individual-level clinical risk factors and heat using primary care records in England. Results presented have important implications for patient medication management during heat events, incorporating heat-risk considerations into other health policies such as suicide prevention plans and highlighted potential differences between clinical vulnerability and patients at risk.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous epidemiological studies exploring individual-level heat risk factors have generally used routine mortality and emergency hospitalisation data to characterise heat risks in persons with pre-existing conditions. This has limited use to support primary care clinicians in protecting their patients as they focus only on severe disease and provide no intelligence on risk associated with ongoing treatment or medication use.

WHAT THIS STUDY ADDS

⇒ This is the first study to explore individual-level clinical risk factors in England using primary care records.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We have demonstrated that primary care data can provide powerful insights that have implications for how clinicians prioritise patients during heatwaves, manage patient medication and implications for other health policy areas such as suicide prevention.

INTRODUCTION

Heatwaves pose a significant risk to health.[1–4] In England, heat-related deaths have been increasing since 2016.[5] In summer 2022, England experienced its first 40°C heatwave and associated level 4 Heat-Health Alert, resulting in 2985 excess deaths.[6] Following the major European heatwave in 2003, many countries introduced Heat-Health Actions Plans (HHAPs), which outline a framework for the health sector to respond to these events.[7] The UK Health Security Agency (UKHSA) launched the revised HHAP for England, the Adverse Weather and Health Plan (AWHP) in 2023 which aims to prevent avoidable harm to health during periods of adverse weather, including extreme heat. A key action for health and social care providers recommended in the AWHP is to 'establish methods to identify, alert and monitor individuals most vulnerable to heat-related illnesses on your caseload'.[8 9] This recommendation is not widely implemented[10 11] and one reason is the absence of an evidence-based process through which healthcare professionals can identify individuals most at risk of dying in a heatwave.[11 12]

The pathophysiological mechanisms of ill health during periods of heat generally result from impaired thermoregulatory (eg, vascular dilation and sweating) or behavioural responses.[13 14] At the population level, subgroups identified as at risk include older adults, children, those with chronic conditions, including mental health conditions, those prescribed certain medication and those unable to adapt their own behaviours or environments.[5 15] Epidemiological studies exploring individual heat risk factors have generally used routine mortality and emergency hospitalisation data to characterise important chronic conditions.[16–20] However, these studies provide limited evidence to support action and clinical decisions. Results from these studies are not specific to the English population leading to uncertainty about the generalisability of the findings. The use of emergency admissions data may not account for individuals receiving treatment or care within the community who do not enter the hospital system prior to death so these studies are likely biased towards individuals with more severe disease. This is especially relevant as a large proportion of deaths during heatwaves occur in the home.[21] Nor do they include information about ongoing treatment, such as prescribed medication. All of these may play a significant role in an individual's overall risk and may underestimate the effects of these risk factors.

Therefore, the aim of this study is to identify individual-level clinical risk factors for heat-related mortality in England by using primary care records and to estimate the potential effect modification of a range of pre-existing conditions, clinical measurements and prescribed medications. Results from this study provide foundational evidence on which methodologies and processes can be established to help healthcare professionals in identifying individuals most at risk so that targeted interventions can be deployed.

METHODS

Study population

We used the Clinical Practice Research Datalink (CPRD) Aurum (ID number 21_000621) to link primary care records, Office for National Statistics (ONS) mortality data and National Health Service hospitalisation data, at the individual level. The outcome was defined as all deaths which occurred between May 2016 and September 2020 using ONS date of death. CPRD Aurum has been shown to be representative of the English population.[22]

Primary care records were used to identify individuals with pre-existing disease, prescribed medication and blood pressure measurements. Records were valid if they were recorded within 2 years of death to better reflect a diagnosis at the time of death (ie, records between 2014 and 2020).[16 17] Where there was more than one relevant record within the 2-year window, the record closest to the date of death was used. Pre-existing conditions and prescribed medication groups were selected a priori based on published evidence, physiological plausibility and data availability as presented in online supplemental file S1. This resulted in a total of 37 priority variables identified (see table 1). Age and gender were not included in this analysis, however, will be in further analysis exploring wider determinants of health. A combination of bespoke and published clinical code lists were used to create risk factor variables for all individuals in the study sample. Clinical code lists are available in online supplemental table S1.

Exposure data

Geographical information on individuals in CPRD is limited to the UK government region of their registered primary care practice. Therefore, a daily mean population-weighted regional temperature series was generated for the study period (May 2016–September 2020). Daily mean temperatures were generated from HadUK-grid daily maximum and minimum temperatures.[23 24] The gridded mean daily temperature series was then combined with 100 m gridded population data using ArcGIS to create regional population-weighted temperature series which were then assigned to each individual based on their general practitioner (GP) practice region. A lag period of 0–2 (3 days) were also calculated and assigned to each individual to estimate delayed and cumulative effects of exposure over the 3 days. Heat effects are known to be mostly immediate so impacts at longer lags were not considered.[25]

Statistical analysis

A time-stratified case-crossover study design was employed to assess the association between temperature and mortality. Within this study design, the temperature on the day of death (event-day) is compared with non-event days. The main relationship under investigation is the association between temperature and risk of death on days above specified temperature thresholds using conditional logistic regression. The case serves as its own control and therefore the potential effect of time-independent confounding factors such as age or gender are automatically controlled for. Case days were determined as the date of death. Control days were selected following a bidirectional referent selection approach, to be the same day of the week of the same month in which the death occurred, resulting in each case having at least three controls, reducing the potential for overlap bias.[26] This is a popular and robust approach used within environmental epidemiology for estimating the association

Table 1 Overview of data used in analysis with a total number of individuals with a valid primary care record that occurred within 2 years of death by subnational region, pre-existing condition, prescribed medication and blood pressure measurements

Variable	Individuals	Proportion
All persons	430682	100.00%
Subnational regions	430682	100.00%
The North (NE, NW and Y&H)	113405	26.33%
Midlands and East (WM, EM, EoE)	102630	23.83%
London	65145	15.13%
The South (SW and SE)	149502	34.71%
Diastolic blood pressure	232518	53.99%
Low (<60 mm Hg)	25416	5.90%
Normal (60–79 mm Hg)	142421	33.07%
Prehypertensive (80–89 mm Hg)	49781	11.56%
Hypertension Stage 1 (90–99 mm Hg)	11588	2.69%
Hypertension Stage 2 (100 mm Hg+)	3336	0.77%
Hypertension (1 and 2)	14924	3.47%
Systolic blood pressure	232785	54.05%
Low (<80 mm Hg)	1442	0.33%
Normal (80–119 mm Hg)	80680	18.73%
Prehypertensive (120–139 mm Hg)	96646	22.44%
Hypertension stage 1 (140–159 mm Hg)	43418	10.08%
Hypertension stage 2 (160 mm Hg+)	10623	2.47%
Hypertension (1 and 2)	54041	12.55%
Alzheimer's and dementia	34610	8.04%
Anxiety	11343	2.63%
Arrythmia	32279	7.49%
Asthma	9118	2.12%
Bipolar disorder	726	0.17%
Cardiac arrest	2084	0.48%
Cardiomyopathy	1148	0.27%
Chronic kidney disease	25252	5.86%
Chronic obstructive pulmonary disease	11655	2.71%
Depression	5705	1.32%
Diabetes	48391	11.24%
Emphysema	1847	0.43%
Haemorrhage	3222	0.75%
Heart failure	22345	5.19%
Hyperthyroidism	718	0.17%
Hypothyroidism	7908	1.84%
Severe learning disability	1135	0.26%
Chronic liver disease	3738	0.87%
Myocardial infarction	7338	1.70%
Occlusion	5841	1.36%
Other cerebrovascular diseases	1407	0.33%
Parkinson's disease	4970	1.15%
Psychosis	6105	1.42%

Continued

Table 1 Continued

Variable	Individuals	Proportion
Schizophrenia	942	0.22%
Severe mental illness	2412	0.56%
Stroke	11 736	2.72%
Prescribed medication		
Ace inhibitors	64 959	15.08%
Beta blockers	59 996	13.93%
Cardio glycols	12 136	2.82%
Diuretics	71 262	16.55%
Non-steroidal anti-inflammatory drugs	49 020	11.38%
Vasoconstrictor drugs	547	0.13%
Anticholinergic drugs	8952	2.08%
Total number of practices contributing to sample=1476 Mean number of patients per practice=291.79		
Maximum number of patients per practice=2400		
Minimum number of patients per practice=6		

Severe mental health is a composite indicator of severe mental health conditions including schizophrenia, bipolar disorder, manic episodes, etc. Clinical code list available in online supplemental materials.
Summary statistics also provided for continuous variables which were categorised for analysis.
EM, East Midlands; EoE, East of England; NE, Northeast; NW, Northwest; SE, Southeast; SW, Southwest; WM, West Midlands; Y&H, Yorkshire and the Humber.

of acute health events in response to short-term exposures.[16 17 27 28]

First the association between temperature and mortality was modelled to assess the dose–response relationship between temperature and risk of death. This was carried out using natural cubic spline functions, with internal knots determined using the Akaike information criterion to define the best model fit. From this initial model, temperature thresholds for analysis were derived. To maximise policy relevance, thresholds were selected using the approach used by the UKHSA for defining the 'low' impact level of the new impact-based Heat-Health Alert system.[29] That is, the temperature is associated with a relative risk (RR) of 1.1, that is, a 10% elevated risk of death. The 'medium' impact threshold, defined as the temperature associated with an RR of 1.2, was used in sensitivity analysis. The 'high' impact threshold was not used in this analysis due to daily temperatures within the study period not reaching the required 40°C temperatures, as defined by UKHSA. Relative thresholds were derived for the national-level analysis and for subnational-level analysis, also carried out as sensitivity analysis. The reference temperature for the conditional logistic regression was taken as the minimum mortality temperature (MMT) which is the temperature at which risk of death is lowest. Results were stratified by condition/medication to assess effect modification.

All results are reported as ORs with 95% CIs and p values. In addition, to aid in the assessment of the potential modifying effect of each subgroup, a relative effect modification (REM) index was calculated as the specific OR of an individual-level factor compared with a reference category. For categorical factors, such as blood pressure, the reference category was taken as 'normal blood pressure'. For all binary variables, however, the reference category was taken as the OR estimate for the whole population. All analyses were carried out in Stata Statistical Software: release V.17.[30]

Sensitivity analysis

First, the analysis was repeated using the 'medium impact' threshold to assess any differences in the patterns of ORs. Second, the analysis was again repeated at subnational level to assess geographical variations in estimated associations. Subnational-level analysis was carried out using the following regional groups: London; The North (combined North East, North West and Yorkshire and Humber); Midlands and East (combined West Midlands, East Midlands and East of England) and the South (South West and South East). Regions were combined to ensure that the frequency of the events (deaths, heat-health alerts) was sufficient to support the analysis and based on study population frequencies, the study population distribution compared with the national distribution over the study period, the number of Heat-Health Alerts issued over the study period, geographical location and climate.

Third, the analysis was carried out using hospitalisation data to define pre-existing disease status. This was to assess any significant differences in the effect sizes observed between the two approaches to defining pre-existing conditions. This was carried out for a limited number of conditions which had strong evidence for a significant association within the main analysis.

And finally, we assessed the potential confounding effect of background air pollutants on a restricted number of pre-existing conditions and medications. Particulate matter (PM_{10}), ozone (O_3) and nitrogen dioxide (NO_2) are all potential confounders in the association of mortality and heatwaves.[31] Due to data limitations, the analysis was restricted to London and daily mean NO_2, PM_{10} and O_3 concentration values. Five urban background air quality monitoring sights were selected across London from the London Air Quality Network.[32] Using daily mean values for each site, a London-wide daily mean background value was derived for each pollutant and assigned to each individual in London. As with temperature, a lag period of 0–2 days was also calculated and assigned to each individual.

Patient and public involvement

The Health Protection Research Unit in Environmental Change and Health has developed a public engagement/involvement group called PLANET, which was established in Autumn of 2020. PLANET stands for Public Led and Knowledge Engagement Team. The group has 30 members and approximately 20 attend the regular meetings every 3–4 months to discuss research projects. A session exploring heat risk and the use of primary care records to identify those at highest risk was carried out in May 2022 where the aims and objectives were presented to the group for feedback. Results from this study will be presented to the PLANET group in early 2024 at an annual meeting, with the meaning and implications of the results discussed, and potential future research in this area explored.

RESULTS

430 682 individuals who died during the study period between May 2016 and September 2020, from 1476 primary care practices were included in the study population. A full breakdown of the numbers of individuals with suitable records for each individual-level factor considered within this analysis is detailed in table 1. The study population age profile indicates that it is heavily skewed towards the older population and is aligned with the national age distribution. Details of exposure data are available in online supplemental table S2.

Temperature thresholds

Figure 1 shows the temperature–mortality relationship derived using the full data series (ie, all individuals within the study population) and the temperature thresholds derived. The policy-relevant thresholds, when rounded to the nearest 0.5°C equate to 17°C (the MMT), 22°C and 24°C. Using these thresholds to identify cases resulted in 13 970 using the ‘low’ impact threshold and 10 187 using the ‘medium’ impact threshold. Temperature thresholds used for subnational-level sensitivity analysis are reported in online supplemental table S3.

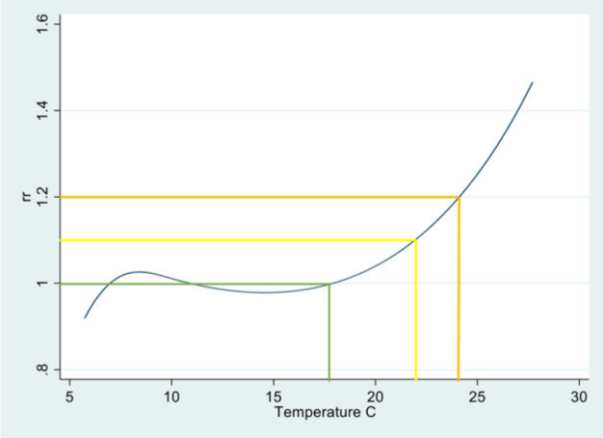


Figure 1 National temperature-mortality relationship plot of relative risk (RR) and mean temperature from which temperature thresholds used in analysis were derived. Green line represents the MMT with an RR=1.0 which equates to about 17°C (rounded to the nearest 0.5°); yellow line indicates the UKHSA defined low impact threshold with an RR of 1.1 which equates to 22°C; amber lines indicate the UKHSA defined Medium Impact threshold with an RR of 1.2 which equates to 24°C which was used in sensitivity analysis. MMT, minimum mortality temperature. UKHSA, UK Health Security Agency

Pre-existing conditions

Heat mortality risk was modified by a large range of pre-existing conditions when comparing the odds of death at the MMT (17°C) and ‘low impact’ temperature (22°C) as illustrated in figure 2, with cardiorespiratory conditions all increasing risk. Most striking, however, is that 4 out of 12 conditions with strong evidence ($p<0.01$) of an association are mental health or cognitive function conditions, including depression, psychosis, severe mental health conditions and Alzheimer’s and dementia. Depression also had the largest relative modification effect of all pre-existing conditions with an REM index value of 1.15 compared with the whole population, however, 95%CIs overlap. Diabetes and Parkinson’s disease were also found to modify risk of death during periods of heat.

Prescribed medication

Individuals prescribed medications to treat heart failure and high blood pressure had increased odds of death during heat episodes (22°C+) compared with non-heat days (≤17°C). The point estimate for patients prescribed vasodilators is relatively high, at 1.83 (1.19 to 2.80), with large 95% CIs due to the small number of individuals in this subgroup. There was also strong evidence that non-steroidal anti-inflammatory drugs increase the odds of death during heatwaves. Evidence for an association with anticholinergic drugs was weak, although the point estimate is comparable to other similar medication classes investigated.

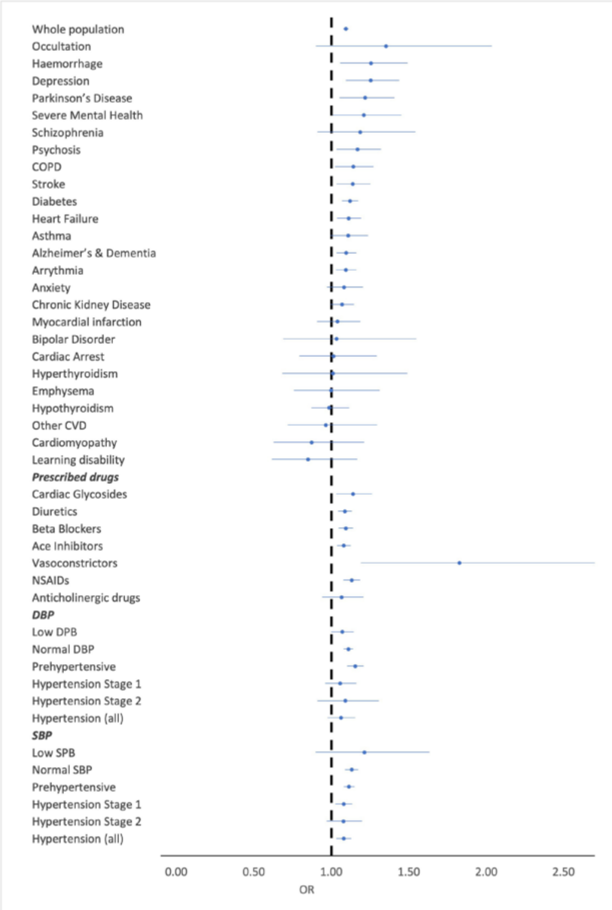


Figure 2 Forest plot of OR and 95% CI for all chronic conditions, prescribed medication groups and results for diastolic and systolic blood pressure categories. COPD, chronic obstructive pulmonary disease; DBP, diastolic blood pressure; SBP, systolic blood pressure; CVD, cardiovascular diseases.

Blood pressure

There appears to be evidence of an increasing trend in ORs for diastolic blood pressure (DBP), with ORs increasing from low DBP up to prehypertensive DBP. However, this trend is not apparent in patients with hypertension 1 and 2 groups, where ORs reduce considerably. The opposite trend is observed for systolic blood pressure (SBP) categories with a decreasing trend in OR estimates with each increasing category of SBP (figure 2).

Sensitivity analysis 1: 'medium' impact threshold and subnational-level analysis

For the first sensitivity analysis, in general, the patterns observed in the study population at national level were mirrored at subnational level. However, some additional conditions displayed strong evidence for an association at subnational level that were not evident at the national level (see online supplemental table S4). Similarly, patterns in ORs reported here were replicated when using the higher, medium impact temperature thresholds. Full results are available in online supplemental table S5.

Sensitivity analysis 2: use of emergency admissions data to define pre-existing disease

Overall, ORs were markedly consistent across the two data sources used to define the conditions investigated (figure 3A). A number of the conditions are associated with higher ORs using GP data than with emergency hospitalisation data, such as haemorrhage, CPOD and particularly depression; however, the 95% CIs for both data sources overlap.

Sensitivity analysis 3: adjusting for background air quality indicators in London

The air-quality-adjusted estimates for all variables were consistent across models. Figure 3B provides comparison of variables with strong evidence of an association between increased odds of death during periods of heat in the unadjusted model and the adjusted model. The adjusted and unadjusted OR estimates are broadly similar for all conditions apart from chronic obstructive pulmonary disease (COPD) which has a considerably reduced OR estimate when adjusted for air pollutant concentrations.

DISCUSSION

To our knowledge, this study is the first to explore individual-level risk factors associated with heatwaves as recorded within primary care records in England. This study has shown very clearly that risk of mortality increases during periods of heat for individuals with a range of pre-existing disease including cardiorespiratory conditions, mental health and cognitive function conditions, diabetes and Parkinson's disease, with the largest increases observed for those with a record of depression and haemorrhage in the 2 years proceeding death. In addition, we also demonstrate that individuals prescribed non-steroidal anti-inflammatory drugs (NSAIDs) and medications used to treat high blood pressure and heart failure are also at increased risk of death during periods of heat. We have explored the role of air pollution as a concurrent risk, providing evidence that it may be the dominant exposure of concern for patients with COPD but may be less important for other patient groups affected by heat exposure. Finally, we have identified an unexpected pattern in risk by DBP groups.

The observed patterns in ORs by individual-level clinical risk factors investigated were remarkably consistent across each sensitivity analysis. It appears our results are unlikely to be confounded by air pollution concentrations, except for COPD. When we compared our results using primary care data to those defined by emergency hospitalisation data, ORs were again very consistent, however, estimates were higher for some individual-level factors defined by primary care records. This suggests that for some pre-existing conditions which are more

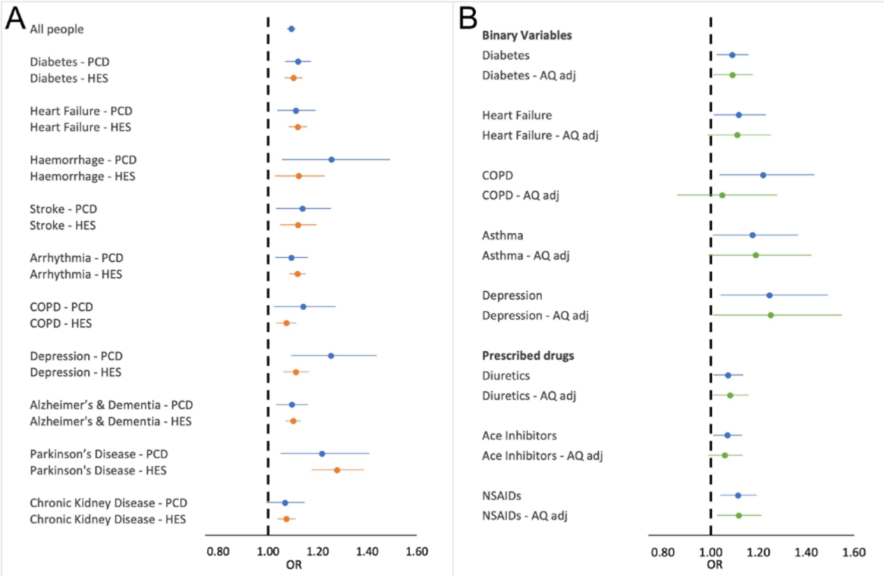


Figure 3 Sensitivity analysis forest plots. (A) The ORs and 95%CI of death on hot day for chronic conditions defined using primary care data (PCD) indicated in blue compared to those same conditions as defined using HES admitted patient care emergency admissions data (HES), indicated as orange. (B) Unadjusted ORs and 95%CI of death on a hot day in London (blue) and the OR and 95%CI estimates adjusting for background daily mean concentrations ozone, nitrogen dioxide and PM10 in London (green). COPD, chronic obstructive pulmonary disease; NSAIDs, non-steroidal anti-inflammatory drugs; PM10, particulate matter; HES, hospital episode statistics.

routinely treated within primary care, primary care records may be more sensitive to identifying a signal for heat risk.

Some findings align with population-level evidence within the broader literature, specifically that odds of death for individuals with circulatory and respiratory diseases is increased during heatwaves.[15 16 33] However, one unexpected observation was that individuals classed as hypertensive (DBP) did not follow the overall trend of increasing ORs with increasing DBP group. This contradicts previous studies that suggest this group may be at increased risk.[34] While physiological pathways suggest this group is potentially more vulnerable to the effects of high temperatures, our results raise questions about potential differences in physiological vulnerability and those actually at-risk during heatwaves. One plausible explanation may be linked to the management of hypertension via clinical interventions. This would align with evidence observed elsewhere suggesting that the level of care received by individuals, regardless of high temperatures, may reduce heat risk in individuals receiving treatment.[17] This potentially has important implications for how clinicians prioritise patients in their care during heat events who might not be the most physiologically vulnerable but at the highest level of risk.

One of the most striking observations was the association between mental health conditions and increased odds of death during heat, with depression particularly standing out. This not only has implications for HHAPs but also extends to suicide prevention strategies, given the link between depression as a risk factor for suicide and the strong evidence for increased risk of suicide during periods of high temperature.[35] Our finding that individuals with Alzheimer's and dementia, Parkinson's disease and diabetes are at risk during heatwaves also aligns with recent population-level research.[28 33 36] These conditions are all associated with older age and emphasise the need for targeted and efficient responses as the number and proportion of older people in our population grows.[37] Integrating heat-risk considerations into broader health agendas, especially within the context of evolving person-focused care in community models, is becoming crucial.

As with the existing literature, our finding of increased risk for individuals prescribed medications for heart failure or blood pressure[38] underscores the need for evidence-based individually-tailored medicine management with presummer reviews to reduce risk during heatwaves. Currently, we are unaware of any guidance for clinicians or patients which addresses this need. Additionally, our finding that NSAIDs appear to increase risk during periods of heat suggests there is a need for evidenced-based clinical guidance beyond just cardiovascular drugs. However, due to the diversity in drug types and modes of action, we could not explore all potential medications that could increase heat-related risks. This is a priority area that requires further research.

Limitations of the study

Geographical resolution of the health data limits precise exposure assignment for each individual within the study. However, previous studies[1] have demonstrated the high correlation between temperature monitoring stations within English regions, and that it is possible to characterise exposure well using a regionally representative temperature series. In addition, inconsistencies in primary care consultation records and the specific terms used when recording the details by clinicians increase the potential for some relevant records to be missed. However, a systematic approach for identifying relevant records which included clinical validation was used and should address this limitation. While CPRD is representative of the English population[22] it does not have full coverage of England and is not geographically representative.[22] Episode analysis of previous heatwaves suggests most heatwave-related deaths occur in the south.[6 39–44] Therefore, it is not anticipated that this limitation would affect our results significantly. Within this analysis, it was not possible to investigate the potential multiple interactions between different individual-level risk factors, and how this may modify an individual's overall risk. Nor did we investigate wider determinants of health which may be recorded within primary care data, that might provide additional intelligence on who is most at risk. While it is well documented that the older population are at increased risk, the focus of this study was to look at specific clinical and diagnostic criteria recorded within primary care records. These additional elements are critical areas which require further research.

CONCLUSIONS

This is the first study to explore individual-level clinical risk factors and heat using primary care records in England and has highlighted important factors associated with increased risk of death during heatwaves. We have demonstrated that primary care data can provide useful insights into important differences between those who might be characterised as clinically vulnerable and those who are at greater risk. Results from this study suggest implications beyond HHAPs alone but transcend into other health policy areas, for example, suicide prevention plans and healthy ageing. In addition, more research is needed on the role that a wide range of medication use potentially has in modifying heat risk. Research is also required to explore other factors recorded within primary care data to gain further insight into heat risk at the individual level. This further intelligence could then be combined with insights from this study to develop an evidenced-based approach to identifying individuals most at risk within the community so that targeted interventions can be deployed.

Author affiliations
[1]NIHR Health Protection Research Unit in Environmental Change and Health, London School of Hygiene & Tropical Medicine, London, UK
[2]Extreme Events and Health Protection, Centre for Climate and Health Security, UK Health Security Agency, London, UK
[3]Climate and Health Assessment Team, Centre for Climate and Health Security, UK Health Security Agency, London, UK
[4]School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

Contributors Conceptualisation: RT, SK, SH and EO'C; Supply of data: HM; Data curation: RT; Formal analysis: RT; Methodology: RT, SK and SH; Analysis interpretation: RT; SK, SH and EO'C; Manuscript drafting and editing: RT, SK, SH, HM and EO'C; Guarantor: RT.

Funding The time of all authors was funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Environmental Change and Health (NIHR200909), a partnership between UK Health Security Agency, the London School of Hygiene & Tropical Medicine, University College London and the Met Office.

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval Ethics approval for this study was provided by the London School of Hygiene & Tropical Medicine Observational/Interventions Research Ethics Committee; Ethics Committee Reference: 27995.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Access to CPRD data, including UK Primary Care Data and linked data such as Hospital Episode Statistics, is subject to protocol approval via CPRD's Research Data Governance (RDG) Process.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

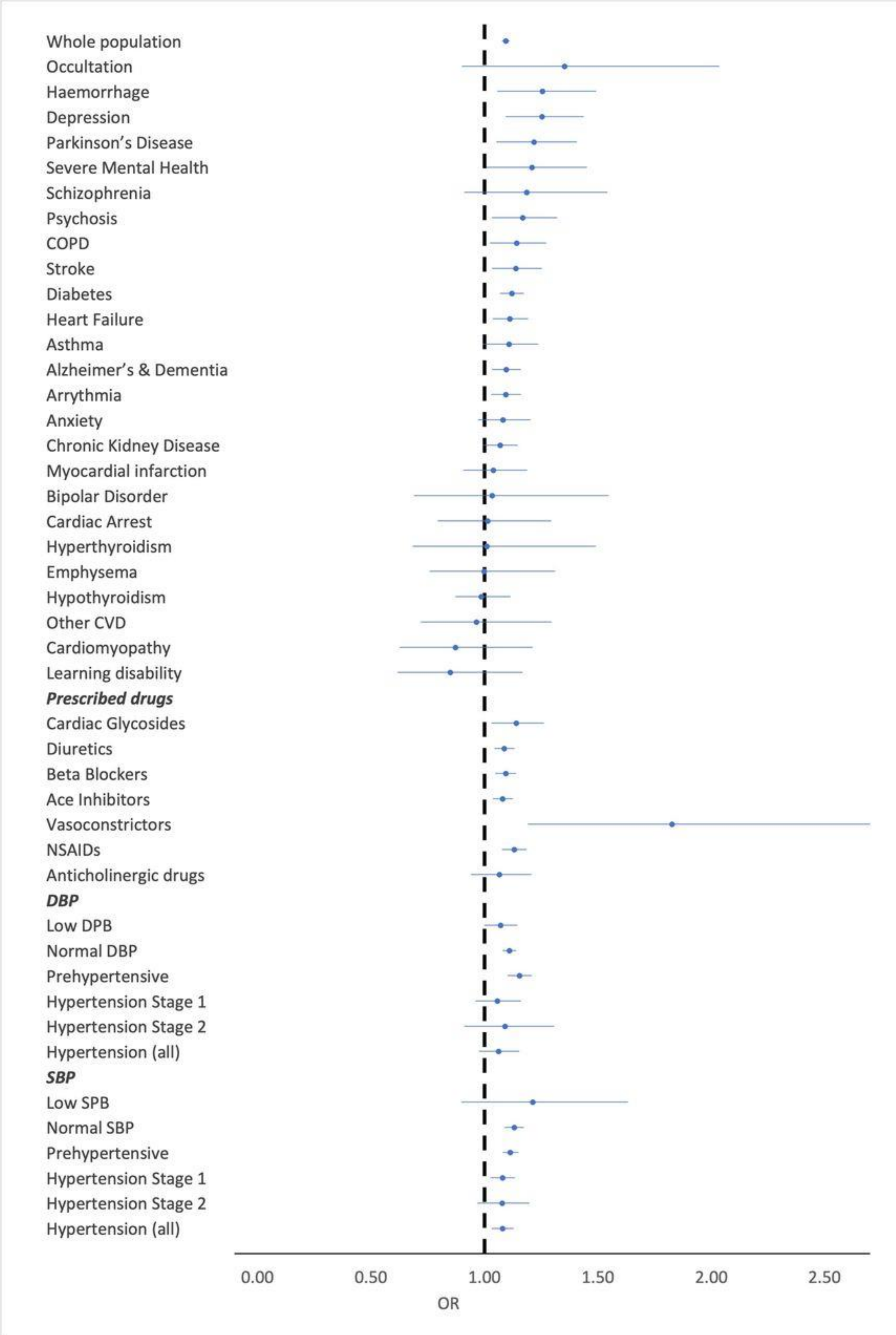
ORCID iD
Sari Kovats http://orcid.org/0000-0002-4823-8099

REFERENCES

1. Armstrong BG, Chalabi Z, Fenn B, *et al*. Association of mortality with high temperatures in a temperate climate: England and Wales. *J Epidemiol Community Health* 2011;65:340–5.
2. Ebi KL, Capon A, Berry P, *et al*. Hot weather and heat extremes: health risks. *Lancet* 2021;398:698–708.
3. Hajat S, Kovats RS, Atkinson RW, *et al*. Impact of hot temperatures on death in London: a time series approach. *J Epidemiol Community Health* 2002;56:367–72.
4. Kovats RS, Hajat S. Heat stress and public health: a critical review. *Annu Rev Public Health* 2008;29:41–55.
5. U.H.S. Agency. Adverse weather and health plan: supporting evidence, U.H.S. 2023. Available: https://www.gov.uk/government/publications/adverse-weather-and-health-plan
6. UK Health Securety Agency. Heat mortality monitoring report: 2022. In: *Heat mortality monitoring*. UK Health Securety Agency: gov.uk, 2022.
7. Matthies F, Bickler G, Marín NC. *Heat-health action plans: guidance*. World Health Organization, 2008.
8. U.H.S. Agency. Adverse weather and health plan; protecting health from weather related harm, HM Government. 2023. Available: https://www.gov.uk/government/publications/adverse-weather-and-health-plan
9. WHO. *Heat and health in the WHO European Region: updated evidence for effective Prevention*. Copenhagen: WHO, 2021.

10 Lorraine Williams BE, Ettelt S, Hajat S, *et al*. Evaluation of the Heatwave plan for England, Final report. London School of Hygiene and Tropical Medicine; 2019. Available: iru.ac.uk
11 Brooks K, Landeg O, Kovats S, *et al*. Heatwaves, hospitals and health system resilience in England: a qualitative assessment of frontline perspectives from the hot summer of 2019. *BMJ Open* 2023;13:e068298.
12 Wilson L, Black D, Veitch C. Heatwaves and the elderly: the role of the GP in reducing morbidity. *Aust Fam Physician* 2011;40:637–40.
13 Osilla EV, Marsidi JL, Shumway KR, *et al*. Physiology, temperature regulation. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, Copyright © 2023, StatPearls Publishing LLC, 2023.
14 Székely M, Carletto L, Garami A. The pathophysiology of heat exposure. *Temperature (Austin)* 2015;2:452.
15 Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk *Occup Environ Med* 2007;64:93–100.
16 Stafoggia M, Forastiere F, Agostini D, *et al*. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology* 2006;17:315–23.
17 Stafoggia M, Forastiere F, Agostini D, *et al*. Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis. *J Epidemiol Community Health* 2008;62:209–15.
18 Rocklöv J, Forsberg B, Ebi K, *et al*. Susceptibility to mortality related to temperature and heat and cold wave duration in the population of Stockholm County, Sweden. *Glob Health Action* 2014;7:22737.
19 Schifano P, Cappai G, De Sario M, *et al*. Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome. *Environ Health* 2009;8:50.
20 Medina-Ramón M, Zanobetti A, Cavanagh DP, *et al*. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environ Health Perspect* 2006;114:1331–6.
21 MASI AT. Potential uses and limitations of hospital data in epidemiologic research. *Am J Public Health Nations Health* 1965;55:658–67.
22 Wolf A, Dedman D, Campbell J, *et al*. Data resource profile: clinical practice research Datalink (CPRD) Aurum. *Int J Epidemiol* 2019;48:1740–1740g.
23 Hollis M, Kendon M, Legg T, *et al*. *HadUK-Grid gridded and regional average climate observations for the UK*. 2018.
24 Hollis D, McCarthy M, Kendon M, *et al*. Haduk-grid—a new UK dataset of gridded climate observations. *Geoscience Data Journal* 2019;6:151–9.
25 Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006;17:624–31.
26 Austin H, Flanders WD, Rothman KJ. Bias arising in case-control studies from selection of controls from overlapping groups. *Int J Epidemiol* 1989;18:713–6.
27 Bhaskaran K, Hajat S, Armstrong B, *et al*. The effects of hourly differences in air pollution on the risk of myocardial infarction: case crossover analysis of the MINAP database. *BMJ* 2011;343:d5531.
28 Hajat S, Haines A, Sarran C, *et al*. The effect of ambient temperature on type-2-diabetes: case-crossover analysis of 4+ million GP consultations across England. *Environ Health* 2017;16:73.
29 U.H.S. Agency. Adverse weather and health plan; user guide - new impact-based weather-health alerting system. 2023. Available: https://www.gov.uk/government/publications/adverse-weather-and-health-plan
30 StataCorp. *Stata statistical software: release 17*. StataCorp LLC: College Station, TX, 2021.
31 Stafoggia M, Michelozzi P, Schneider A, *et al*. Joint effect of heat and air pollution on mortality in 620 cities of 36 countries. *Environ Int* 2023;181:108258.
32 Impirial College London. London air. 2023. Available: https://www.londonair.org.uk/
33 Thompson R, Landeg O, Kar-Purkayastha I, *et al*. Heatwave mortality in summer 2020 in England: an observational study. *Int J Environ Res Public Health* 2022;19:6123.
34 Liu J, Varghese BM, Hansen A, *et al*. Heat exposure and cardiovascular health outcomes: a systematic review and meta-analysis. *Lancet Planet Health* 2022;6:e484–95.
35 Thompson R, Hornigold R, Page L, *et al*. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. *Public Health* 2018;161:171–91.
36 Linares C, Martinez-Martin P, Rodríguez-Blázquez C, *et al*. Effect of heat waves on morbidity and mortality due to Parkinson's disease in Madrid: a time-series analysis. *Environ Int* 2016;89–90:1–6.
37 Department of Health and Social Care. Chief medical officer's annual report 2023: health in an ageing society. 2023.
38 Layton JB, Li W, Yuan J, *et al*. Heatwaves, medications, and heat-related hospitalization in older medicare beneficiaries with chronic conditions. *PLoS ONE* 2020;15:e0243665.
39 Public Health England. PHE Heatwave mortality monitoring: summer 2016. In: *PHE Heatwave Mortality Monitoring*. Public Health England, 2016.
40 Public Health England. PHE Heatwave mortality monitoring: summer 2017. In: *PHE Heatwave Mortality Monitoring*. Public Health England, 2017.
41 Public Health England. PHE Heatwave mortality monitoring; summer. In: *PHE Heatwave Mortality Monitoring. 2018*. Public Health England: gov.uk, 2018.
42 Research and analysis. PHE Heatwave mortality monitoring: summer 2019. 2019.
43 Public Health England. Heatwave mortality monitoring report: 2020. In: *PHE heatwave mortality monitoring*. Public Health England, 2020.
44 UK Health Securety Agency. Heat mortality monitoring report: 2021. In: *Heat mortality monitoring*. 2021.

High resolution version of figure 2 from Thompson R, *et al. BMJ Public Health 2024;2:e000927*



Chapter 3 - Social determinants of heat related mortality in England – a time-stratified case-crossover study using primary care records

3.1 Introduction

This second study forms the third chapter of the PhD and builds upon the results from the first study. The pathway to ill health during periods of heat is complex and involves a combination of exposures, individual-level risk factors and adaptive capacity of the individual. Population-level epidemiological studies exploring individual-level heat risk have generally used routine mortality and emergency hospitalisation data, which do not consider the role of socio-environmental factors. Where wider determinants of health and heat risk have been investigated, these have generally been linked to restricted registries or using a small number of proxy measures where individuals are assigned a relevant category based solely on their geographic location.

Whilst patient record systems in primary care are predominantly used for managing clinical care, they also contain other, non-clinical types of data that may be relevant to heat risk, and which are currently used within primary care practice for other health assessments, such as Electronic Frailty Index and QRISK. Building upon results from the first study presented in Chapter 2, the aim of this second study was to identify individual-level socio-environmental risk factors for heat-related mortality in England using primary care records and to use these data to estimate the potential effect modification of a range of wider determinants of health. The same methodological approach was undertaken for this study, using the case-crossover study design to assess the association between temperature and mortality, with results stratified by sub-population groups.

The study manuscript was submitted to the BMJ Public Health journal on 01 March 2024 and is currently out for external peer review. Supplementary materials to accompany this study are available in Annex 4.

3.2 Research paper

Cover page and research paper on subsequent pages.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	2006076	**Title**	MR
First Name(s)	Ross		
Surname/Family Name	Thompson		
Thesis Title	Heatwaves and clinical vulnerability in England; development of a risk stratification tool for use in primary care		
Primary Supervisor	Sari Kovats		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

<u>SECTION C – Prepared for publication, but not yet published</u>

Where is the work intended to be published?	BMJ Public Health Submitted 17 Aug 2024
Please list the paper's authors in the intended authorship order:	Ross Thompson, Sari Kovats, Shakoor Hajat, Emer O'Connell
Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Undertook all data formatting, coding and analysis, interpretation of results, manuscript drafting, submission and responding to reviewer comments etc. Co-authors advised throughout the process

SECTION E

Student Signature	[redacted]
Date	28/09/2024

Supervisor Signature	[redacted]
Date	28/09/2024

Social determinants of heat related mortality in England – a time-stratified case-crossover study using primary care records

Authors:

Ross Thompson[1,2], Sari Kovats[1], Shakoor Hajat[1], Emer O'Connell[1]

Affiliations:

[1] NIHR Health Protection Research Unit in Environmental Change and Health, London School of Hygiene and Tropical Medicine, London

[2]Extreme Events and Health Protection Team, UK Health Security Agency, London

Abstract:

Background: Despite increases in heat-related deaths in England, there has been limited progress in developing interventions in primary care that identify and target individuals at risk. Lack of understanding of individual-level socio-environmental risk factors limit development of an evidence-based approach to targeted prevention.

Objective: To identify individual-level non-clinical risk factors for heat-related mortality in England using primary care records and to assess the potential of these socio-environmental factors as effect modifiers for the association between ambient temperature and death.

Methods: A time-stratified case-crossover analysis was undertaken of 9 potential risk factors at the individual-level and categorised into risk factor sub-groups. 430,682 patients with valid records were included in the study population, obtained from the Clinical Practice Research Datalink. Conditional logistic regression was used to characterise associations between temperature and the risk of death on hot days and to investigate the modifying effect of each risk factor.

Results: Older ages, females, ethnic minorities, and those living in the most deprived areas all had increased risk of death during periods of heat. An increasing trend in odds ratios were observed with increasing amounts of alcohol intake and increasing body mass index, excluding the obese-3 group. No differences in risks were observed by marital status or frailty category.

Conclusions: This is the first study in England to assess the role of socio-environmental factors in modifying heat risk at an individual level. The results provide important evidence on the role of disadvantage in driving the inequitable distribution of climate change impacts, and the need for better socio-economic data linked to health records. For clinical practice, the findings highlight the importance of incorporating an assessment of individual socio-environmental circumstances when prioritising patients at highest risk during heat events.

Key Messages:

What is already known about the subject?

The pathway to ill health during periods of heat is complex and involves a combination of exposures, individual level risk factors and adaptive capacity of the individual. Population-level epidemiological studies exploring individual-level heat risk have generally used routine mortality and emergency hospitalisation data which do not contain data on wider social determinants of health which are known to influence risk.

What are the new findings?

We explore socio-environmental factors modifying heat risk at an individual-level using primary care records for the first time in England. We identify ethnicity and deprivation as significant risk modifying factors in England for the first time, along with unexpected patterns in risk by BMI and frailty.

How might these results change the focus of future research or clinical practise?

We demonstrate that primary care data can provide powerful insights that have implications for patient management during heat events, highlight the complexity of heat risk and the role of socio-environmental factors in driving that risk and further underscore the urgency of policy action that is required to address health inequalities observed in heat associated mortality during heat events in England.

Introduction

Heatwaves and high temperatures pose significant risks to health.[1-5] In England, there is an increasing trend in total heat-associated mortality.[6-8] 2022 observed the highest heat mortality value following the first 40°C heatwave and associated Level 4 Heat-Health Alert and RED Extreme Heat warning, and resulted in 2,985 heat associated deaths.[8, 9] Following the pan-European heatwave in 2003, many countries and cities introduced Heat-Health Action Plans which set out a framework to plan for and respond to these adverse weather events.[10] In 2023, the UK Health Security Agency (UKHSA) launched the Adverse Weather and Health Plan (AWHP)[11] which aims to prevent avoidable harms to health during adverse weather events, including during periods of increased heat. A key action for health and social care providers recommended in the AWHP is to "establish methods to identify, alert and monitor individuals most vulnerable to heat-related illnesses on your caseload".[10, 11] Evidence suggests that this particular recommendation is not widely implemented[12, 13] and one potential contributing factor is the absence of an evidenced-based approach through which health care professionals can identify individuals most at risk of dying in a heatwave.[13, 14]

The pathway to ill health during periods of heat is complex and involves a combination of exposures, individual level risk factors and adaptive capacity of the individual. At the population level, at-risk groups include older people, the very young and people with pre-existing medical conditions as well as those whose social, housing or economic circumstances put them at greater risk of harm during periods of heat. [3, 9] However, such broadly defined sub-groups along with poorly specified aspects of heat risk do not allow for the highest risk individuals to be identified and targeted for intervention before adverse health effects occur. Population-level epidemiological studies exploring individual level heat risk have generally used routine mortality and emergency hospitalisation data.[15-17] Where wider determinants of health and heat risk have been investigated, these data have generally been linked to restricted registries[15] or use a small number of proxy measures where individuals are assigned a relevant category based solely on their geographic location.[18] While this does provide some evidence on area-level risk factors, such as level of vegetation cover in London[19], there are a number of assumptions made which may miss some of the individual-level context of heat risk.

We have recently explored individual-level clinical risk factors and heat using primary care records in England and highlighted important clinical factors associated with increased risk of death during heatwaves.[20] Whilst patient record systems in primary care are predominantly used for managing clinical care, they also contain other types of data that are relevant to heat risk. In England, vulnerability assessments undertaken as part of routine primary care practice, such as the electronic frailty index[21] and the QRISK prediction algorithm[22] consider a range of factors in addition to clinical aspects to derive risk scores for frailty and cardiovascular risk, respectively. Therefore, this study aims to identify individual-level socio-environmental risk factors for heat-related mortality in England using primary care records, and to use these data to estimate the potential effect modification of a range of wider determinants of health. Results from this study will build upon previous work and provide foundational evidence for the development of methodologies for effectively identifying individuals at risk of heat-related mortality in England, so that targeted interventions can be deployed.

Methods

Study population

We used the Clinical Practice Research Datalink (CPRD) Aurum (ID number 21_000621) to link primary care records, Office for National Statistics (ONS) mortality data and NHS hospitalisation data, at the individual-level. The outcome was defined as all deaths which occurred between May and September, 2016-2020 using ONS date of death. CPRD Aurum has been shown to be representative of the English population.[23]

Primary care data were used to identify individuals with existing records of relevant individual-level socio-environmental determinants of health. Records were valid if they were recorded within two years of death to better reflect the individual status at the time of death.[15, 24] Where there was more than one relevant record within the two-year window, the record closest to the date of death was used. Individual-level factors investigated were selected a priori based on published evidence, plausibility, and data availability as presented in S1 in the supplementary materials. This resulted in a total of nine priority variables, which were subsequently categorised into multiple sub-groups (see Table 1). Published and bespoke clinical code lists were developed and used to create risk factor variables for all individuals in the study sample. Clinical code lists are available in the supplemental materials table S1.

Exposure data

Geographical information on individuals in CPRD is limited to UK government region of their registered primary care practice. Therefore, a daily mean population weighted regional temperature series was generated for the study period. Daily mean temperatures were generated from HadUK-grid daily maximum and minimum temperatures.[25, 26] The gridded mean daily temperature series was then combined with 100m gridded population data using ArcGIS to create regional population weighted temperature series which were then assigned to each individual based on their GP practice region. A lag period of 0-2 (3-days) was also calculated and assigned to each individual to estimate delayed and cumulative effects of exposure over the three days. Heat effects are largely immediate so impacts at longer lags were not considered.[27]

Statistical analysis

A time-stratified case-crossover study design was used to assess the association between temperature and mortality. Within this study design, temperature on the day of death (event-day) is compared to non-event days. The main relationship under investigation is the association between temperature and risk of death on days above specified temperature thresholds using conditional logistic regression. Each case serves as its own control and therefore the potential effect of time independent confounding factors such as age or gender are automatically controlled for. Case days were determined as the date of death. Control days were selected following a bidirectional referent selection approach, to be the same day of the week of the same month in which the death occurred, resulting in each case having at least three controls, reducing potential for overlap bias.[28]

First, the association between temperature and mortality was modelled to assess the dose-response relationship between temperature and risk of death. This was carried out using natural cubic spline functions, with internal knots determined using the Akaike Information Criterion (AIC) to define the best model fit. From this initial model, temperature thresholds for analysis were derived. To maximise policy-relevance, thresholds were selected using the approach used by the UKHSA for defining the “Low” impact level of the new impact-based Heat-Health Alert system.[29] That is, the temperature associated with a relative risk (RR) of 1.1, i.e. a 10% elevated risk of death. The “Medium” impact threshold, defined as the temperature associated with a RR of 1.2 was used in sensitivity analysis. The “high” impact threshold was not used in this analysis due to daily temperatures within the study period not reaching the required 40°C temperatures, as defined by UKHSA. Relative thresholds were derived for the national level analysis and for sub-national level analysis, also carried out as sensitivity analysis. The reference temperature for the conditional logistic regression was taken as the minimum mortality temperature (MMT) which is the temperature at which risk of death is lowest. Results were stratified by sub-population categories to assess effect modification.

All results are reported as odds ratio’s (OR) with 95% confidence intervals and p-values. In addition, a relative effect modification (REM) index was calculated as the specific OR of an individual-level factor compared to a reference category to aid interpretation of OR estimates. All analysis was carried out in Stata Statistical Software: release 17.[30]

Sub-national level analysis

The analysis was repeated at sub-national level to assess potential geographical variations in estimated associations. Sub-national level analysis was carried out using the following regional groups: London; The North (combined North East, North West and Yorkshire and Humber); Midlands and East (combined West Midlands, East Midlands and East of England); and the South (South West and South East). Regions were combined to ensure that the frequency of the events (deaths, heat-health alerts) was sufficient to support the analysis and based on: study population frequencies, the study population distribution compared to the national distribution over the study period, number of Heat-Health Alerts issued over the study period, geographic location and climate.

Sensitivity analyses

To assess the robustness and generalisability of the results of the analysis three separate sensitivity analyses were undertaken. First, the analysis was repeated using the "medium impact" threshold to assess any differences in the patterns of ORs. Second, we assessed the potential confounding effect of background air pollutants on a restricted number of variables. Particulate matter (PM_{10}), ozone (O_3) and nitrogen dioxide (NO_2) are all potential confounders of the association between heatwaves and mortality .[31] Due to data limitations, this analysis was restricted to London using daily mean NO_2, PM_{10} and O_3

concentration. Five urban background air quality monitoring sites were selected across London from the London Air Quality Network[32]. Using daily means for each site, a London-wide daily mean background value was derived for each pollutant and assigned to cases in London. As with temperature, a 0 to 2-day lag period was also calculated and assigned to each individual.

Patient and Public Involvement

The Health Protection Research Unit in Environmental Change and Health has developed a public engagement/involvement group called PLANET, which was established in Autumn of 2020. PLANET stands for Public Led and Knowledge Engagement Team. The group has 30 members and approximately 20 attend the regular meetings every 3-4 months to discuss research projects. A session exploring heat risk and use of primary care records to identify those at highest risk was carried out in May 2022 where the aims and objectives were presented to the group for feedback. Results from this and linked studies will be presented to the PLANET group in early 2024 at an annual meeting, with the meaning and implications of the results discussed, and potential future research in this area explored.

Results

430,682 individuals that died over the study period (May to September, 2016-2020) from 1,476 primary care practices were included in the analysis. An overview of all individuals with a suitable record for each variable is provided in Table 1. Details of exposure data (temperature and air pollutant concentrations) are provided in S2 in the supplemental materials.

Table 1 Overview of data used in analysis

Variable	**Observations**	**Proportion**
All persons	430,682	100.00%
Male	*211,651*	49.14%
Female	*219,029*	50.86%
Age	430,682	100.00%
Sub-national regions	430,682	100.00%
The North (NE, NW & Y&H)	113,405	26.33%
Midlands and East (WM, EM, EoE)	102,630	23.83%
London	65,145	15.13%
The South (SW & SE)	149,502	34.71%
Alcohol intake category	99,099	**23.01%**
Non-Drinker	*3,386*	0.79%
Light Drinker	*4,624*	1.07%
Moderate Drinker	*88,000*	20.43%
Heavy Drinker	*3,089*	0.72%
Ethnicity	45,263	**10.51%**
White	42,280	9.82%
Black	*985*	0.23%
Asian	1,317	0.31%
Other ethnicity	681	0.16%

Variable	**Observations**	**Proportion**
Living arrangement	16,874	**3.92%**
Living Alone	*11,609*	2.70%
Cohabiting	*4,950*	1.15%
Homeless	*315*	0.07%
Marital Status	17,229	**4.00%**
Single/Divorced/Widowed	*5,773*	1.34%
Married/Has Partner	*11,456*	2.66%
Body Mass Index category	134,884	**31.32%**
Underweight	*15,602*	3.62%
Normal Weight	*54,742*	12.71%
Overweight	*37,044*	8.60%
Obese 1	*17,146*	3.98%
Obese 2	*6,504*	1.51%
Obese 3	*3,864*	0.90%
Frailty category (eFI)	83,968	**19.50%**
Fit	*4,285*	0.99%
Mildly Frail	*15,805*	3.67%
Moderately Frail	*28,290*	6.57%
Severely Frail	*35,604*	8.27%
Index of Multiple Deprivation	430,682	100.00%
1 (least deprived)	*41,757*	9.70%
2	*41,611*	9.66%
3	*43,581*	10.12%
4	*44,693*	10.38%
5	*40,801*	9.47%
6	*43,300*	10.05%

Variable	Observations	Proportion
7	*43,157*	10.02%
8	*40,784*	9.47%
9	*44,597*	10.36%
10 (most deprived)	*46,397*	10.77%
Total number of practices contributing to sample = 1,476 Mean number of patients per practice = 291.79 Maximum number of patients per practice = 2,400 Minimum number of patients per practice = 6		
NW = Northwest; NE = Northeast; Y&H = Yorkshire and the Humber; WM = West Midlands; EM = East Midlands; EoE = East of England; Lon = London; SE = Southeast; SW = Southwest		

Temperature thresholds

Figure 1 illustrates the temperature-mortality relationship derived using the full data series (i.e. all individuals within the study population) and the temperature thresholds derived. The policy relevant thresholds, when rounded to the nearest 0.5°C equate to 17°C (the MMT), 22°C and 24°C. Using these thresholds to identify cases resulted in 13,970 using the "Low" impact threshold and 10,187 using the "Medium" impact threshold. Temperature thresholds used for sub-national level sensitivity analysis are reported in table S3 in the supplementary materials.

Age, sex, ethnicity

Heat mortality risk was modified by age, sex and ethnicity when comparing the odds of death at the MMT (17°C) and “low impact” temperature (22°C). Risk of death during heat episodes increased with age, regardless of size of age groups, however 95% CIs overlap across age groups with the OR estimates for over 65-years groups relatively consistent. Females have somewhat higher risk than males within the study population, but this was not statistically significant.

Of the 10.5% of individuals for whom ethnicity was recorded, those of black or asian ethnicity had substantially higher risk than those who are white, with a REM index of 1.27 for those of black ethnicity and of 1.10 for those with asian ethnicity (white ethnicity as the reference group) (Figure 2 and Table S4 in supplementary materials).

Marital status and living arrangements

Risk of death during heat episodes for those who are single, divorced or widowed are slightly higher than those for individuals who are married or who have a partner, but overlapping CIs indicate that there is likely to be little difference in the risk between these groups. There is some evidence that those who are living alone have increased risk of death on hot days compared to those who are cohabiting. Unfortunately, the numbers of individuals categorised as homeless was too small to provide meaningful OR estimates.

Electronic frailty index (eFI)

Overall, no trend in risk by eFI category was observed. The similarity of the OR point estimates and overlapping CIs for each category of the eFI indicates no difference in the risk profile across eFI categories (fit, mildly frail, moderately frail and severely frail), see Figure 2.

Alcohol intake

Those classed as heavy drinkers had the highest risk of death during heat episodes compared to the other classes of alcohol intake, with moderate drinkers also having increased risk compared to light and non-drinkers, who appear to have no evidence of an association between mortality and high temperatures.

Body Mass Index

Individuals who are categorised as underweight and overweight (including obese 1-3) have increased risk of heat-related death compared to those who are considered normal weight. There is a clear J-shaped trend in ORs by BMI category, except for the highest BMI category, obese 3, where the OR estimate reduces considerably (see Figure 2). However, when all obesity sub-categories are combined into a single group, the OR of the combined obese group is raised, and this finding is statistically significant.

Deprivation

Patterns observed in risk of heat-related death by deprivation group highlight that the highest risk in heat-related deaths occurs for those within the two most deprived groups, while those in the least deprived groups have the lowest risk, with the difference between the ORs for the most deprived groups (IMD 9 and 10) with the least deprived groups (IMD 1 and 2) statistically significant. While the relationship does not appear to be linear across all groups, with those living in areas with an IMD value of 3 up to 7 having comparable ORs, there does appear to be a general trend in increasing risk with increasing deprivation score.

Sub-national and Sensitivity analysis

In general, the patterns observed for the estimated ORs described above were largely consistent when the analysis was repeated using the medium impact threshold temperatures (Table S4) and at sub-national level (Table S5). However, the patterns observed for deprivation at the sub-national level did not quite match those of the national patterns and there were two specific differences observed for the London analysis which are notable. First, the OR estimates by age group for London are reasonably constant across all groups, unlike the results from the national level analysis and other sub-national areas where OR estimates generally increase with age (see Table S5). Second, the difference in OR by sex is particularly pronounced in London using the "medium" impact threshold. When the model for London was adjusted for daily mean concentrations of background PM_{10}, NO_2 and O_3, OR estimates and patterns were consistent with the unadjusted estimates as can be seen in figure S1 in the supplementary materials.

Discussion

This study shows that amongst the study population there are clear patterns across sub-group populations for a range of health determinants. This included increasing risk with age, differences in risk by sex, ethnicity, living arrangement, alcohol intake, BMI and deprivation. However, no real differences in risk were observed within or between frailty sub-groups, illustrating that this routine measure is not a good proxy for heat-risk. We demonstrated that our findings are unlikely to be due to confounding from concurrent exposure to air pollution. We also identified some regional nuances in patterns of risk in London that differ to that of the national picture.

The results of increasing risk for older adults and potential differences in risk by sex align with the population-level epidemiological evidence.[3, 33-43] However, our study identified different patterns in London, where risk was more uniformly distributed across the different age groups. Reasons for this are unclear and likely to be a complex combination of many factors. These may include London's unique population profile compared to other parts of the UK[44] and complex migratory patterns of movement into and out of London across age groups.[45]

Observed temperatures are generally higher in London than in other parts of the England, with the additional heat burden of the urban heat island effect which likely increase exposure further,[46] via indoor overheating risk.[47] This risk may be further compounded by reduced capacity for adaptive behaviours to reduce overheating risk, for example security concerns or through necessity of income, or by other socio-economic factors unique to the capital, such as high cost of living[48] and high rates of household overcrowding.[49] Analysis of inequality in the UK from 2020 suggests that inequality in London is far higher than in other regions of the UK, with over a quarter of Londoners living in poverty and over 15% in the top 10% of earners nationally.[50] Recent evidence suggests that in London more affluent areas also have more access to green space which is also linked with cooler urban environments, potentially reducing risk.[19] The potential complexity of contributing factors to heat risk highlighted here just demonstrates the difficulty of ensuring any interventions that are deployed are both adequately targeted and equitable.

Sex-based differences in the ORs are also noteworthy, with women exhibiting higher ORs than men, a finding that is consistent with previous studies.[51] Results by sex are age adjusted by study design, therefore reasons behind this remain unclear. There is evidence that there may be physiological differences with thermoregulatory responses to exogenous and endogenous heat loads, including core body temperature variation by sex, sweat volume discrepancies and hormonal influences associated with the menstrual cycle[52]. However, there are a number of other factors which may also contribute to this apparent difference in risk, such as social and cultural influences[53] and other socioeconomic factors and comorbidities which are more prevalent in older females that may increase their risk as they age.[54, 55] It is plausible however that all of the above play a role in differences in risk observed here and further demonstrates the complexity of contributing factors to heat risk. More research is required to fully explore this trend and the potential causes for these observed differences by sex. This is particularly important in delivering equitable health and care services and reducing health inequalities.[56]

This is the first study to evidence ethnicity as an important risk factor for death during hot days in England. Black and Asian individuals experienced higher risk of mortality on hot days, contrary to previous UK studies.[3] However this result may be the consequence of circumstances and structural racism experienced by ethnic minority groups which lead to increased health inequalities.[57] In addition, individuals living in the most deprived areas experienced increased risk of death, while those in the least deprived areas displayed the lowest. This trend aligns with evidence globally[18] and strengthens emerging evidence that deprivation may be a significant risk factor during heat periods in England.[19] However the relationship observed was not consistently linear, therefore caution should be taken when considering deprivation alone as a way of characterising heat risk.

The results for ethnicity and deprivation within this study reflect well-documented evidence on the importance of socio-environmental factors for health equity.[58] The domains from which IMD is calculated include income, employment, education, health, crime, barriers to housing and services and living environment.[59] All of these domains affect an individual's under-lying physical and mental health as well as their capacity to adapt either their environments or behaviours when temperatures increase. The ability to adapt to changing conditions is one of the key domains of heat risk. Existing health disparities will also play an important role with clear evidence that those in the lowest IMD groups have significantly poorer health overall, with shorter healthy life-expectancy and higher prevalence of long-term conditions many of which are associated with increased risk during heat events.[60] In addition, there is evidence that poor housing, both current and in the past, are significantly associated with poor health outcomes[61] and that disadvantaged households are less likely to have adaptive approaches to maintain cool indoor temperatures.[62] This is particularly important as our buildings are one of the core mediating factors of the temperatures to which we are exposed. These findings underline the importance of considering social determinants of health in assessing heat-related risks, and further strengthens the case for addressing health inequalities as part of wider climate adaptation strategies and policy. The findings illustrate the role of climate justice at a local level, with those experiencing the highest risk contributing the least in terms of green-house gas emissions: in the UK, the top 1% of earners average 76.6 tons of CO2 equivalent per capita, compared to 5.6 tons of CO2 equivalent per capita for the bottom 50%.[63]

Results for eFI and BMI analysis suggest there may be subtle distinctions between those who are considered clinically vulnerable and those at-risk of death during periods of heat. There is considerable overlap between factors used to calculate eFI and those identified within the literature associated with increased heat risk.[18] Therefore, it is plausible to assume that a high eFI score could be a reasonable proxy for heat risk. Similarly, its plausible that those in the highest BMI category could be prioritised due to their risk profile from a physiological perspective.[64] But neither measure was a reliable proxy in this study. For eFI, the ORs align with the estimates for older age groups, which indicates no change in effect, as eFI assessments are only carried out on those aged 65+. This absence of a trend could stem from the fact that moderately frail individuals may be receiving care to limit their transition into the higher category of frailty, while those classed as severely frail will be in receipt of clinical review and management.[65] In addition, individuals classed as 'Obese

3' may be referred to tier 3 weight management—a clinically-led approach to reducing an individual's weight issues.[66] Thus, the level of care being received by individuals within these groups, may lead indirectly or directly to an overall reduction in risk during a heatwave. The distinction between clinical vulnerability and the risk identified in the current study has important implications for how patients are prioritised by clinicians during periods of heat. These findings align with previous studies[20, 24] which highlighted that individuals at greatest risk are not limited to those with the most severe disease, and that those considered more resilient in general are also at risk during periods of heat.

Limitations of the study

Geographical resolution of the health data meant that precise exposure assignment for each individual within the study was not possible. However, previous studies[1] have demonstrated high correlation between temperature monitoring stations within English regions, and that it is possible to characterise exposure well using a regionally representative temperature series. Inconsistencies in primary care consultation records and the specific terms used when recording details by clinicians increases potential for some relevant records to be missed. However, our systematic approach for identifying relevant records, which included clinical validation, should address this limitation. It's also unlikely that these inconsistencies or missing records are correlated with exposure or outcome, therefore unlikely to be a source of bias. While CPRD is representative of the English population[23] it does not have full coverage of England, and is not geographically representative.[23] Episode analysis of previous heatwaves suggests most heatwave related deaths occur in the south.[6-8, 67-70] Therefore, it is not anticipated that this limitation would affect our results significantly.

These limitations are compounded by data challenges in relation to the completeness of primary care records. For example, only 10.2% of patients in the study had a valid record for ethnicity – the findings of this study suggest that there may be important differences in the risk profile of ethnic minority individuals but it is difficult to draw strong conclusions as the majority of patients lacked these data, despite ethnicity being one of the indicators of the Quality and Outcomes Framework in the UK in which payments are made to encourage recording of such information.[71] Similarly only 3.9% in individuals had a record of living arrangement, with only 0.07% with a record of being homeless. Such small proportions make gaining any insight about this specific group very challenging, despite evidence that this is an at risk group during periods of heat.[72] There are also potential issues with collinearity of some of the wider determinants of health (e.g. ethnicity and IMD) that may be important to consider in the development of any risk stratification approach, however that was not the primary objective of this study. We have also highlighted a potential limitation of relying solely on primary care records, or on clinical risk factors alone, for comprehensive heat risk assessment. The importance of the wider determinants in relation to an individual's heat risk has been well demonstrated; however, as these important data are not routinely collected through electronic records or noted within clinical consultations, it will be difficult to design targeted, evaluated and cost-effective policy interventions in the absence of this foundational evidence. The role of housing (affordable, accessible, healthy) as a core determinant of health and as a mediator of environmental exposures is well defined[73]; therefore, generation and integration of housing data with clinical and non-clinical records should be a priority area for research funders.

Conclusion

This study has demonstrated that information related to socio-environmental determinants of heath as recorded within primary care records can provide important insight about an individual's heat risk. The role of intersecting risks is nuanced and complex, with aspects that are not apparent in the population level data. A key finding of this study is the subtle difference between clinical vulnerability and risk during periods of heat, which has important implications for the identification and management of priority patients during heatwaves. While these results demonstrate the utility of primary care data when assessing non-clinical risk factors, the completeness of records remains a significant challenge. Our results illustrate the complexity of factors that drive heat-related health outcomes and the necessity for evidence-based approaches for assessing risk that accounts for both the clinical and contextual factors influencing an individual's overall risk. This study has also indicated the inequitable burden of impacts experienced by those of non-white ethnicity and those with the least adaptive capacity. Climate change will widen health inequalities, and heat-related harms in particular will be an important outcome of compounding inequalities across exposure, clinical vulnerability and wider socio-environmental disadvantage. The urgency for policy to address these factors is increasing as the climate continues to change and heat events occur more frequently, last longer and are more intense. This research provides foundational evidence for the development of risk management strategies that target those at greatest risk for the deployment of effective interventions. This is an essential step to tackle the increasing trend in heat-related mortality.

Funding Statement

The time of all authors were funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Environmental Change and Health (NIHR200909), a partnership between UK Health Security Agency, the London School of Hygiene & Tropical Medicine, University College London and the Met Office. The views expressed are those of the author(s) and not necessarily those of the NIHR, UK Health Security Agency or the Department of Health and Social Care.

Ethics Approval

Ethics approval for this study was provided by the London School of Hygiene & Tropical Medicine Observational / Interventions Research Ethics Committee; Ethics Committee Reference: 27995

Contributors

Conceptualisation: RT, SK, SH, EOC; Data curation: RT; Formal Analysis: RT; Methodology: RT, SK, SH; Analysis interpterion: RT; SK, SH, EOC; Manuscript drafting and editing: RT, SK, SH, EOC.

Competing Interest

None declared.

References

1. Armstrong, B.G., et al., Association of mortality with high temperatures in a temperate climate: England and Wales. Journal of Epidemiology and Community Health, 2011. 65(4): p. 340-345.
2. Hajat, S., et al., Impact of hot temperatures on death in London: a time series approach. Journal of Epidemiology and Community Health, 2002. 56(5): p. 367.
3. Hajat, S., R.S. Kovats, and K. Lachowycz, Heat-related and cold-related deaths in England and Wales: who is at risk? Occupational and environmental medicine, 2007. 64(2): p. 93-100.
4. Kovats, R.S. and S. Hajat, Heat stress and public health: a critical review. Annu Rev Public Health, 2008. 29: p. 41-55.
5. Leonardi, G.S., et al., Syndromic surveillance use to detect the early effects of heat-waves: an analysis of NHS direct data in England. Soz Praventivmed, 2006. 51(4): p. 194-201.
6. Heatwave mortality monitoring report: 2020, in PHE heatwave mortality monitoring. 2020, Public Health England: Gov.uk.

7. Heat mortality monitoring report: 2021, in Heat mortality monitoring, E.E.a.H.P. Team, Editor. 2021, UK Health Securety Agency gov.uk.
8. Heat mortality monitoring report: 2022, in Heat mortality monitoring, E.E.a.H.P. team, Editor. 2022, UK Health Securety Agency: gov.uk.
9. Adverse Weather and Health Plan: Supporting evidence, U.H.S. Agency, Editor. 2023, HM Government: https://www.gov.uk/government/publications/adverse-weather-and-health-plan.
10. Heat and health in the WHO European Region: updated evidence for effective prevention, W.R.O.f. Europe, Editor. 2021, WHO: Copenhagen.
11. Adverse Weather and Health Plan; Protecting health from weather related harm, U.H.S. Agency, Editor. 2023, HM Government: https://www.gov.uk/government/publications/adverse-weather-and-health-plan.
12. Lorraine Williams, B.E., Stefanie Ettelt, Shakoor Hajat, Tommaso Manacorda, Nicholas Mays, Evaluation of the Heatwave Plan for England, FINAL REPORT. 2019, London School of Hygiene and Tropical Medicine: piru.ac.uk.
13. Brooks, K., et al., Heatwaves, hospitals and health system resilience in England: a qualitative assessment of frontline perspectives from the hot summer of 2019. BMJ Open, 2023. 13(3): p. e068298.
14. Wilson, L., D. Black, and C. Veitch, Heatwaves and the Elderly: the Role of the GP in Reducing Morbidity. Australian Family Physician, 2011. 40(8): p. 637-640.
15. Stafoggia, M., et al., Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. Epidemiology, 2006. 17(3): p. 315-23.

16. Rocklöv, J., et al., Susceptibility to mortality related to temperature and heat and cold wave duration in the population of Stockholm County, Sweden. Global health action, 2014. 7: p. 22737-22737.
17. Schifano, P., et al., Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome. Environmental health : a global access science source, 2009. 8: p. 50-50.
18. Son, J.-Y., J.C. Liu, and M.L. Bell, Temperature-related mortality: a systematic review and investigation of effect modifiers. Environmental Research Letters, 2019. 14(7): p. 073004.
19. Murage, P., et al., What individual and neighbourhood-level factors increase the risk of heat-related mortality? A case-crossover study of over 185,000 deaths in London using high-resolution climate datasets. Environment International, 2020. 134: p. 105292.
20. Thompson, R., et al., Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records. Submitted to BMJ Public Health 2024.
21. Clegg, A., et al., Development and validation of an electronic frailty index using routine primary care electronic health record data. Age and Ageing, 2016. 45(3): p. 353-360.
22. Hippisley-Cox, J., C. Coupland, and P. Brindle, Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ, 2017. 357: p. j2099.
23. Wolf, A., et al., Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. International Journal of Epidemiology, 2019. 48(6): p. 1740-1740g.

24. Stafoggia, M., et al., Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis. J Epidemiol Community Health, 2008. 62(3): p. 209-15.
25. Met Office; HOLLIS, D.M., M.; KENDON, M.; LEGG, T.; SIMPSON, I., HadUK-Grid gridded and regional average climate observations for the UK, C.f.E.D. Analysis, Editor. 2018.
26. Hollis, D., et al., HadUK-Grid—A new UK dataset of gridded climate observations. Geoscience Data Journal, 2019. 6(2): p. 151-159.
27. Armstrong, B., Models for the Relationship Between Ambient Temperature and Daily Mortality. 2006, Epidemiology. p. 624-631.
28. Austin, H., W.D. Flanders, and K.J. Rothman, Bias arising in case-control studies from selection of controls from overlapping groups. Int J Epidemiol, 1989. 18(3): p. 713-6.
29. Adverse Weather and Health Plan; User guide - New impact-based Weather-Health Alerting System, U.H.S. Agency, Editor. 2023, HM Government https://www.gov.uk/government/publications/adverse-weather-and-health-plan.
30. StataCorp, Stata Statistical Software: Release 17. 2021, StataCorp LLC.: College Station, TX:.
31. Stafoggia, M., et al., Joint effect of heat and air pollution on mortality in 620 cities of 36 countries. Environment International, 2023. 181: p. 108258.
32. London Air, I.C. London, Editor. 2023, Impirial College London https://www.londonair.org.uk/.
33. Almeida, S.P., E. Casimiro, and J. Calheiros, Effects of apparent temperature on daily mortality in Lisbon and Oporto, Portugal. Environmental Health, 2010. 9(1): p. 12.

34. Breitner, S., et al., Short-term effects of air temperature on mortality and effect modification by air pollution in three cities of Bavaria, Germany: A time-series analysis. Science of The Total Environment, 2014. 485-486: p. 49-61.
35. Gasparrini, A., et al., The effect of high temperatures on cause-specific mortality in England and Wales. Occup Environ Med, 2012. 69(1): p. 56-61.
36. Goodman, P.G., D.W. Dockery, and L. Clancy, Cause-specific mortality and the extended effects of particulate pollution and temperature exposure. Environmental health perspectives, 2004. 112(2): p. 179-185.
37. Huynen, M.M., et al., The impact of heat waves and cold spells on mortality rates in the Dutch population. Environmental health perspectives, 2001. 109(5): p. 463-470.
38. Iñiguez, C., et al., Relation between temperature and mortality in thirteen Spanish cities. International journal of environmental research and public health, 2010. 7(8): p. 3196-3210.
39. Gasparrini, A., et al., Projections of temperature-related excess mortality under climate change scenarios. The Lancet Planetary Health, 2017.
40. Oudin Åström, D., et al., High Summer Temperatures and Mortality in Estonia. PloS one, 2016. 11(5): p. e0155045-e0155045.
41. Pattenden, S., et al., Ozone, heat and mortality: acute effects in 15 British conurbations. Occupational and Environmental Medicine, 2010. 67(10): p. 699.
42. Ragettli, M.S., et al., Exploring the association between heat and mortality in Switzerland between 1995 and 2013. Environmental Research, 2017. 158: p. 703-709.

43. Rocklöv, J. and B. Forsberg, The effect of temperature on mortality in Stockholm 1998—2003: A study of lag structures and heatwave effects. Scandinavian Journal of Public Health, 2008. 36(5): p. 516-523.
44. UK population pyramid interactive, O.o.N. Statistics, Editor. 2023: Online.
45. London, T.f., London’s Poverty Profile 2021. 2021, Trust for London: trustforlondon.org.
46. Kolokotroni, M. and R. Giridharan, Urban heat island intensity in London: An investigation of the impact of physical characteristics on changes in outdoor air temperature during summer. Solar Energy, 2008. 82(11): p. 986-998.
47. Kovats, S.a.B., R., Health, communities and the built environment. In: The Third UK Climate Change Risk Assessment Technical Report in The Third UK Climate Change Risk Assessment Technical Report 2021, Committee on Climate Change.
48. Padley, M., A Minimum Income Standard for London 2019. 2020, Centre for Research in Social Policy, Loughborough University.
49. Alice Munro, J.A.a.M.M., Evidence Review: Housing and Health Inequalities in London,. London: Institute of Health Equity.
50. Agrawal, S. and D. Phillips, Catching up or falling behind? Geographical inequalities in the UK and how they have changed in recent years 2020, The Institute for Fiscal Studies ifs.org.
51. van Steen, Y., et al., Sex differences in mortality after heat waves: are elderly women at higher risk? Int Arch Occup Environ Health, 2019. 92(1): p. 37-48.
52. Kaciuba-Uscilko, H. and R. Grucza, Gender differences in thermoregulation. Curr Opin Clin Nutr Metab Care, 2001. 4(6): p. 533-6.

53. Kovats, R.S. and S. Hajat, Heat Stress and Public Health: A Critical Review. Annual Review of Public Health, 2008. 29(1): p. 41-55.
54. Aggarwal, N.R., et al., Sex Differences in Ischemic Heart Disease: Advances, Obstacles, and Next Steps. Circ Cardiovasc Qual Outcomes, 2018. 11(2): p. e004437.
55. Vlassoff, C., Gender differences in determinants and consequences of health and illness. J Health Popul Nutr, 2007. 25(1): p. 47-61.
56. Winchester, N., Women's health outcomes: Is there a gender gap? 2021: House of Lords Library.
57. Gee, G.C. and C.L. Ford, STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions. Du Bois Rev, 2011. 8(1): p. 115-132.
58. Gronlund, C.J., Racial and socioeconomic disparities in heat-related health effects and their mechanisms: a review. Curr Epidemiol Rep, 2014. 1(3): p. 165-173.
59. The English Indices of Deprivation 2019 (IoD2019) C.a.L.G. Ministry of Housing, Editor. 2019, HM Government: gov.uk.
60. Toby Watt, A.R., Laurie Rachet-Jacquet REAL Centre: Quantifying health inequalities in England. 2022, The Health Foundation.
61. Marsh, A., et al., Housing Deprivation and Health: A Longitudinal Analysis. Housing Studies, 2000. 15(3): p. 411-428.
62. Taylor, J., et al., Ten questions concerning residential overheating in Central and Northern Europe. Building and Environment, 2023. 234: p. 110154.
63. Chancel, L., Piketty, T., Saez, E., Zucman, G. et al., WORLD INEQUALITY REPORT 2022. 2022, World Inequality Lab.
64. Speakman, J.R., Obesity and thermoregulation. Handb Clin Neurol, 2018. 156: p. 431-443.

65. Turner, G. and A. Clegg, Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age and Ageing, 2014. 43(6): p. 744-747.
66. Gupta, S. and M. Chen, Medical management of obesity. Clinical Medicine, 2023. 23(4): p. 323-329.
67. PHE Heatwave Mortality Monitoring: Summer 2016, in PHE Heatwave Mortality Monitoring. 2016, Public Health England: gov.uk.
68. PHE Heatwave Mortality Monitoring: Summer 2017, in PHE Heatwave Mortality Monitoring. 2017, Public Health England: gov.uk.
69. PHE heatwave mortality monitoring; Summer 2018, in PHE Heatwave Mortality Monitoring. 2018, Public Health England: gov.uk.
70. PHE Heatwave Mortality Monitoring: Summer 2019. Research and analysis 2019.
71. Alshamsan, R., et al., Effect of a UK pay-for-performance program on ethnic disparities in diabetes outcomes: interrupted time series analysis. Ann Fam Med, 2012. 10(3): p. 228-34.
72. Hajat, S., et al., Ambient Temperature and Emergency Hospital Admissions in People Experiencing Homelessness: London, United Kingdom, 2011–2019. American Journal of Public Health, 2023. 113(9): p. 981-984.
73. Mwoka, M., et al., Housing as a Social Determinant of Health: Evidence from Singapore, the UK, and Kenya: the 3-D Commission. J Urban Health, 2021. 98(Suppl 1): p. 15-30.

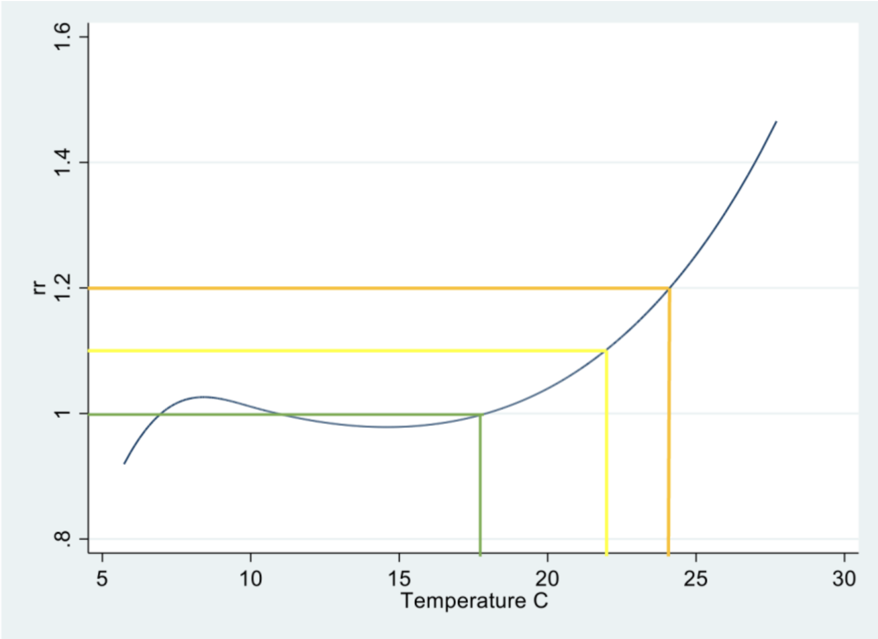


Figure 1 National temperature-mortality relationship plot of relative risk (RR) and mean temperature from which temperature thresholds used in analysis were derived. Green line represents the MMT with a RR=1.0 which equates to about 17°C (rounded to the nearest 0.5°); Yellow line indicate the UKHSA defined Low Impact threshold with a RR of 1.1 which equates to 22°C; Amber lines indicate the UKHSA defined Medium Impact threshold with a RR of 1.2 which equates to 24°C which was used in sensitivity analysis

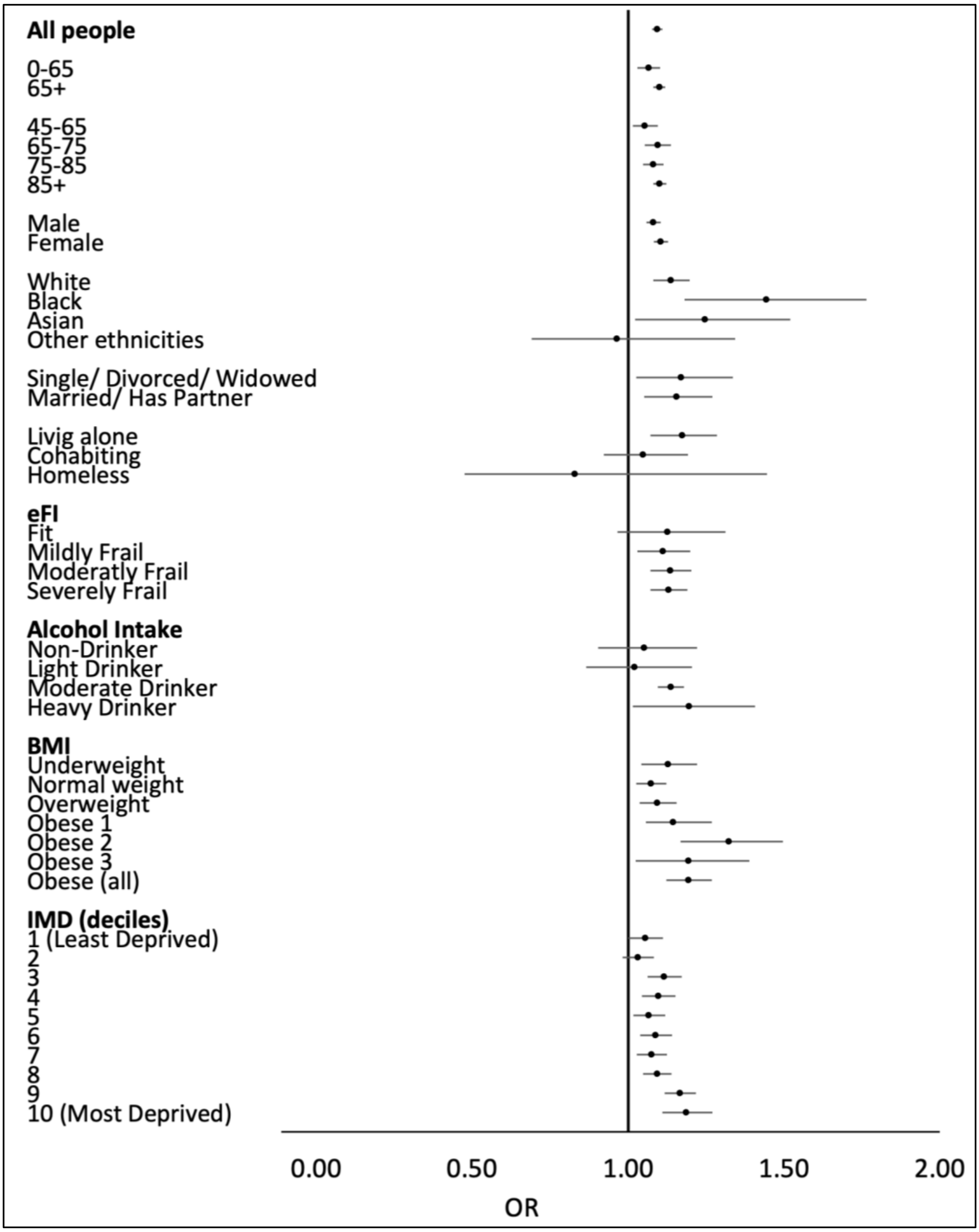


Figure 2 Forest plot showing the estimated OR and 95%CIs by age group, gender, ethnicity, marital status living arrangement, frailty (electronic frailty index), alcohol intake, BMI and deprivation (IMD) using the UKHSA defined HHA low impact thresholds at the national level

.

Chapter 4 - Feasibility of using machine learning and primary care data to predict heat mortality risk in England

4.1 Introduction

The third study undertaken forms the fourth chapter of this PhD and attempts to build upon the results from chapters two and three. Identifying individuals at risk during periods of extreme heat is a core recommendation within Heat-Health Action Plans, yet widespread implementation of this specific recommendation is limited. One contributing factor to this is the lack of evidence-based approaches for accurately identifying which individuals are at risk so that targeted interventions can be deployed. Previous attempts to address this issue have used routinely available population-level data to generate heat vulnerability maps. However, the aggregated nature of the data means they are inaccurate in predicting where the impacts will occur when evaluated against health data and do not provide any intelligence on the individuals within those areas that are most at risk. In addition, potential risk factors may be highly correlated and have complex interactions with each other, prompting the need for new methods to explore this complex area.

The use of machine learning to predict health risks with a high degree of accuracy based on a range of clinical factors has exploded in recent years. But to date, no such attempt has been made to assess the feasibility of such an approach to identifying individuals at high risk during periods of heat using individual-level clinical data that could be used to deploy targeted interventions. Therefore, the aim of this study was to explore the feasibility of using machine learning, and specifically Random Forest, as a tool to predict those at risk of death during periods of heat in England based on risk factors as recorded with primary care records.

The manuscript for this study is currently being prepared for submission to Health Services Research, aiming for submission in October 2024.

4.2 Research paper

Cover page and research paper on subsequent pages.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	2006076	**Title**	MR
First Name(s)	Ross		
Surname/Family Name	Thompson		
Thesis Title	Heatwaves and clinical vulnerability in England; development of a risk stratification tool for use in primary care		
Primary Supervisor	Sari Kovats		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Health Services Research
Please list the paper's authors in the intended authorship order:	Ross Thompson, Sari Kovats, Shakoor Hajat, Emer O'Connell
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Undertook all data formatting, coding and analysis, interpretation of results, manuscript drafting, submission and responding to reviewer comments etc. Co-authors advised throughout the process.

SECTION E

Student Signature	[redacted]
Date	28/09/2024

Supervisor Signature	[redacted]
Date	28/09/2024

Feasibility of using machine learning and primary care data to predict heat mortality risk in England.

Authors:

Ross Thompson[1,2], Sari Kovats[1], Shakoor Hajat[1], Emer O'Connell[1,3]

Affiliations:

[1] NIHR Health Protection Research Unit in Environmental Change and Health, London School of Hygiene and Tropical Medicine, London

[2]Extreme Events and Health Protection Team, UK Health Security Agency, London

[3] Public Health Unit, Greater London Authority

Abstract

Heatwaves pose significant risk to health. Identifying individuals at risk is a core recommendation within Heat-Health Action Plans, yet widespread implementation is limited. Despite the recent boom in the use of machine learning techniques in health research, no attempt has been made to assess the feasibility of using these approaches to identify individuals at risk during periods of heat. This study aimed to explore the feasibility of using Random Forest, to identify those at risk of death during heatwaves in England based on clinical risk factors recorded within primary care records. We used health data from Clinical Practice Research Datalink to identify individual-level risk factors potentially associated with heat risk for 131,305 individuals. Two risk prediction models were developed, Model-1 using all available predictor variables and Model-2 restricted to predictor variables shown to have a strong association with risk of death during heatwaves in previous studies. Models and hyperparameters were determined using the k-folds cross validation approach and variable importance estimated. Model performance was assessed via accuracy, false-negative rates, Kappa coefficient, Area Under the Curve and ROC. Both models performed extremely poorly across all assessment metrics (Model-1 AUC=55.01%; Model-2 =53.99%). However, predictor importance indicated that age was the most important predictor, followed by circulatory system diseases, medications to treat those and non-steroidal anti-inflammatory drugs. Ultimately, we have demonstrated that the use of clinical records alone is unlikely to allow accurate prediction of individuals at risk of death during periods of heat. Potential contributing reasons for this include how heatwave deaths are defined, the complexity of heat risk itself which is a mix of individual susceptibility, environmental, social and behavioural factors. It's recommended that future research explores the linkage of individual level clinical data with a wide range of data sources that represent contextual factors which may influence heat risk.

Author Summary:

Climate change is happening now. We are already seeing more frequent, intense, and longer lasting heatwave episodes. This has resulted in increasing trends in observed heat associated mortality in recent years. Electronic health records (EHR) and the use of machine learning approaches has started to become widespread and is having an impact on patient

care. Yet methods for identifying individuals most at risk of heat related mortality is still lacking. This work explores the utility of EHR in identifying those most at risk of death during periods of heat by using random forest classifying model to predict risk of heat related mortality. We found that the use of clinical data as recorded within primary care records alone is unlikely to provide enough information to reliably identify which individuals are most at risk of death during heatwaves. This is likely due to limitations of the data used, complexity of heat risk and methodological issues around how we define a heat death. This work should be built upon to combine individual-level clinical data with relevant data on other risk factors known to be associated with heat risk.

Introduction

Heat poses a significant risk to health[1-4], with trends in heat associated mortality appearing to increase in recent years.[5 6] In England, deaths associated with periods of high temperature are constantly occurring in people's homes, in the community, in hospitals and care homes.[7 8] The number of excess deaths are predicted to increase further as the climate continues to change and the population ages.[9] There is a large evidence base on which sub-population groups are at increased risk of death during heatwaves, however information on which of those factors contribute the most to the overall risk and attempts to incorporate this into risk stratification approaches is lacking. [10-16]

Identifying individuals at risk is a core recommendation within Heat-Health Action Plans internationally[17], yet widespread implementation is limited.[18] One contributing factor to this is lack of evidence-based approached for accurately identifying which individuals are at risk so that targeted interventions can be deployed.[18 19] Previous attempts to address this issue have used routinely available population level data to generate heat vulnerability maps.[20-22] But the aggregated nature of the data mean they are inaccurate in predicting where the impacts will occur when evaluated against health data and do not provide any intelligence on the individuals within those areas that are most at risk.[20 23 24] In addition, potential risk factors may be highly correlated and have complex interactions with each other, prompting the need for new methods to explore this complex area.[24 25]

A recent study suggested that primary care records are an extremely useful source of data to explore individual level heat risk factors, such as chronic conditions, prescribed medication and certain personal details.[15 16] Primary care professionals may also be well placed to undertake assessment of heat risk for their patient's as more, and more complex care is being provided at home and in the community.[26] Early identification of a patient likely to experience heat stress may result in reduced likelihood of progression of heat illnesses, reduce potential for that individual to require emergency medical attention and even prevent death with early and targeted intervention.[27] Furthermore, primary care professionals will have invaluable knowledge of the population they serve.[26]

The use of machine learning to predict health risk with a high degree of accuracy based on a range of clinical factors has exploded in recent years.[28] Recent examples include applications to risks associated with COVID-19[29], diabetes[30] and cancer[31]. A recent study explored the use of machine learning for investigating the geographic distribution of heat-related mortality using routinely available data with the aim of estimating where the risk of heat associated mortality was highest.[32] But to date, no such attempt has been made to assess the feasibility of such an approach to identifying individuals at high risk during periods of heat using individual-level clinical data that could be used to deploy targeted interventions. Therefore, the aim of this study is to explore the feasibility of using machine learning, and specifically Random Forest[33], as a tool to predict those at risk of death during periods of heat in England based on risk factors as recorded with primary care records.

Methods

Random Forest

Random Forest (RF) is a widely used ensemble machine learning approach used to predict outcomes based on a range of predicting variables. The approach expands on that of decision trees, and develops hundreds of decision trees on random subsamples of the data, with the result for classification models then determined by majority vote of predicted outcome from each decision tree. Each tree within the forest is developed with a random subset of the training sample splitting randomly using predefined number of predictor variables. A separate random subset of the training data is then used for internal evaluation to provide an estimate of model error, termed out-of-bag (OOB) error.[34] Rational for selecting RF as a method include its degree of high predictive accuracy, its robustness to overfitting, allows users to inspect the relative importance of each predictor variable, its user friendliness and relatively low computational cost.[34]

Data

Health data used was obtained from Clinical Practise Research Datalink (CPRD) Aurum (ID number 21_000621). CPRD Aurum has been shown to be representative of the English population.[35] This included primary care records, Office of National Statistics (ONS) mortality data and practise level index of multiple deprivation (IMD) linked at the individual level. The study population was defined as all deaths which occurred between May and September, 2016-2020 using ONS date of death.

Chronic conditions, prescribed medication groups and personal information (age and gender, etc) were selected a priori based on the literature and include a total of 39 variables (not including sub-categories) which are presented in table 1. Data used was adapted from previous studies for use in Random Forest.[15 16] Primary care records were considered valid for inclusion if they were within two years of death. Where there was more than one relevant record within the two-year window, only the record closest to date of death was used. Where a record for the specific condition/medication was not found, it was assumed that the individual did not have the condition or were not prescribed the medication in question.

Due to difficulties of RF dealing with missing values, and the complexity of statistical approaches to imputing synthetic values for sub-groups of categorical variables, the initial study population was reduced based on the following restrictions. First, categorical variables with small proportions of completeness (<15%) of the total sample were removed. This included variables such as living arrangement, marital status, ethnicity, and alcohol intake. Second, individuals that did not contain a valid record for the remaining categorical variables were also removed. Binary variables were simply the presence or not of a condition/prescribed medication, and therefore lack of a value was classed as the condition not present or medication not prescribed.

Geographical information on individuals within the study population is limited to UK government region in which the individual's primary care practice is located. Therefore, a daily mean population weighted regional temperature series that covered the study period and was generated for a previous study[15] was used. Each individual was assigned the population weighted daily mean temperature exposure on their day of death based on their primary care practice location.

Model development

All analysis was carried out in R version 4.2.0.[36] RF approach was used to develop a classification model to predict which individuals within the study population are at high risk of death during heatwave episodes. Definition of heatwave associated death used in this study is based on the temperature on the day of death. To maximize the policy relevance of the research, thresholds used to define heatwave and non-heatwave deaths were aligned with the approach used by the UK Health Security Agency to derive the threshold temperatures at which a yellow alert may be issued.[37] That is the temperature at which the relative risk of all-cause mortality is 1.1.[37] These were determined in a previous study.[15]

The study sample was split into training and testing sub-samples with a 7:3 ratio as is the standard approach, with 70% of the data being used for training, and 30% retained for testing. Due to the unbalanced nature of the sample population (97% non-hot day deaths to 3% hot day deaths) the Synthetic Minority Over-sampling Technique (SMOTE) was employed to generate a synthetic balanced training sample. SMOTE is a widely used approach within machine learning to address highly unbalanced data.[38] The algorithm balances to the data by both under sampling the majority class (non-hot day deaths) and oversampling the minority class (hot day deaths). This new, balanced data was then used to train the RF model.

The RF developed here was trained using the k-fold cross validation technique, in which the training sample is split into k-folds with the model trained and validated k-times. Each sub-model uses all but one of the data folds to train the model and uses the remaining fold to test sub-model performance. Repeated cross-validation approach is a robust method for performance evaluation, can provide important information on model hyperparameters which can be fine-tuned to reduce the risk of model overfitting and finally increases the robustness of the model to data variability and better generalisation when introduced to new data.[39] Ten-fold cross validation was undertaken using the CARET package[40] within R to derive the number of variables to try at each decision node (mtry). The final model was selected based on model accuracy and kappa coefficient. Predictor variable importance was calculated via the CARET package within R[40] and was plotted to indicate the importance of each predictor variable in relation to predicting heatwave or non-heatwave deaths and plotted.

Model performance using the testing data set (30% withheld data) was assessed using a range of statistical tests and calculations, including the confusion matrix output to calculate the overall model accuracy (and 95% CIs), sensitivity and specificity scores to assess the rate of false negatives and positives, Kappa coefficients used to measure how closely the models prediction match the true classes across models and ROC plots and area under the curve (AUC) estimates to assess overall model performance.

Two models were developed following the approach as outlined above. The model-1 was developed using all available predictor variables while model-2 was developed using variables identified as important risk factors associated with increased odds of death in previous studies.[15 16] Evaluation metrics described above for each model were then compered.

Results

The initial data set extracted from CPRD contained records for 430,682 individuals that died over the study period. 131,305 individual records (30%) remained after removal of

Individuals with missing data for included predictors. Table 1 contains details for all potential predictor variables for the included sample, and an indication of the number of patients with the condition of interest or the proportion of patients within sub-categories included in analysis.

Table 1 Overview of the number and proportion of individuals within the study with a positive record for all predictor variables included in the analysis after removal of variables with substantial missing data and observations with missing data. For binary variables each individual was classes as disease present/medication prescribed (value given) or not.

Variable	Frequency	Proportion	Mean	Min	Max
Whole population	131,305	100.0%			
Heat-death	4077	3.1%			
Non-heat death	127228	96.9%			
Age	131,305	100.0%	80	2	110
Sex	131,305	100.0%			
Male	67,866	51.7%			
Female	63,439	48.3%			
Alzheimer's and dementia	18,161	13.8%			
Anxiety	6,328	4.8%			

Variable	Frequency	Proportion	Mean	Min	Max
Arrythmia	20,047	15.3%			
Asthma	6,344	4.8%			
Bipolar disorder	489	0.4%			
Cardiac arrest	1,250	1.0%			
Cardiomyopathy	764	0.6%			
Chronic Kidney disease	16,795	12.8%			
COPD	8,633	6.6%			
Depression	3,269	2.5%			
Emphysema	1,359	1.0%			
Haemorrhage	1,742	1.3%			
Heart failure	14,198	10.8%			
Hyperthyroidism	428	0.3%			
Hypothyroidism	4,463	3.4%			
Severe learning disabilities	800	0.6%			
Liver disease	2,230	1.7%			
Occlusion	3,565	2.7%			
Other CVD	837	0.6%			

Variable	Frequency	Proportion	Mean	Min	Max
Parkinson’s disease	2,550	1.9%			
Psychosis	3,516	2.7%			
Schizophrenia	640	0.5%			
Severe mental illness	1,590	1.2%			
Stroke	6,518	5.0%			
Body mass index category	131,305	100.0%			
Underweight	14,826	11.3%			
Normal weight	53,137	40.5%			
Overweight	36,320	27.7%			
Obese 1	16,858	12.8%			
Obese 2	6,398	4.9%			
Obese 3	3,766	2.9%			
Diastolic blood pressure	131,305	100.0%			
Low DBP	14,852	11.3%			
Normal DBP	81,526	62.1%			
Pre-hypertensive DBP	27,143	20.7%			
DBP Hypertensive 1	6,055	4.6%			

Variable	Frequency	Proportion	Mean	Min	Max
DBP Hypertensive 2	1,729	1.3%			
Systolic blood pressure	131,305	100.0%			
Low SBP	821	0.6%			
Normal SBP	46,210	35.2%			
Pre-hypertensive SBP	55,060	41.9%			
SBP Hypertensive 1	23,780	18.1%			
SBP Hypertensive 2	5,434	4.1%			
Diabetes	36,862	28.1%			
Myocardial infarction	4,638	3.5%			
Ace Inhibitors	39,370	30.0%			
Beta Blockers	33,960	25.9%			
Cardio glycosides	6,907	5.3%			
Diuretics	39,982	30.4%			
NSAIDs	26,905	20.5%			
Vasoconstrictors	279	0.2%			
Anticholinergics	5,035	3.8%			
Index of multiple depravation	131,305	100.0%			

Variable	Frequency	Proportion	Mean	Min	Max
(Least) IMD-1	12,134	9.2%			
IMD-2	12,256	9.3%			
IMD-3	13,355	10.2%			
IMD-4	13,421	10.2%			
IMD-5	12,252	9.3%			
IMD-6	13,119	10.0%			
IMD-7	13,397	10.2%			
IMD-8	12,708	9.7%			
IMD-9	13,780	10.5%			
(Most) IMD-10	14,883	11.3%			

Random forest models

Optimised model parameters following 10-fold cross validation approach for each model are reported in table 2 including the out-of-bag error, indicating the internal assessment of model accuracy. The OOB error across both models are fairly consistent with model 1 (using all available predictors) having a slightly lower OOB error value.

Table 2 Model parameters following k-fold cross-validation approach with 10-folds.

Model parameter	Model 1	Model 2
Number of trees	500 (default)	500 (default)
Number of variable tried at each decision node (*mtry*)	30	18
Predictor variables available (*inlc sub-categories*)	58	35
Out-of-bag Error	20.46%	23.32%
Class error (heatwave death)	0.206	0.234
Class error (non-heatwave death)	0.204	0.232
Note: lower error values in the table indicate better model performance		

The order of relative importance of predictor variables across models are consistent, with age clearly the most important predictor for both, followed by circulatory system diseases and medications used to treat heart failure and high blood pressure. Fig 1 illustrates the top 10 predictors of each model. As the importance ranking reduces, the difference in importance values also reduces resulting in most variables having a comparable importance value below the top 10.

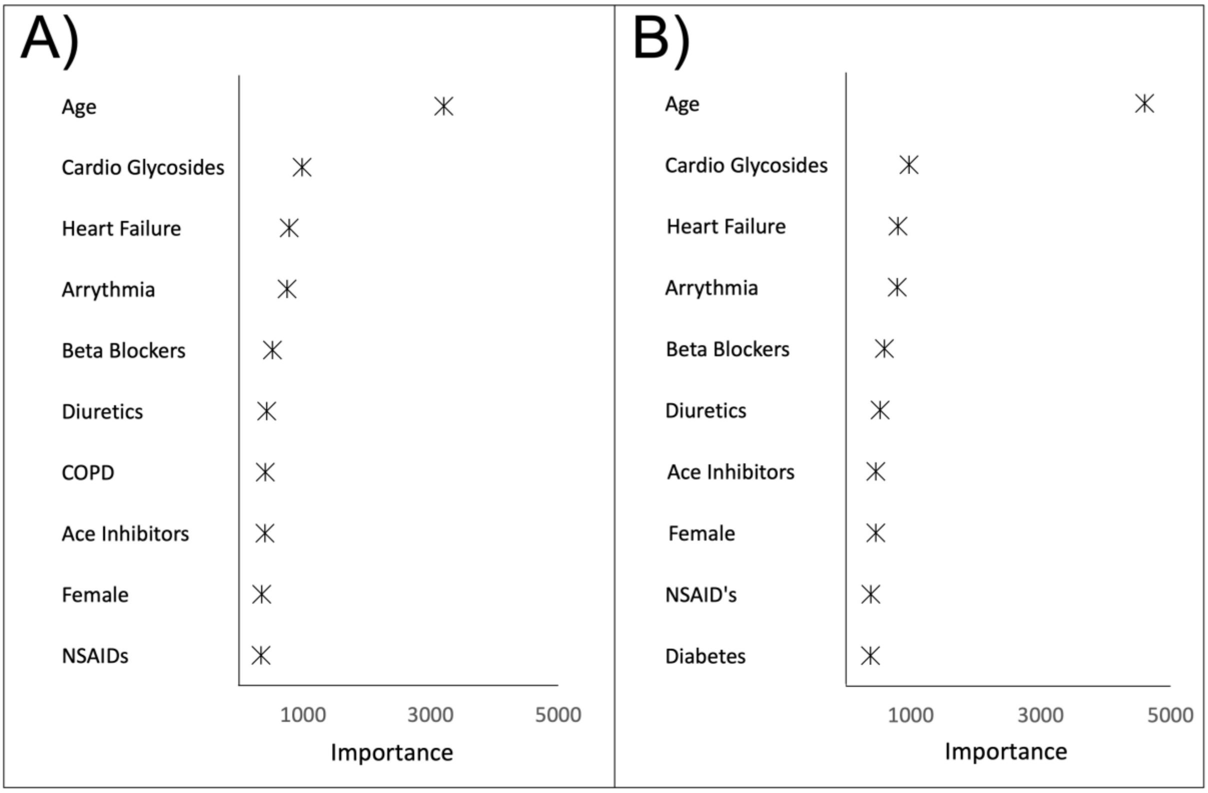
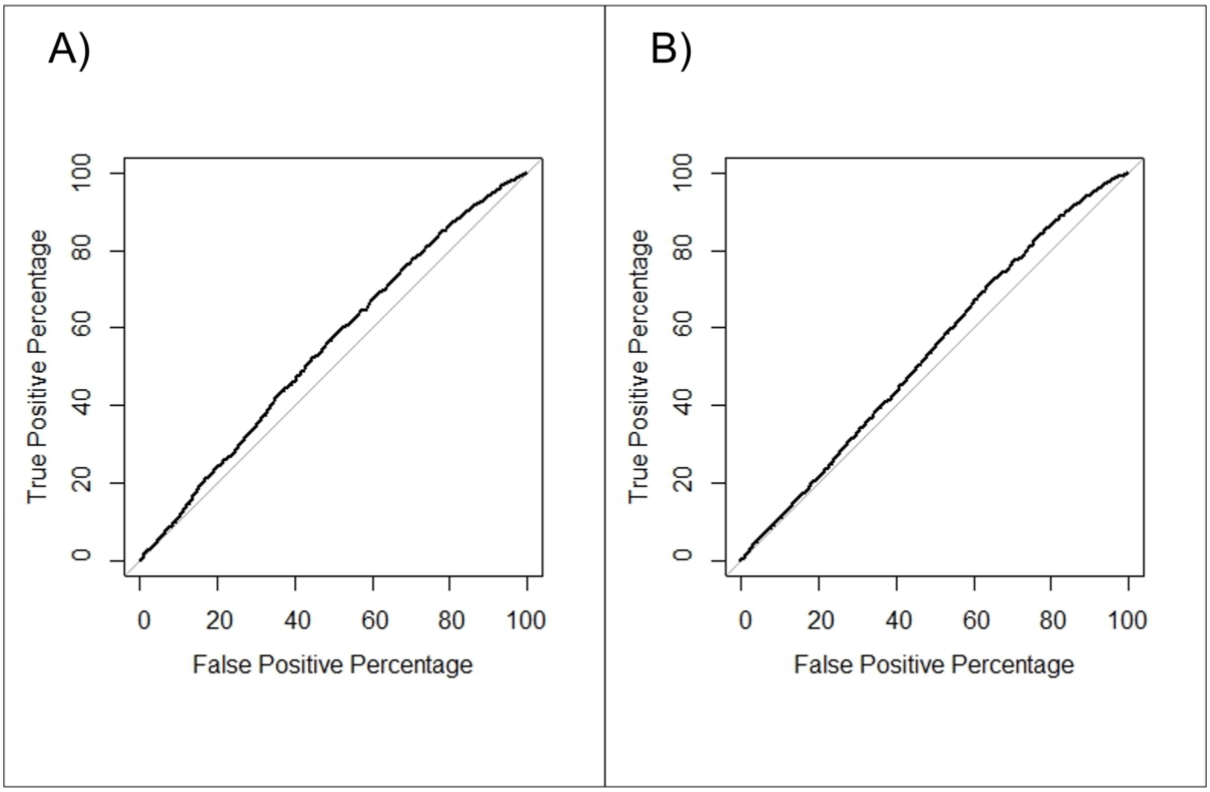


Fig 1 – Variable importance plot of the top-10 variables for both models – panel A for model 1 using all predictor variables and panel B for model 2 using restricted number of variables. Importance value calculated via CARET package in R, values are unitless.

Table 3 shows the model assessment metrics for each of the models developed using the test data. Both models performed poorly in terms of predictive ability to identify the outcome of interest, namely died on a hot day as indicated by AUC estimates, sensitivity values, Kappa coefficients and false negative rate estimates. The poor performance of each model is confirmed via ROC plot as presented in Fig 2.

Table 3 – Model assessment metrics for model 1 (all available predictors) and model 2 (restricted predictors)

Assessment metric	Model 1	Model 2
AUC	55.01%	53.99%
Accuracy (95%CI)	0.75 (0.742 to 0.7506)	0.721 (0.716 to 0.725)
Sensitivity	0.274	0.293
Specificity	0.761	0.734
Kappa coefficient	0.0083	0.0059
False negative rate	72%	71%



Archive

Fig 2 – ROC plot of false positive percentage by true positive percentage for model 1 (panel A) and model 2 (panel B) indicating both models as having poor performance

Discussion

The use of machine learning in health care and clinical medicine to predict the occurrence of rare outcomes has increased dramatically in recent years.[41] Yet this is the first time ML has been used in attempting to identify individuals at increased risk of death on hot days using clinical data as recorded within primary care records.[32 42 43] However, results from this analysis suggest that the use of clinical data alone is not sufficient for accurately predicting which individuals are likely to die during heatwaves. There are perhaps two major limiting factors which contribute to this.

First, heat risk is multifactorial, and a patients individual heat risk will depend not just on their clinical vulnerability, but also their environmental exposures, including the indoor environment where people spend most of their time; their ability to adapt their own behaviours or environment; and social norms and contexts.[4 13 24 44] Therefore, attempting to identify individuals at risk using only one of these strands will lead to poor precision in any prediction model, as has been demonstrated by attempts to develop heat-vulnerability maps based on aggregated social and environmental factors alone and their lack of public health utility.[20 21 23 24]

There are several limitations of the data used in this analysis which unfortunately mean that this first limiting factor cannot be addressed here. These include the geographical resolution available for each individual only being available at the UK government region, meaning that it is not possible to link to other potentially relevant data with location specific detail that could provide further insight into an individual's overall risk. Additional to the inability to accurately assign location specific temperature exposures, there is limited availability of housing type data that would allow indoor temperature exposure estimates at the individual level, which may drive individual level heat risk, given people spend the majority of their time indoors.[45] Linked to these spatial resolution points, deprivation (IMD), which has recently been identified as a potentially important factor for individual level heat risk,[13 16] was based on the individuals primary care practise location, and not the individuals place of residence. This may lead to miss-classification of some individuals, particularly in urban areas where deprivation levels can fluctuate greatly in small geographic areas.

And finally, a range of non-clinical data can be recorded within primary care records such as ethnicity and living arrangement which have been shown to be associated with increased odds of death on hot days.[16] However, the frequency of these types of detail being recorded within the study sample was extremely small, indicating they are not uniformly recorded across primary care.[46] Therefore, where this information was recorded for individuals included within the study sample we were not able to include within the model given the high proportion of missing values, which random forest is not able to deal with efficiently, and the assumption required to impute missing data that values are missing at random is also implausible. These excluded variables could potentially bring important information to the RF approach and increase its predictive performance.

The second limiting factor is potentially related to how a heat-death was defined, here as any person who died on a hot day. Not all deaths that occur on hot days are the result of the increased temperatures, with some deaths likely to occur regardless of the onset of adverse weather. Therefore, without the ability to identify which deaths on a hot day are additional to the deaths that would otherwise have occurred despite the high temperatures, any prediction model is likely to have issues with accuracy in identifying heat-associated deaths. In contrast, where machine learning approaches have been shown to provide robust estimates of individuals at risk, such as diabetes[30] or COVID-19[29], the target health outcome is clear, which is not the case in this present study. This is arguably a methodological challenge that is unlikely to be resolved with the current data available as in the UK deaths are not coded in such a way that would allow easy identification of which deaths were additional during periods of heat.[47]

However, while the results of this analysis are disappointing from a risk prediction perspective, the analysis provides some interesting insights. First, the predictor variables with the highest importance values remained consistent across both models. This suggests, that from a clinical factor perspective at least, there is some confidence that the most important factor of those included is age, followed by the presence of cardiovascular disease or use of medication used to treat heart failure and hypertension. This resonates with previous analysis of episode specific excess mortality in England, where the highest heat-associated death estimates were observed for the older populations with the strongest signal observed for spikes in daily deaths were for those with the underlying cause of death recorded as cardiovascular diseases.[7]

In addition, there is a clear signal that medication use is an important risk factor associated with heat risk. This further demonstrates that this under researched area of heat-health risk requires more attention from the research and medical communities so that appropriate and actionable evidence based clinical guidance can be created to help clinicians in determining how best to manage their patient's medication use during periods of high temperatures.

The use of individual level clinical records in isolation, when attempting to predict heat risk, and the lack of predictive ability identified here resonates with the shortcomings identified when assessing the utility of HVI maps as previously highlighted.[20 23 24 48] With HVI maps, and as demonstrated here, predicting where and who is at risk during a heatwave is a complex issue. The maps fail due to their lack of detailed information about the individual and their personal risk factors, while the approach attempted here failed likely do the fact that wider contextual factors, which are important for overall heat risk, were not included due to data limitations. Therefore, the logical next step is to retain the individual level granularity of personal risk factors contained within primary care records and to combine with available data sets which are either direct risk factors or serve as proxies for relevant contextual factors, such as housing quality or other social or economic factors of relevance. Individual-level behaviours are also likely to play a significant role in an individuals overall risk, yet data on behaviours is likely to be a significant barrier. Nonetheless, future research should explore random forest and other machine learning approaches further, combining different data sources to capture a holistic view of heat risk. However, while precise approaches to identifying individuals most at risk remain limited, effective risk communication with those potentially at increased risk should be prioritised, and the relationships between primary care providers and their patients utilised.

The current approach attempted to utilise machine learning to predict individuals at high risk of death during heatwave periods using primary care records for the first time. Ultimately, it has been demonstrated that the use of clinical records alone is unlikely to be able to predict individuals at risk of death during periods of heat, with any accuracy.

Potential contributing reasons for this lack of predictive ability include how heatwave deaths are defined, and the complexity of heat risk itself. Therefore, it is recommended that future research explore the linkage of individual level clinical data with a wide range of data sources that represent contextual factors which may influence heat risk. This approach would retain the individual level resolution, but also consider wider environmental and social factors. As the climate continues to warm, the response of the health and social care sector will also be required to increase, therefore it is vital that strategies are developed now in which targeted interventions can be deployed to individuals most at risk when periods of adverse weather occur.

Acknowledgments

With special thanks to Dr Stephen O'Neil at London School of Hygiene and Tropical Medicine for sacrificing time and energy to discuss and provide constructive feedback on this work.

References

1. Gasparrini A, Armstrong B, Kovats S, et al. The effect of high temperatures on cause-specific mortality in England and Wales. Occup Environ Med 2012;69(1):56-61. doi: 10.1136/oem.2010.059782 [published Online First: 2011/03/11]
2. Armstrong BG, Chalabi Z, Fenn B, et al. Association of mortality with high temperatures in a temperate climate: England and Wales. Journal of Epidemiology and Community Health 2011;65(4):340-45. doi: 10.1136/jech.2009.093161

3. Ebi KL, Capon A, Berry P, et al. Hot weather and heat extremes: health risks. The Lancet 2021;398(10301):698-708. doi: https://doi.org/10.1016/S0140-6736(21)01208-3
4. Kovats RS, Hajat S. Heat stress and public health: a critical review. Annu Rev Public Health 2008;29:41-55. doi: 10.1146/annurev.publhealth.29.020907.090843 [published Online First: 2007/11/23]
5. Adverse Weather and Health Plan: Supporting evidence. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2023.
6. Wallemacq P, Unisdr, Cred. Economic Losses, Poverty and Disasters 1998-20172018.
7. Thompson R, Landeg O, Kar-Purkayastha I, et al. Heatwave Mortality in Summer 2020 in England: An Observational Study. Int J Environ Res Public Health 2022;19(10) doi: 10.3390/ijerph19106123 [published Online First: 20220518]
8. Kovats S, Johnson, H., and Griffiths, C. . Mortality in southern England during the 2003 heat wave by place of death. Health Statistics Quarterly 2006;29:6-8.
9. Hajat S, Vardoulakis S, Heaviside C, et al. Climate change effects on human health: projections of temperature-related mortality for the UK during the 2020s, 2050s and 2080s. Journal of Epidemiology and Community Health 2014 doi: 10.1136/jech-2013-202449
10. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? Occupational and environmental medicine 2007;64(2):93-100. doi: 10.1136/oem.2006.029017 [published Online First: 2006/09/21]

11. Hajat S, Kovats RS, Atkinson RW, et al. Impact of hot temperatures on death in London: a time series approach. Journal of Epidemiology and Community Health 2002;56(5):367. doi: 10.1136/jech.56.5.367
12. Murage P, Hajat S, Kovats RS. Effect of night-time temperatures on cause and age-specific mortality in London. Environmental Epidemiology 2017;1(2):e005. doi: 10.1097/ee9.0000000000000005
13. Murage P, Kovats S, Sarran C, et al. What individual and neighbourhood-level factors increase the risk of heat-related mortality? A case-crossover study of over 185,000 deaths in London using high-resolution climate datasets. Environment International 2020;134:105292. doi: https://doi.org/10.1016/j.envint.2019.105292
14. Kovats SaB, R. Health, communities and the built environment. In: The Third UK Climate Change Risk Assessment Technical Report The Third UK Climate Change Risk Assessment Technical Report Committee on Climate Change, 2021.
15. Thompson R, Kovats S, Hajat S, et al. Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records. Submitted to BMJ Public Health 2024
16. Thompson R, Kovats RS, Hajat S, et al. Social determinants of heat related mortality in England – a time-stratified case-crossover study using primary care records. Submitted to BMJ Public Health 2024
17. Heat and health in the WHO European Region: updated evidence for effective prevention. In: Europe WROf, ed. Copenhagen: WHO, 2021.

18. Lorraine Williams BE, Stefanie Ettelt, Shakoor Hajat, Tommaso Manacorda, Nicholas Mays. Evaluation of the Heatwave Plan for England, FINAL REPORT. piru.ac.uk: London School of Hygiene and Tropical Medicine, 2019.
19. Brooks K, Landeg O, Kovats S, et al. Heatwaves, hospitals and health system resilience in England: a qualitative assessment of frontline perspectives from the hot summer of 2019. BMJ Open 2023;13(3):e068298. doi: 10.1136/bmjopen-2022-068298
20. Wolf T, McGregor G. The development of a heat wave vulnerability index for London, United Kingdom. Weather and Climate Extremes 2013;1:59-68. doi: https://doi.org/10.1016/j.wace.2013.07.004
21. Taylor J, Wilkinson P, Davies M, et al. Mapping the effects of urban heat island, housing, and age on excess heat-related mortality in London. Urban Climate 2015;14:517-28. doi: https://doi.org/10.1016/j.uclim.2015.08.001
22. Reid CE, O'Neill MS, Gronlund CJ, et al. Mapping community determinants of heat vulnerability. Environ Health Perspect 2009;117(11):1730-6. doi: 10.1289/ehp.0900683 [published Online First: 20090610]
23. Reid CE, Mann JK, Alfasso R, et al. Evaluation of a heat vulnerability index on abnormally hot days: an environmental public health tracking study. Environ Health Perspect 2012;120(5):715-20. doi: 10.1289/ehp.1103766 [published Online First: 20120131]
24. Chuang WC, Gober P. Predicting hospitalization for heat-related illness at the census-tract level: accuracy of a generic heat vulnerability index in Phoenix, Arizona (USA). Environ Health Perspect 2015;123(6):606-12. doi: 10.1289/ehp.1307868 [published Online First: 20150130]

25. Son J-Y, Liu JC, Bell ML. Temperature-related mortality: a systematic review and investigation of effect modifiers. Environmental Research Letters 2019;14(7):073004. doi: 10.1088/1748-9326/ab1cdb
26. Imison C CN, Holder H, Castle-Clarke S, Nimmons D, Appleby J, Thorlby R and Lombardo S Shifting the balance of care: great expectations: Nuffield Trust, 2017.
27. Adverse Weather and Health Plan; Protecting health from weather related harm. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2023.
28. Habehh H, Gohel S. Machine Learning in Healthcare. Curr Genomics 2021;22(4):291-300. doi: 10.2174/1389202922666210705124359
29. Dabbah MA, Reed AB, Booth ATC, et al. Machine learning approach to dynamic risk modeling of mortality in COVID-19: a UK Biobank study. Scientific Reports 2021;11(1):16936. doi: 10.1038/s41598-021-95136-x
30. Ooka T, Johno H, Nakamoto K, et al. Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. BMJ Nutrition, Prevention & Health 2021:bmjnph-2020-000200. doi: 10.1136/bmjnph-2020-000200
31. Hunter B, Hindocha S, Lee RW. The Role of Artificial Intelligence in Early Cancer Diagnosis. Cancers (Basel) 2022;14(6) doi: 10.3390/cancers14061524 [published Online First: 20220316]
32. Kim Y, Kim Y. Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models. Sustainable Cities and Society 2022;79:103677. doi: https://doi.org/10.1016/j.scs.2022.103677

33. Breiman L. Random Forests. Machine Learning 2001;45(1):5-32. doi: 10.1023/A:1010933404324
34. Fratello M, Tagliaferri R. Decision trees and random forests. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics 2018;374
35. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. International Journal of Epidemiology 2019;48(6):1740-40g. doi: 10.1093/ije/dyz034
36. R: A language and environment for statistical computing. [program]. Vienna, Austria, 2021.
37. Adverse Weather and Health Plan; User guide - New impact-based Weather-Health Alerting System. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government 2023.
38. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics 2013;14(1):106. doi: 10.1186/1471-2105-14-106
39. Trevor Hastie RT, Jerome Friedman. The Elements of Statistical Learning; Data Mining, Inference, and Prediction, Second Edition. 2 ed: Springer New York, NY 2009.
40. Kuhn M. Caret package. Journal of statistical software 2008;28(5):1-26.
41. Weissler EH, Naumann T, Andersson T, et al. The role of machine learning in clinical research: transforming the future of evidence generation. Trials 2021;22(1):537. doi: 10.1186/s13063-021-05489-x
42. Boudreault J, Campagna C, Chebana F. Machine and deep learning for modelling heat-health relationships. Science of The Total Environment 2023;892:164660. doi: https://doi.org/10.1016/j.scitotenv.2023.164660

43. Jian L, Patel D, Xiao J, et al. Can we use a machine learning approach to predict the impact of heatwaves on emergency department attendance? Environmental Research Communications 2023;5(4):045005. doi: 10.1088/2515-7620/acca6e
44. Jay O, Capon A, Berry P, et al. Reducing the health effects of hot weather and heat extremes: from personal cooling strategies to green cities. Lancet 2021;398(10301):709-24. doi: 10.1016/s0140-6736(21)01209-5 [published Online First: 2021/08/23]
45. Petrou G, Hutchinson E, Mavrogianni A, et al. Home energy efficiency under net zero: time to monitor UK indoor air. BMJ 2022;377:e069435. doi: 10.1136/bmj-2021-069435
46. Petersen I, Welch CA, Nazareth I, et al. Health indicator recording in UK primary care electronic health records: key implications for handling missing data. Clin Epidemiol 2019;11:157-67. doi: 10.2147/clep.S191437 [published Online First: 20190211]
47. Millares-Martin P. Death certification in England must evolve (Considering current technology). Journal of Forensic and Legal Medicine 2020;69:101882. doi: https://doi.org/10.1016/j.jflm.2019.101882
48. Jonathon Taylor CS, Anna Mavrogianni. URBAN HEAT VULNERABILITY MAPPING TEAM PLACEMENT: Hounslow London Borough Council: Institute for Environmental Design and Engineering, The Bartlett Faculty of the Built Environment, University College London, 2017.

Chapter 5 – Discussion

5.1 – Context of this thesis

The latest State of the Climate report highlights unprecedented acceleration in anthropogenic global warming. [1] Climate adaptation policies and Heat Health Action Plans have been introduced across Europe and around the globe following the 2003 pan-European Heatwave, but heat-associated deaths persist. [2-10] Identifying vulnerable individuals is a key recommendation within heat-health action plans, yet implementation is limited by a lack of evidence-based approaches and the complexity of individual and contextual drivers of heat risk itself. [11-13] While population-level epidemiological studies dominate the literature on heat-related mortality, individual-level risk modification factors remain understudied.[14 15] Italy is currently the only country which has established methodologies for identifying and implementing targeted action towards vulnerable populations.[16]

In England, a significant proportion of heat-associated deaths occur outside healthcare settings and in the community, suggesting a role for primary care professionals to deploy targeted interventions. [17 18] Existing heat vulnerability mapping efforts lack precision due to limitations in the data used to develop them, and therefore are of little utility in the response phase of a heat event. [19-21] Emerging tools using electronic health records offer promise in identifying high-risk individuals, which might enable evidenced-based approaches to be developed so that targeted interventions can be deployed, and ultimately reduce the burden on health during heat events.

Therefore, this research aimed to address two key questions: First, are primary care records a viable source of intelligence on individual-level heat-risk factors? And if so, what type of information contained within primary care records can be used to improve our understanding of individual-level risk. And second, can we use information on individual-level risk factors as recorded within primary care data to predict which individuals are at risk of death during periods of high temperature.

To address these two broad questions and to address the evidence gaps identified, three research objectives were set which were presented in subsequent chapters:

- Chapter 2 - an epidemiological analysis of individual-level risk factors associated with heat risk using electronic health records (EHR) containing primary care data, focusing on clinical risk factors such as pre-existing conditions, prescribed medications and clinical measurements

- Chapter 3 – a follow-up epidemiological analysis of individual-level risk factors associated with heat risk using EHR containing primary care data, focusing on a range of wider determinants of health as recorded within primary care data

- Chapter 4 – determine the feasibility of using primary care records to identify those considered at high risk of death during heatwaves using Random Forest, a machine learning classification model

The final discussion chapter of the thesis will summarise the results of each objective of the PhD, discuss the limitations and assess the strengths and novel contributions of this work to the wider evidence base. In addition, the policy relevance of the findings is explored and potential themes for future research presented.

5.2 – Summary of Key results

5.2.1 – Chapter 2 – "*Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records*"

Chapter two explored part of the research question if primary care records are a viable source of intelligence on individual level heat-risk factors, and what type of information contained within primary care records can be used to improve our understanding of individual-level risk, focusing on individual-level clinical risk factors.

Key results from this study suggest that mortality risk increases during periods of heat for individuals with various chronic conditions, including cardio-respiratory, mental health, cognitive function, diabetes, and Parkinson's disease. Individuals prescribed NSAIDs and medications for high blood pressure and heart failure were also identified as being at an increased risk of death during heatwaves. However, the observation that the risk for the highest categories for diastolic blood pressure groups (hypertension 1 and 2) was lower than might have been expected. The study's consistency across different sensitivity analyses suggests that the results observed within this study are robust and unlikely to be affected by potential confounding of air pollutants.

5.2.1.1 - Cardio-respiratory conditions

Cardio-respiratory conditions that were identified in this study as potential risk factors for heat mortality align with the wider literature. Individuals within the study population with a record of haemorrhage, stroke, heart failure, arrythmia and COPD within two years of death had increased risk. The likely physiological mechanisms for increased risk from circulatory system conditions include increased strain on the heart from moving blood to areas of the body for heat loss, which can be further increased as blood viscosity increases with dehydration, or increased likelihood of burst blood vessels or bleeding on the brain, or lack of oxygen to brain cells. Potentially associated, the likely mechanism for COPD-related risk may include difficulty in supplying enough oxygen to cells due to increased cellular activity associated with thermoregulation.

Cardiovascular, stroke and respiratory diseases are all identified as key indicators within the NHS's Quality and Outcomes Framework (QOF).[22] QOF is a pay-for-performance scheme in which primary care practices are rewarded financially for meeting targets. This has been credited with reducing variation in outcomes between practices, better recording within electronic health records and improved interventions.[23] Therefore, from a practical sense, and given the incentivised nature to identify individuals with these conditions identified within the QOF, increased awareness amongst general practitioners of the risks to these groups of individuals who are already identifiable within primary care systems may enable action.

5.2.1.2 – Hypertension

The first unexpected observation was that hypertensive individuals who had lower than expected odds ratios (hypertension 1 and 2 groups for DBP) and raises questions about physiological vulnerability, the role of clinical interventions already deployed and an individual's overall risk when heatwaves occur. From a physiological perspective, individuals classed as hypertensive may be considered vulnerable given that the heart and circulatory system is already under some strain as outlined above, that when a heat event occurs that individual's body will need to work harder to achieve thermoregulation via increased blood flow to the skin while also coping with potential thickening of blood as moisture is lost due to increased sweating. But this was not observed.

This may of course be a chance result, however when viewed in context of results of other risk factors (see section 5.2.2 for frailty and BMI results) this is a pattern that is repeated and has been observed in previous studies.[15] This suggests that patients who have physiological vulnerability during heat events who are being actively managed via clinical intervention, this intervention may mitigate those individuals risks. This could have implications for how clinicians prioritise patients during periods of heat as it may not always be those who are physiologically the most vulnerable who will have the highest risk.

5.2.1.3 – Mental health conditions

The study highlighted a strong association between mental health conditions, especially depression, and increased odds of death during heatwaves. The underlying mechanisms behind these results remain unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk; or combination of all three. This demonstrates the importance of carers and their role in helping those with some mental health conditions cope when high temperatures arrive, regardless if that care is formal or informal. For those in formal care and in residential settings, adapting the environment may be particularly difficult given that there may be other barriers such as security concerns for both the patient and staff when opening windows when its cooler outside than in, lack of knowledge by carers on what effective actions can be taken and poor quality of care. Therefore, the context in which the individual resides will be an important factor to consider in terms of actions taken to reduce risk when heat events arrive.

NHS data suggests that the number of individuals in contact with mental health services is increasing, with the rate of increase particularly high amongst the younger age groups,[24] while analysis from the British Medical Association suggests that investment in the capacity of mental health services is not keeping up with demand.[25] Projections of patterns of illness in England from the Health Foundation estimate that anxiety, depression and Alzheimer's prevalence will also likely increase significantly by 2040[26]. These statistics and estimates of burden, in combination with increasing frequency, intensity and duration of heat events suggest this specific sub-population group are and will be at increased risk as the climate continues to change, and that effective plans for responding are urgently needed. These results have other potential implications beyond heat-health action planning and could overlap with other health priority areas, such as suicide prevention, given the strong links between depression and suicide, and the evidence of increased risk of suicide deaths and high temperatures.[27]

5.2.1.4 – Prescribed medication

Finally, the results for all drug types investigated highlight the need for further research on the interaction of specific medications, disease, and high temperatures on an individual's overall heat risk. As highlighted, risk is increased across some circulatory system diseases, yet risk is also increased for those who are prescribed medication to control symptoms related to some of these conditions, such as heart failure or high blood pressure. However, also as demonstrated, those identified as being hypertensive (1 and 2) had lower than expected risk when compared to other blood pressure groups. This demonstrates that there are likely to be complex interactions between conditions and drugs used to treat them during periods of heat that may lead to changes in risk. However, this would require more precise data on a range of related factors such as exact drug type, dosage prescribed, severity of disease, which was beyond the scope of this work. This complexity is also one of the leading factors for why there is currently a lack of evidence-based guidance for clinicians to implement individually tailored medicine management during heat events globally.

This first phase of the study explored some of the plausible clinical factors which may increase an individual's risk of death during periods of heat as documented within clinical records. However, heat risk is potentially a multifactorial issue, with risk associated with the individual's susceptibility, social and contextual factors, their environment and behaviours. Therefore, the next step was to investigate to what extent social and contextual factors are recorded within primary care records and to explore which of those factors may modify heat risk at the individual-level.

5.2.2 – Chapter 3 – "*Social determinants of heat-related mortality in England – a time-stratified case-crossover study using primary care records*"

Chapter three built upon results presented in chapter two to establish the utility of primary care records for providing data on individual-level socio-environmental determinants of health, which could be of use when developing evidenced based approaches to identifying individuals at risk during periods of heat. While some of the results from this second study align with the wider literature such as increasing risk with age, some interesting patterns emerged which were not quite expected.

5.2.2.1 – Age

First, we found that risk of death during periods of heat increased with age, except in London where the risk was more evenly distributed across age groups. This discrepancy may be linked to a range of factors such as differing age structure within the capital in addition to complex population movements into and out of London, housing quality, density and general affordability issues that may influence an individual's ability to adapt their behaviours or environment, and ultimately their risk. In fact, London consistently has the highest poverty rate in England when compared to the other regions[28], with the material poverty rate, which considers housing costs, even higher in London amongst pensioners.[29] Overheating of homes is also of particular concern as generally temperatures experienced in London and the southeast of England are higher than in other parts of the country, with temperatures further exasperated by the urban heat island effect. This result demonstrates that the local context can play a significant role in overall risk.

5.2.2.2 – Sex

We also found slight differences in risk by sex, with risk for females being slightly higher risk than males, which was adjusted for age by study design. The reasons for this may be related to some physiological differences between men and woman including hormonal influences on thermoregulation, and other social and behavioural responses to heat.[30] However, these remain unclear. Furthermore, it's unclear if this is of operational significance. For example, UKHSA's 2023 Heat mortality monitoring report suggests there was no significant difference in observed heat-associated deaths by sex, however these estimates are crude and unstandardised.[31] The potential difference in risk identified here requires further research to fully understand what might be influencing any difference by sex and should be a priority area of research to address this potential gender health inequality issue.

5.2.2.3 – Inequalities

One of the key findings was related to ethnicity and deprivation. For the first time in England, we found heat mortality risk was significantly modified by ethnicity group with those classed as white having the lowest ORs, while those who are black and Asian having higher ORs. Linked to this we also found that those with the highest IMD score (i.e. most deprived) had the highest risk, while those with the lowest IMD score (i.e. the least depraved) had the lowest risk. These two findings together are reflective of general equity issues in England that are apparent across health outcomes and highlights the fact the effects of climate change are not equally distributed across the population, but that some within our communities will be disproportionality impacted.

A complex mixture of factors influence our health, these include accessibility and quality of health care, individual behaviours and a range of wider determinants such as quality of housing, income and education attainment, all of which are inter-related. These structural drivers of inequality have been shown to be disproportionately associated with poorer health outcomes for some sub-groups within the population, including ethnic minority groups .[32] The impacts observed on individuals because of these inequalities are not proportionate and the increased likelihood of poor health outcomes are avoidable. However, without addressing them now, these differences in risk are likely to increase as the population grows, ages and the equality gap continues to increase into the future. This highlights why interventions need to be developed through an equity lens so as to ensure that any interventions deployed are both addressing the risks associated with high temperature, and not compounding inequitable outcomes. In addition, this also demonstrates the need for better data on such factors so that we can better monitor any unintended consequences of interventions and adapt where needed.

5.2.2.4 – Frailty and obesity

Another key finding from this study and linked to patterns observed for diastolic hypertension groups in the first study, was the suggestion that there may be a difference between those who may have the highest physiological vulnerability, and those who experience the highest risk during a period of heat. Unexpected patterns were observed in ORs for categories of eFI and BMI. First, no pattern was observed in relation to frailty category and risk of death on a hot day, which was unexpected. Second, there appears to be a J-shaped trend for BMI categories and their OR estimates, with OR decreasing between underweight and normal weight categories, which increases again and continues to with overweight and obese categories. However Obese 3 category does not follow this general trend, as the OR for obese 3 was considerably lower than obese 2. Generally, the higher categories for both eFI and BMI experience poorer health, and therefore are likely to be receiving clinical management due to their clinical status in order to either reduce likelihood of progression to severe frailty or to reduce BMI.[33 34] It is plausible that some of these patients needs are significant enough to require a high degree of clinical care. It is perhaps this clinical management/intervention or contact with clinicians during a period of heat that either directly or indirectly reduces that individual's risk overall. This could have significant implications for how clinicians and medical staff prioritise patients during periods of heat.

Results from this second study clearly demonstrate the importance of wider determinants of health in driving heat risk at the individual level, the importance of local and contextual population intelligence in understanding how risk might be distributed across a local area, the potential lack of high quality socio-environmental information recorded within the primary care records and finally the importance of patient management during periods of heat, and that those most at risk may not necessarily be those with the most advanced disease, or poor health.

5.2.3 – Chapter 4 – "*Feasibility of using machine learning and primary care data to predict heat mortality risk in England*"

Chapter four set about to explore the feasibility of using primary care data to identify individuals at risk of death on hot days, by employing Random Forest (RF), a machine learning (ML) approach. The use of machine learning in health care and clinical medicine to predict the occurrence of rare outcomes has increased dramatically in recent years. Yet this is the first time ML has been used in attempting to identify individuals at risk of death on hot days using clinical data as recorded within primary care records. However, results from this analysis suggest that the use of clinical data alone is not sufficient for accurately predicting which individuals are likely to die during heatwaves. The results do suggest that from a clinical perspective at least however, the factors included within the model which could be considered most important in terms of increased risk of death on hot days are advanced age, followed by the presence of circulatory system diseases and/or prescription of medications used to treat circulatory system issues, such as heart failure or hypertension.

Two major limiting factors were identified which limited this approach but could be explored further in the future with the aim of improving prediction model performance with improved data availability. The first limiting factor was the geographical resolution of the health data used in this study, which limited our ability to link health data to more area based risk factors which could play a significant role in overall heat risk such as housing information. The second, perhaps more fundamental limitation is how a heat-death is defined, as not all deaths which occur on a hot day are directly associated with the increased temperatures. The inability to identify those additional deaths that are the result of the high temperatures either directly or indirectly means that the background noise within a risk stratification tool, may lead to poor performance, as the models are also not able to differentiate between those who died on a hot day because of the heat strain on their physiological mechanisms, or simply would have died anyway. When previous applications of RF models are considered, the outcome which they are trying to predict are very clear; diabetes, COVID-19 infection or cancer, and so do not suffer from this same issue. One of the benefits of RF is its ability to predict outcomes with a high degree of accuracy when the outcome has a large number of contributing factors, and interactions between factors, which may lead to the outcome of interest. Having searched the literature for studies in which researchers have experienced similar issues, no such studies were found. This may be either due the fact that random forests have not previously been used in such a way, or that where they have, and predictive ability of the resultant models are poor, they may not have been published due to the negative nature of their results – i.e. negative bias.

Overall, the results from this third study suggest that further research is required before any heat-risk stratification tool is developed that addresses the main consideration raised by primary care professionals and the public alike, i.e. any tool would have to have a high degree of accuracy of identifying those at high risk of ill health during heatwaves.

5.3 – Limitations of research

Several limitations were identified within this PhD which have been briefly described within each chapter. These included limitations of the data available and used and some more fundamental methodological challenges which would need to be addressed before an accurate risk stratification tool for heat risk could be developed. The sections below explore some of the most important limitations identified and are split by data and methodological limitations. To reduce duplication, the below expands on some of the key limitations identified and highlights their importance.

5.3.1 – Data limitations

5.3.1.1 – Individual-level geographical location

One of the major limitations flagged within each study was the lack of geographic location information available for patients included in each study and obtained from CPRD. This had impacts on the analysis undertaken including the exposures assigned to each individual (temperature and air pollution concentrations). The use of regional population weighted temperature series was unlikely to impact results significantly as previous studies have demonstrated temperatures within regions are highly correlated and that its possible to characterise temperature exposures well at regional level.[35] While we were not able to assign each individual with air pollutant concentrations to assess potential confounding of heat and air pollutants on health outcomes investigated, however, we were able to include background concentrations of key air pollutants for individuals in London to include in sensitivity analysis, which suggested that the adjusted Ors for most stratified sub-population groups were unaffected, except for CPOD, and therefore we are confident that the estimates obtained are robust. Finally, we did not assess the potential effect of humidity or rainfall in this analysis, as there is limited evidence that humidity plays a significant role in heat associated mortality in England.[36] While outside the scope of this Project, future research should aim to gain finer geographic resolution of individual-level data and attempt to link with wider meteorological and weather related data sets, and other more general data sets which may provide further intelligence on some of the important area-level risk factors that modify overall heat risk at the individual-level.

5.3.1.2 – lack of individual-level data on socio-environmental factors

The use of primary care data on its own for non-clinical and wider social determinants of health that affect overall heat risk, is unlikely to allow accurate risk prediction. Patient record systems used in primary care settings are primarily for the collection of clinical information to help clinicians manage their patients and record all consultations and outcomes of those consultations. However, contextual and social drivers of health can also be recorded within these systems. But as we have demonstrated, these types of non-clinical drivers of health are not routinely recorded. It's well documented that certain types of social determinants are not adequately recorded within clinical data, despite a move to encourage better recording of this information, like ethnicity for example.[37] It is possible however that the use of other electronic health record databases may have more complete data on some of these factors. For example, representativeness studies of OpenSafely, has demonstrated that only 9.4% of patient records from one of the patient management systems which feed into Open Safely did not have a record for ethnicity.[38] However, it's unclear the extent to which other non-clinical factors are recorded and represented within these other systems, and this may still prove to be a challenge. Another potential challenge is quality of data at the individual level for non-clinical factors. For example, while deprivation was included in this analysis, the IMD assigned to each individual was based on primary care practice location, rather than the individual's area of residence. Even if IMD had been assigned based on the individual's post code of address, there is evidence that suggests that area-based deprivation measures like IMD have limited sensitivity and specificity for identifying individuals who are considered deprived.[39]

For example, from the data we obtained, we demonstrated that OR estimates by living arrangement clearly show that those living alone had higher odds of death on hot days than those cohabiting. This aligns with the wider literature that those potentially socially isolated are at increased risk. However, the proportion of individuals included within the study population with a valid record for living arrangement was only 3.92%. Only two of the categories of this variable provided useful ORs (living alone and cohabiting), with the numbers for those with a valid record for homelessness or sleeping rough, being too small to generate sensible OR estimates, despite this population group being known to be at risk.[40] Regarding homelessness specifically, reasons for the low numbers is unclear given that research by Homeless Link suggests that 93% of rough sleepers surveyed said they were registered with a GP.[41]

According to the 2021 UK census, just under 25% of the homeless population in England are in London[42], however, London is overrepresented more generally within the CPRD data, with a higher proportion of deaths occurring in the capital than in the full national figures over the study period. Therefore, the lack of data on homelessness observed here is not likely to be associated with the representativeness of the data. Perhaps then, this information is not adequately recorded by clinicians with consistent coding, or records are not updated to reflect when individuals may transition to sleeping rough. In addition, there are several barriers to the homeless population accessing health care services which may lead to clinicians being unaware of this transition and recording such information, such as location of practices, ability to get an appointment, health service coordination once they have transitioned to homelessness and negative experience with health care amongst others, potentially all contributing to lack of consistent recording of homelessness.[43] Regardless, this demonstrates the potential limitation of relying on primary care data for information on wider determinants of health which can impact an individual's overall risk during heatwaves, and is a further example of the issues around data completeness, quality and availability.

This particular limitation also impacted the development of the Random Forest prediction model in that the wider determinants of health predictor variables which were low in frequency across the study population due to lack of a record meant that the number of potentially relevant factors had to be omitted from the RF model due to missing values. This almost certainly would have reduced the model performance given the evidence that we generated that some of those wider contextual and social factors play in driving individual-level heat risk. Thus, the RF models developed only considered part of the overall sphere of factors that drive heat risk. Identifying and linking such data sets to clinical data may enhance our understanding of the degree to which socio-environmental factors and clinical factors contribute to overall heat risk, and may allow improved risk prediction, however, the additional data requirement would likely mean that primary care system are not appropriate for hosting a risk stratification tool until such time linked data is readily available within patient record management systems.

There is growing effort to integrate different data sets which provide information on some of the wider determinants of health within clinical records, given that they have such a bearing on health outcomes, and as such can become clinically relevant for specific outcomes.[44-46] But until such time that this type of information is routinely recorded by clinicians, researchers will need to identify relevant data sets and link to clinical records where possible.

5.3.1.3 – Post-pandemic trends and temperature extremes

In addition to the above, the study period used in this project included May to September 2020, during the COVID-19 pandemic, and within the analysis we did not have the ability to remove COVID-associated deaths. However, heat-associated mortality report from UKHSA for summer 2020 demonstrated that COVID-19-related deaths were minimal during the summer months.[8] Therefore, as only one year during the pandemic was included within the study, the effect on these results of not removing COVID-19 deaths is likely to be minimal. It was attempted to identify patients with a record of COVID-19 infection prior to death, however the numbers were extremely low and so this was not included as a factor. In addition, we did not attempt to investigate, not adjust for other circulating infectious diseases, as these tend to be negligible in England during the summer months.

Linked to this, the study period also does not include the most recent, impactful, and extreme heat season on record in England, summer 2022. In fact, from 2020 up until 2023, there appears to have been a step change in heat-associated mortality in England, with annual heat-associated deaths exceeding 1,500-2,000 deaths per year, compared to just under 1,000 deaths between 2016 and 2019.[3] Reasons behind this remain unclear but may be associated with the general increases in excess mortality more broadly observed in the

UK in the post-pandemic world, particularly around cardiovascular diseases as the underlying cause of death.[47] Reasons for this general trend are thought to be many, including direct effects of COVID-19 infection, pressures on NHS services, and disruption to the detection and management of chronic conditions as an indirect effect of the pandemic. In addition, there is emerging evidence of a link between cardiovascular complications following acute COVID-19 infection and the underlying mechanisms that may lead to poor health outcomes.[48 49] In addition, there may also be questions about acclimatisation of the population. However, the observations from the UKHSA annual Heat Mortality Reports would suggest that risk is increasing and not decreasing due to acclimatisation. It would be important that any heat risk stratification tool that is developed uses the most up-to-date data to ensure that subtle changes in general mortality, potential indirect drivers of heat risk and any acclimatisation of the population are fully incorporated into the model so that these new trends can be accounted for. Unfortunately, this limitation is down to timing of the project and data availability.

The limitations of the data identified above could be addressed via use of different data sources and linking several data sets at the individual level which could help to build a better overall picture of individual and area-level factors that drive risk during heatwaves. However, there are perhaps more fundamental limitations which need to be addressed in order to gain a holistic view of individual-level heat-risk that are beyond the scope of this PhD but are essential to address if any risk stratification tool can be developed that has a high degree of accuracy that would be required as highlighted by both members of the PLANET group (stands for Public Led and Knowledge Engagement Team, established by the

NIHR funded Health Protection Research Unit in Environmental Change) and primary care community of practice.

5.3.1.4 – Housing data

First and foremost, accurate and reliable housing data is a serious limitation. People spend on average over 90% of their time indoors[50 51], and therefore individual-level exposures will be significantly mediated by the indoor environment. Evidence in England suggests that even now about 20% of dwellings are overheating in the current climate.[52] However, this vital data simply does not exist in any meaningful form. Previous research has attempted to model indoor temperatures[53]; however, these models are limited by the scarcity of empirical data upon which they are based, estimates are not universally available and due to the aggregated nature of the estimated indoor temperatures, mismatch of exposures is likely. For example, in a multiple occupancy building, internal temperatures may vary widely across dwellings in one building. However, only the mean modelled indoor temperature across the whole building are used in these assessments, which could mean that individuals living in dwellings prone to much higher temperatures than the mean indoor temperature may not be assigned the correct temperature exposures, and their overall risk ultimately underestimated, leading to a high degree of false negatives.

There are perhaps proxies that could be used in analysis and risk stratification tool development such as the Energy Performance Certificate (EPC) rating, which is available for each address, where the dwelling has been sold since 2007. However, as of 2023, around

67% of all residential dwellings in England have an EPC, so coverage is not universal.[54] It is also possible that those dwellings without an EPC are owned and occupied by individuals who have been living in those dwellings since before 2007, and potentially are of older age, one of the prominent risk factors for heat. In addition, there are concerns about the accuracy of EPCs, with one study suggesting the error rate of EPCs to be between 36% and 62%.[55] Given that housing and the built environment mediate an individual's exposure to not just high (and low) temperatures, but also air pollutants and other indoor hazards to health such as mould and dust, research funders should make this an absolute priority area, so that reliable data sets can be developed that can be linked to health data at the individual-level so that the full picture of how our buildings impact on our health can be studied, quantified and ultimately develop and prioritise effective interventions deployed to address the most important risks.

5.3.1.5 – Risk profiles by place of death

In this analysis, patients within the study population were not stratified by place of death. Previous research in England has suggested that additional deaths associated with periods of heat occur consistently for deaths happening at home and in care homes, while increased heat-associated deaths occur in hospital settings during the first period of heat of the season, or during truly extreme conditions.[18] It is plausible that the populations within these different settings have different risk profiles. For example, individuals within care homes or in hospitals may have advanced comorbidities compared to those at home and who might be considered to have more resilient health. This could potentially mean that factors that are driving risk of mortality are slightly different by place of death, which would not be accounted for within this analysis. This would potentially fit with our finding that those who might be considered vulnerable based on their physiological profile, might not have the greatest risk during heatwaves, and the findings of an Italian study that risk within hospitals is highest amongst those in general medicine wards.[15] In addition, within some settings, such as nursing homes, there is potential that there could be multiple cohorts at risk for very different reasons. For example, those with dementia or Alzheimer's who are otherwise considered in good health, compared to other in the same setting who may have a number of comorbidities, leading to different pathways to poor health during periods of heat. However, this level of information was not available within CPRD linked data, which provided information on whether or not an individual died in an NHS facility only. Due to the limited intelligence that this particular variable might provide, this was not a priority variable and therefore was not investigated.

5.3.2 – Methodological limitations

Both of the random forest models performed extremely poorly as reported in chapter four, and in addition to the limitations identified above which may have contributed to the poor predictive ability, perhaps the most fundamental methodological limitation identified within this study is how heat deaths were defined, i.e. any death which occurred on a hot day (using the UKHSA low-impact threshold). For the case-crossover studies, this definition is completely appropriate given that this method assessed the relationship between deaths occurring on a hot day compared to a non-hot day. However, the random forest approach is reliant on a clear defining characteristic which the potential predictor variables may contribute.[56] Using all deaths which occurred on a hot day to define the outcome of interest is perhaps too broad for a risk stratification algorithm to discriminate between deaths that would have happened anyway, despite the heat, and those additional deaths driven by the high temperatures. However, how one would be able to identify a death that is the direct result of heat in England is not clear, as this type of information is not recorded on death certificates. Nor would this type of information be available via any other type of registry. In addition, selecting a higher temperature threshold at which to define a heat-associated death would also not have been possible due to reduced number of deaths and statistical power required. Therefore, this appears to be a fundamental challenge that needs further consideration.

Several other factors were not considered within the analysis presented within this study which should be mentioned. First, we did not explore the role of humidity, either absolute or relative. There is limited evidence for humidity actually playing a significant role when in predicting heat associated mortality associations in the UK.[36] There are potentially several reasons for this which have been presented elsewhere,[36] and so its unlikely that this would have impact the results presented here to any meaningful degree. Within this study we did not consider the potential effect of mortality displacement during heat events, where a spike in deaths might be observed but are infect deaths which are brought forward. Over the study period, UKHSA annual heat mortality reports present daily timeseries of heat associated deaths, and there is no suggestion of such mortality displacement within the records,[4,5,6,7,8] and as such we do not believe this would have impacted the results presented here. Other potential issue that were outside the scope of this PhD, but could be explored in any subsequent work in this area include the effect of the previous winter and summer mortality on observed heat-associated mortality in the year if interest, and what effect that might have on risk profiles across years; similarly assessing risk by underlaying cause of death could also be investigated, as the risk profiles for those who die due to cardiovascular diseases due to heat events may be different to those with respiratory disease listed as underlaying cause of death. However, sample size would likely be a serious consideration, and would not have been possible with the same available for the present work.

5.4 – Strengths of the PhD and it's contribution to knowledge

This was the first study to explore individual-level heat risk factors using primary care data in England, with the ultimate aim of attempting to use this information to develop an approach for identifying those most at risk during a heatwave. While the results of the ultimate aim were somewhat disappointing given that the risk prediction model performed so poorly, here it has been demonstrated that the use of clinical data alone is not a viable approach to identifying individuals at risk of death during periods of heat. As machine learning gains prominence in the field of clinical medicine and patient management, this is an important first step to understanding what is and isn't possible. In addition, a number of the findings from this project have uncovered new insights into the nuances of heat risk at the individual-level, providing valuable contributions to the field of climate and health.

5.4.1 – Physiological vulnerability vrs heat risk

First, across both case-crossover studies, we identified subtle nuances between those who may be clinically vulnerable during periods of heat and those potentially most at risk across sub-population groups. These nuances were observed for diastolic blood pressure, BMI and eFI. From a physiological perspective, the patterns observed were unexpected, given that we might expect those in the highest categories for DBP hypertension, obesity and frailty to have the highest heat risk compared to the other, and lower, groups. However, this was not the case, and for each of these sub-populations, there are potential clinical interventions which may be contributing to this. Such as careful medication management of hypertension, referral of individuals classed as obese-3 to clinical interventions and increasing level of clinical management of patients from moderately frail to severely frail. Evidence from Italy aligns with this finding, in that when risk of death during periods of heat within hospitals was stratified by hospital ward type, it was the general medicine ward which had the highest risk, with the intensive care unit having the lowest risk, and likely temperature controlled.[15] It is plausible these counterintuitive findings are related to the fact that those with the most severe disease are also receiving clinical interventions and a high degree of care which either directly or indirectly reduce their overall risk during periods of heat, thus demonstrating the importance of context in relation to heat risks, even when considering clinical risk factors. The potentially indirect nature of lowering risk also demonstrates the impact interventions can have in terms of heat risk at the individual-level which should not be underplayed.

5.4.2 – Context is key to heat mortality risk

Second, this PhD also revealed that there may be slightly different risk profiles in different parts of England, further demonstrating the importance of context in relation to heat risk. For example, increasing risk with age was observed at the national level, and in most sub-national areas, which aligns with the literature. However, when we looked at patterns of ORs by age group in London, the risk was more evenly spread across the age groups. While London may experience higher temperatures than other parts of the country due to its location and the urban heat island effect, this suggests that the local population dynamics and socioeconomic situation play a significant role in determining individual level heat risk, and that patterns which we think may be universal in terms of heat risk, might be mediated by other contextual factors that are either hard or not possible to quantify.

This resonates with the wider literature on heat mortality that suggests contextual factors are vitally important in driving heat risk. For example, the development and evaluation of a heat vulnerability index (HVI) in Phoenix Arizona, suggested that factors which are considered to be important for heat risk at the national level, may not be replicated at the local level.[21] Indeed, a recent analysis comparing observed heat-associated mortality in England between 2003 and 2020 with modelled heat mortality estimates based on the observed temperatures suggested that the observed mortality each year in England may be driven by more than temperature alone, and that the role of contextual factors is significant in driving the observed impacts.[57]

5.4.3 – Inequitable distribution of heat mortality risk in England

We also demonstrated the potentially inequitable distribution of mortality during periods of heat in England. Differences in risk by sex is a common observation across studies[14 15 58-61] , however, the reasons for this difference remain unclear but underlines the importance of considering social drivers of health in assessing heat-related risks. The results for ethnicity and deprivation within this study reflect well-documented evidence on the importance of socio-environmental factors for health equity.[62] Black and Asian individuals experienced a higher risk of mortality when compared to white individuals on hot days contrary to previous UK studies.[58] However this result may be the consequence of circumstances and structural racism experienced by ethnic minority groups leading to increased health inequalities, and thus increased heat risk.[63] The domains from which IMD is calculated include income, employment, education, health, crime, barriers to housing and services and living environment.[64] All of these domains affect an individual's underlying physical and mental health as well as their capacity to adapt either their environments or behaviours when temperatures increase. Existing health disparities amongst communities also play an important role with clear evidence that those is the lowest IMD groups have significantly poorer health overall, with shorter healthy life expectancy and higher prevalence of long-term conditions many of which are associated with increased risk during heat events.[65] These findings illustrate the role of climate justice at a local level, with those experiencing the highest risk contributing the least in terms of greenhouse gas emissions: in the UK, the top 1% of earners average 76.6 tons of CO_2 equivalent per capita, compared to 5.6 tons of CO_2 equivalent per capita for the bottom 50%.[66]

5.4.5 – Prominence of mental heat conditions

A previous systematic literature review suggested that there was limited evidence on the effect of high ambient temperatures and mortality on those with known mental health conditions, with knowledge gaps on the impact of high temperatures on those with common mental health outcomes such as depression and anxiety.[27] Out of the 12 chronic conditions investigated, which showed a significant association with increased risk of death on hot days, depression had the highest OR estimate, at 1.25 (95%CI 1.09 to 1.44; $p<0.001$) and a relative effect modification index (REM) of 1.15. In addition, severe mental health conditions (composite indicator which includes conditions such as schizophrenia, bipolar disorder etc), Psychosis and Alzheimer's and Dementia also displayed a significant association with heat mortality. The results presented here help in addressing some of the gaps previously identified and has implications for both policy and practice.

5.4.6 – Medication and heat mortality risk

We demonstrated that a range of medications prescribed may also increase risk during periods of heat. The majority of prescribed medications investigated here are used to treat heart failure and hypertension patients, however, a significant association was also found for individuals prescribed non-steroidal anti-inflammatory drugs (NSAIDs), which are a very common drug prescribed for a range of reasons. While we were not able to investigate the extent to which the prescribed medication may drive the observed impacts as outlined above, the evidence identified here, across both the case-crossover study and the random forest study, strengthens the case for the need for evidenced-based clinical guidance to allow health care professionals to manage their patient's medication. Future research in this area would need to address questions about the role of the medication, condition for which it is being prescribed to treat and any potential interaction between the two and overal heat risk. For example, it was demonstrated in chapter 2 that individuals with hypertension (DBP) had lower than expected ORs, yet those who were prescribed medication to treat hypertension also had increased ORs. Investigating what is actually driving this risk will be key for any future clinical guidance. This particular topic was raised several times by primary care professionals during ongoing engagement across the life course of this PhD.

5.4.7 – Clinical code lists

Finally, and in addition to the new insights generated from this thesis, this project also saw the generation of a number of validated clinical code lists which are required to extract useful data on pre-existing conditions, prescribed medications and where possible, some wider determinants of health as recorded within primary care data. These lists are freely available (see appendixes 2 and 3) and can be adapted by other researchers depending on their study needs. The development of the clinically validated code lists required a significant amount of effort and time, and therefore, by making these available, including an overview of the methodology behind their development, other researchers can focus more time to method development and analysis, without having to spend a significant proportion of the time developing these types of code lists.

5.5 – Relevance to public health policy and practice

Ultimately, we have demonstrated that using random forest as a method of predicting individuals at risk of heat associated mortality using primary care data is not feasible with the data available within this project. While it was not possible to use individual-level clinical data alone to identify those most at risk of death during a period of heat, there are consistent findings which could be used within primary care and by health care professionals more generally to help identify groups of patients in their care who may be at increased risk during periods of heat. While it may not be as simple as identifying those with specific conditions, the evidence generated in this PhD, in combination with the wider literature could be used to help develop guidance specifically for primary care professionals on how they might engage with at-risk groups of patients, either on a seasonal basis or in response to the issuing of Heat-Health Alerts by UKHSA. Of course, this would require a considerable level of engagement with primary care professionals to ensure that any advice and guidance was both acceptable to them and actionable to ensure sufficient buy-in for this approach to have an impact.

This study provides evidence on a range of individual-level factors associated with increased risk of death during periods of heat. This evidence can be incorporated into UKHSA's Adverse Weather and Health Plan and supporting materials and help the agencies thinking in terms of future research priorities around identifying vulnerable individuals and some of the challenges which may be associated with this type of work. The results from this PhD have policy relevance beyond just England and the UK, but also at the international level.

One of the main findings of this PhD was the subtle nuances between those who may be considered clinically vulnerable from a physiological perspective and those who appear to have the highest risk have clear implications for providers of health and social care. Previous qualitative research has suggested that prioritisation of patients is not carried out because of the perception from healthcare providers that all those within their care are at risk, and therefore prioritisation is a redundant exercise.[12 13 67] However, our results and that of an Italian study[15] clearly demonstrate that it is not always those with the most severe disease that are at the highest risk, and that those otherwise considered more resilient may be at increased risk during periods of heat. Aside from practical implications for clinicians on how they prioritise patients during periods of heat, there is also a capacity-building exercise that is required to ensure that healthcare providers know of the drivers of heat risk and feel empowered to address those risks for their patients. UKHSA advice and guidance for health and care professionals could also be updated to reflect this as a first step, by working with health professionals to co-develop practical, effective and actionable guidance that would lead to improved patient prioritisation during heat events.

We demonstrated that there is an element of climate justice in terms of risk of death during periods of heat in England. This so-called climate gap is not a new finding but has been clearly demonstrated for heat risk in England in this PhD. Achieving more equitable outcomes is a core priority of UKHSA, and therefore this has very clear policy implications, not just on how health care is provided, but also for the frameworks on which government organisations attempt to address these inequalities and protect health, including the

Adverse Weather and Health Plan, and local heat-health action plans, to ensure that any interventions that are developed and deployed, at both national and local level, are equitable in both delivery and impact.

5.6 – Key areas for future research

In addition to some areas already identified which require further research, there are other potential future research needs which have been identified within this PhD. The limitations revealed here when attempting to identify vulnerable individuals mirror that of the limitations identified for other heat vulnerability identification approaches (e.g. HVI maps), in that both use data that only partially represents factors that contribute to overall heat risk.[19] In the case of this PhD, only factors as captured in clinical records were used, while in heat risk mapping products, population-level health data linked to neighbourhood factors which may contribute to heat risk are used, both of which are not sufficient to accurately characterise where or who is at risk with sufficient accuracy. Therefore, to build upon the work of this PhD, future research should be developed which attempts to retain the individual-level granularity of health data, but which is available at much finer geographic resolution and is linked to wider data sets which may better characterise some of the more contextual nuances of neighbourhood and built environment factors and their overall contribution to individual-level heat risk and more accurate environmental exposures.

In the absence of a validated data-driven approach to targeting patients at the highest risk during heat events, there is still a role for primary care professionals to communicate with their patients and provide useful information and actionable advice to reduce the burden on health. However, engagement within primary care in general for research is challenging due to the increasing workload being placed upon them from both an aging population and a shifting models of care in the community[68], and as the delivery of health care shifts to a more care at home model in England, the asks on primary care will continue to increase, meaning less time for health protection and health improvement activities which are not mandated within the NHS contract. Future research should be undertaken with the aim of identifying levers that can be harnessed to incorporate heat-health (and climate and health more generally) responses into the work already being carried out by primary care on a day-to-day basis. By leveraging available mechanisms, and by providing easily cascaded information, and suggesting easily actionable responses, primary care could potentially deliver a significant intervention that could boost the level of awareness amongst the general population and reduce the overall burden on health during periods of heat.

During engagement sessions with primary care clinicians, the role of medication management and heat risk was consistently identified as being a priority topic of concern, with every primary care professional stressing this as a major gap in their being able to provide the best care for their patients. While a tool that would identify those most at risk would generally be welcomed, those who engaged suggested that even without this, they would most likely have an idea of which patients might need additional help, but that the lack of clinical guidance means there is little action. Lack of specific clinical guidance on patient medicine management during periods of heat is not just an issue in England and the UK but is an international issue. One of the contributing factors to this is a lack of pharmacopathological evidence about the impact of heat on the effectiveness of different drug types, and their potential adverse effects on patients during periods of heat. Given that this is a priority area of clinical concern in regard to heat risk internationally, the World Health Organisation should take a leading role in initiating a work programme to establish the state of the evidence and to set out a road map for the development of clinical guidance for managing patients medication during periods of heat. This would certainly align with the recent call to action by the UN Secretary-General in July 2024 which called for an urgent and coordinated effort to enhance international cooperation to address extreme heat risks, including in the key area of caring for the vulnerable.

5.7 – Concluding statements

Within this PhD I set out to explore if primary care records were a viable source of intelligence on individual-level heat-risk factors, and if so, could this type of information improve our understanding of individual-level heat risks. And second, if we could use this information to predict which individuals were at risk of death during periods of high temperature. The first two studies addressed the first question and demonstrated a number of novel insights into individual-level heat risk in England. This included the nuanced differences between clinical vulnerability and those most at risk which has clear implications for how clinicians are managing their patients during periods of hot weather and for UKHSA in terms of guidance for health and social care providers;, the importance of context in driving heat risk at the individual-level which must be considered when assessing risk and vulnerability to high temperatures; clearly demonstrated the inequitable distribution of heat mortality in England which will only increase as the effects of climate change continue without immediate action to address it; highlighted the critical importance of mental health conditions and the need for high temperatures to be incorporated into other health policies by governmental departments, and finally identified key gaps in the evidence that limit the development of clinically relevant guidance on patient medication management during periods of heat that was identified as a key concern by clinicians. The third study ultimately demonstrated that the use of clinical records alone is unlikely to lead to accurate predictions of individuals at risk of death during periods of heat. Potential contributing reasons for this lack of predictive ability include how heatwave deaths are defined, and the complexity of heat risk itself. Future research should explore the linkage of individual-level clinical data with a wide range of data sources that represent area-level contextual factors which may influence overall individual-level heat risk. In the absence of a data-driven approach to identifying individuals most at risk during periods, mechanisms and levers should be explored that could help clinicians and primary care professionals more generally to engage with their patients before and during heat events that would boost patient awareness and potentially change behaviours reducing the overall burden on health during heatwaves. As the climate continues to warm and the population continues to age, the level of response of the health and social care sector will also be required to also increase, therefore it is vital that a range of strategies are developed now in which interventions can be equitably deployed to individuals most at risk when periods of adverse weather do occur.

References

1. Ripple WJ, Wolf C, Gregg JW, et al. The 2023 state of the climate report: Entering uncharted territory. BioScience 2023 doi: 10.1093/biosci/biad080
2. Robine J-M, Cheung SLK, Le Roy S, et al. Death toll exceeded 70,000 in Europe during the summer of 2003. Comptes Rendus Biologies 2008;331(2):171-78. doi: https://doi.org/10.1016/j.crvi.2007.12.001
3. Adverse Weather and Health Plan: Supporting evidence. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2024.
4. PHE Heatwave Mortality Monitoring: Summer 2016. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2016.
5. PHE Heatwave Mortality Monitoring: Summer 2017. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2017.
6. PHE heatwave mortality monitoring; Summer 2018. PHE Heatwave Mortality Monitoring. gov.uk: Public Health England, 2018.
7. PHE Heatwave Mortality Monitoring: Summer 2019: Public Health England; 2019 [
8. Heatwave mortality monitoring report: 2020. PHE heatwave mortality monitoring. Gov.uk: Public Health England, 2020.
9. Heat mortality monitoring report: 2021. In: Team EEaHP, ed. Heat mortality monitoring. gov.uk: UK Health Securety Agency 2021.
10. Heat mortality monitoring report: 2022. In: team EEaHP, ed. Heat mortality monitoring. gov.uk: UK Health Securety Agency, 2022.

11. Adverse Weather and Health Plan; Protecting health from weather related harm. In: Agency UHS, ed. https://www.gov.uk/government/publications/adverse-weather-and-health-plan: HM Government, 2023.
12. Lorraine Williams BE, Stefanie Ettelt, Shakoor Hajat, Tommaso Manacorda, Nicholas Mays. Evaluation of the Heatwave Plan for England, FINAL REPORT. piru.ac.uk: London School of Hygiene and Tropical Medicine, 2019.
13. Brooks K, Landeg O, Kovats S, et al. Heatwaves, hospitals and health system resilience in England: a qualitative assessment of frontline perspectives from the hot summer of 2019. BMJ Open 2023;13(3):e068298. doi: 10.1136/bmjopen-2022-068298
14. Stafoggia M, Forastiere F, Agostini D, et al. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. Epidemiology 2006;17(3):315-23. doi: 10.1097/01.ede.0000208477.36665.34 [published Online First: 2006/03/30]
15. Stafoggia M, Forastiere F, Agostini D, et al. Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis. J Epidemiol Community Health 2008;62(3):209-15. doi: 10.1136/jech.2007.060715 [published Online First: 2008/02/15]
16. Heat and health in the WHO European Region: updated evidence for effective prevention. In: Europe WROf, ed. Copenhagen: WHO, 2021.
17. Syndromic surveillance: weekly summaries for 2020. GOV.UK: UKHSA, 2020.
18. Thompson R, Landeg O, Kar-Purkayastha I, et al. Heatwave Mortality in Summer 2020 in England: An Observational Study. Int J Environ Res Public Health 2022;19(10) doi: 10.3390/ijerph19106123 [published Online First: 20220518]

19. Wolf T, Chuang WC, McGregor G. On the Science-Policy Bridge: Do Spatial Heat Vulnerability Assessment Studies Influence Policy? Int J Environ Res Public Health 2015;12(10):13321-49. doi: 10.3390/ijerph121013321 [published Online First: 20151023]
20. Wolf T, McGregor G, Analitis A. Performance Assessment of a Heat Wave Vulnerability Index for Greater London, United Kingdom. Weather, Climate, and Society 2014;6(1):32-46. doi: https://doi.org/10.1175/WCAS-D-13-00014.1
21. Chuang WC, Gober P. Predicting hospitalization for heat-related illness at the census-tract level: accuracy of a generic heat vulnerability index in Phoenix, Arizona (USA). Environ Health Perspect 2015;123(6):606-12. doi: 10.1289/ehp.1307868 [published Online First: 20150130]
22. Sutcliffe D, Lester H, Hutton J, et al. NICE and the Quality and Outcomes Framework (QOF) 2009-2011. Qual Prim Care 2012;20(1):47-55.
23. Forbes LJ, Marchand C, Doran T, et al. The role of the Quality and Outcomes Framework in the care of long-term conditions: a systematic review. Br J Gen Pract 2017;67(664):e775-e84. doi: 10.3399/bjgp17X693077 [published Online First: 20170925]
24. Mental Health Services Monthly Statistics , Performance January 2024. In: NHS Digital, ed.: National Health Service, 2024.
25. British Medical Association. Mental health pressures in England: BMA; 2024 [Available from: https://www.bma.org.uk/advice-and-support/nhs-delivery-and-workforce/pressures/mental-health-pressures-data-analysis#:~:text=Between%202017%20and%202022%2C%20rates,10%20to%201%20in%204.

26. Watt T, Raymond A, Rachet-Jacquet L, et al. Health in 2040: projected patterns of illness in England: The Health Foundation, 2023.
27. Thompson R, Hornigold R, Page L, et al. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. Public Health 2018;161:171-91. doi: https://doi.org/10.1016/j.puhe.2018.06.008
28. Poverty in London, before and during the Covid-19 pandemic: Legatum Institute, 2022.
29. Trust for London. London's Poverty Profile 2021. trustforlondon.org: Trust for London, 2021.
30. Gifford RM, Todisco T, Stacey M, et al. Risk of heat illness in men and women: A systematic review and meta-analysis. Environmental Research 2019;171:24-35. doi: https://doi.org/10.1016/j.envres.2018.10.020
31. Heat mortality monitoring report: 2023. In: team EEaHP, ed. Heat mortality monitoring. govuk: UK Health Security Agency, 2024.
32. Williams E, Buck D, Babalola G, et al. What are health inequalities? Insight and analysis, The Kings Fund: The Kings Fund, 2022:https://www.kingsfund.org.uk/insight-and-analysis/long-reads/what-are-health-inequalities.
33. Gupta S, Chen M. Medical management of obesity. Clinical Medicine 2023;23(4):323-29. doi: 10.7861/clinmed.2023-0183
34. Turner G, Clegg A. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age and Ageing 2014;43(6):744-47. doi: 10.1093/ageing/afu138

35. Armstrong BG, Chalabi Z, Fenn B, et al. Association of mortality with high temperatures in a temperate climate: England and Wales. Journal of Epidemiology and Community Health 2011;65(4):340-45. doi: 10.1136/jech.2009.093161
36. Baldwin JW, Benmarhnia T, Ebi KL, et al. Humidity's Role in Heat-Related Health Outcomes: A Heated Debate. Environ Health Perspect 2023;131(5):55001. doi: 10.1289/ehp11807 [published Online First: 20230531]
37. Iqbal G, Johnson MRD, Szczepura A, et al. UK ethnicity data collection for healthcare statistics: the South Asian perspective. BMC Public Health 2012;12(1):243. doi: 10.1186/1471-2458-12-243
38. Andrews C, Schultze A, Curtis H, et al. OpenSAFELY: Representativeness of electronic health record platform OpenSAFELY-TPP data compared to the population of England. Wellcome Open Res 2022;7:191. doi: 10.12688/wellcomeopenres.18010.1 [published Online First: 20220718]
39. McCartney G, Hoggett R, Walsh D, et al. How well do area-based deprivation indices identify income- and employment-deprived individuals across Great Britain today? Public Health 2023;217:22-25. doi: 10.1016/j.puhe.2023.01.020 [published Online First: 20230224]
40. Hajat S, Sarran CE, Bezgrebelna M, et al. Ambient Temperature and Emergency Hospital Admissions in People Experiencing Homelessness: London, United Kingdom, 2011–2019. American Journal of Public Health 2023;113(9):981-84. doi: 10.2105/ajph.2023.307351
41. Homeless Link. The unhealthy state of homelessness: Health audit results 2014: Homeless Link, 2014.

42. Office of National Statistics. 2021 Census: Safeguarded Household Microdata Sample (England and Wales). UK Data Service, 2023.

43. McNeill S, O'Donovan D, Hart N. Access to healthcare for people experiencing homelessness in the UK and Ireland: a scoping review. BMC Health Services Research 2022;22(1):910. doi: 10.1186/s12913-022-08265-y

44. Lai Y, Moseley E, Salgueiro F, et al. Integrating Non-clinical Data with EHRs. In: Data MITC, ed. Secondary Analysis of Electronic Health Records. Cham: Springer International Publishing 2016:51-60.

45. Gottlieb L, Tobey R, Cantor J, et al. Integrating Social And Medical Data To Improve Population Health: Opportunities And Barriers. Health Affairs 2016;35(11):2116-23. doi: 10.1377/hlthaff.2016.0723

46. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep 2014;129 Suppl 2(Suppl 2):19-31. doi: 10.1177/00333549141291s206

47. Pearson-Stuttard J, Caul S, McDonald S, et al. Excess mortality in England post COVID-19 pandemic: implications for secondary prevention. The Lancet Regional Health – Europe 2024;36 doi: 10.1016/j.lanepe.2023.100802

48. Parhizgar P, Yazdankhah N, Rzepka AM, et al. Beyond Acute COVID-19: A Review of Long-term Cardiovascular Outcomes. Can J Cardiol 2023;39(6):726-40. doi: 10.1016/j.cjca.2023.01.031 [published Online First: 20230206]

49. Xie Y, Xu E, Bowe B, et al. Long-term cardiovascular outcomes of COVID-19. Nature Medicine 2022 doi: 10.1038/s41591-022-01689-3

50. Klepeis NE, Nelson WC, Ott WR, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. Journal of

Exposure Science & Environmental Epidemiology 2001;11(3):231-52. doi: 10.1038/sj.jea.7500165

51. Schweizer C, Edwards RD, Bayer-Oglesby L, et al. Indoor time–microenvironment–activity patterns in seven regions of Europe. Journal of Exposure Science & Environmental Epidemiology 2007;17(2):170-81. doi: 10.1038/sj.jes.7500490
52. Beizaee A, Lomas KJ, Firth SK. National survey of summertime temperatures and overheating risk in English homes. Building and Environment 2013;65:1-17. doi: https://doi.org/10.1016/j.buildenv.2013.03.011
53. Taylor J, Wilkinson P, Davies M, et al. Mapping the effects of urban heat island, housing, and age on excess heat-related mortality in London. Urban Climate 2015;14:517-28. doi: https://doi.org/10.1016/j.uclim.2015.08.001
54. Office of National Statistics. Energy efficiency of housing in England and Wales: 2023. In: (ONS) OfNS, ed. ONS Website, 2023.
55. Hardy A, Glew D. An analysis of errors in the Energy Performance certificate database. Energy Policy 2019;129:1168-78. doi: https://doi.org/10.1016/j.enpol.2019.03.022
56. Breiman L. Random Forests. Machine Learning 2001;45(1):5-32. doi: 10.1023/A:1010933404324
57. Lo YTE, Mitchell DM, Thompson R, et al. Estimating heat-related mortality in near real time for national heatwave plans. Environ Res Lett 2022;17(2):024017-24017. doi: 10.1088/1748-9326/ac4cf4 [published Online First: 20220203]
58. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? Occupational and environmental medicine 2007;64(2):93-100. doi: 10.1136/oem.2006.029017 [published Online First: 2006/09/21]

59. Oudin Åström D, Åström C, Rekker K, et al. High Summer Temperatures and Mortality in Estonia. PloS one 2016;11(5):e0155045-e45. doi: 10.1371/journal.pone.0155045
60. Ragettli MS, Vicedo-Cabrera AM, Schindler C, et al. Exploring the association between heat and mortality in Switzerland between 1995 and 2013. Environmental Research 2017;158:703-09. doi: https://doi.org/10.1016/j.envres.2017.07.021
61. Schifano P, Cappai G, De Sario M, et al. Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome. Environmental health : a global access science source 2009;8:50-50. doi: 10.1186/1476-069X-8-50
62. Gronlund CJ. Racial and socioeconomic disparities in heat-related health effects and their mechanisms: a review. Curr Epidemiol Rep 2014;1(3):165-73. doi: 10.1007/s40471-014-0014-4
63. Gee GC, Ford CL. STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions. Du Bois Rev 2011;8(1):115-32. doi: 10.1017/s1742058x11000130
64. The English Indices of Deprivation 2019 (IoD2019) In: Ministry of Housing CaLG, ed. gov.uk: HM Government, 2019.
65. Toby Watt AR, Laurie Rachet-Jacquet REAL Centre: Quantifying health inequalities in England: The Health Foundation, 2022.
66. Chancel L, Piketty, T., Saez, E., Zucman, G. et al. WORLD INEQUALITY REPORT 2022: World Inequality Lab., 2022.
67. Roberts T, Seymour V, Brooks K, et al. Stakeholder perspectives on extreme hot and cold weather alerts in England and the proposed move towards an impact-based approach. Environmental Science & Policy 2022;136:467-75. doi: https://doi.org/10.1016/j.envsci.2022.07.012

68. Stephenson S, Tang EYH, Tang E, et al. Barriers and facilitators to primary care research: views of GP trainees and trainers. BJGP Open 2022;6(2) doi: 10.3399/bjgpo.2021.0099 [published Online First: 20220830]

Appendix 1 – Re-review or Son et al studies evidence tables

Table A1. 1 Evidence tables from re-review papers included in Son et al review.

Title & author	Year	Health data used and analysis methods	Exposure	Factors assessed	How effect size is reported	Effect modifier effect sizes reported	Notes
Effects of apparent temperature on daily mortality in Lisbon and Oporto, Portugal; Almeida et al	2010	Timeseries analysis. Daily deaths from April to September 200-2004 obtained from Portugal National Institute of Statistics. Generalized additive models (GAM), with a quasi- Poisson link function, in the warm period (April to September). Controlled for seasonality, day of the week	Daily apparent mean temperature derived from a single monitoring station for each city (Central Porto and Oporto Airport). Daily mean estimates of PM10 and O3 derived from 3 local background stations	Age • All • >65 Cause of death • All causes • Cardiovascular • Respiratory AQ lower effect when adjusting for PM_{10} in Lisbon, and a decrease in the effect with PM_{10} and ozone in Oporto **Confounding** Long terms trends; day of the week; AQ (PM10 & O3)	Percentage increase in daily mortality for a 1C increase in mean apparent temperature during the warm season (April to September)	***Lisbon*** All causes • All ages = 2.1 (1.6 to 2.5) • >65 = 2.7 (2.1 to 3.2) Cardiovascular • All ages = 2.4 (1.7 to 3.1) • >65 = 2.8 (2.1 to 3.6) Respiratory • All ages = 1.7 (0.1 to 3.4) • >65 = 2.3 (0.5 to 4.1) ***Oporto*** All causes • All ages = 1.5 (1.0 to 1.9)	Unsurprisingly, older age is identified as being the main risk factor. Risk appears to be higher in Lisbon than in Oporto for all-cause mortality with risk of CVD deaths higher than respiratory deaths. However, risk of death from respiratory disease was higher in Oporto than in Lisbon. Authors propose that the difference observed is due to the make-up of the cities (Lisbon more urban but little heavy industry,

	• >65 = 1.8 (1.2 to 2.3) Cardiovascular • All ages = 2.1 (1.3 to 2.9) • >65 = 2.2 (1.3 to 3.0) Respiratory • All ages = 2.7 (1.2 to 4.3) • >65 = 3.0 (1.4 to 4.7)	industrial fabric of Oporto and resultant poorer AQ leading to the difference. This is interesting as it suggests that even within the same country risk of death from particular cause of death during a HW can fluctuate, based on both the location and built fabric of the city, and the impact the main industries may have on the population. **Main messages**: 1. Older age has stringer association with mortality 2. Risk profile of populations modified by their environment 3. Poor AQ may increase respiratory death risk

Short-term effects of air temperature on mortality and effect modification by air pollution in three cities of Bavaria, Germany: A time-series analysis Breitner et al	2014	Daily deaths from Bavarian cities between 1990 and 2006 (Munich, Nuremberg & Augsburg) with cause of death and age. Generalized semi- parametric quasi-Poisson models As potential confounders, long-term trend/seasonality (by using day of study), weekday variations, influenza epidemics (weekly doctor's practice index), relative humidity, and barometric pressure. Time trend and same-day relative humidity were forced into the models Analysis carried out at -0-2, and 0-14 day lags	Daily mean air temperature – met variables derived for analysis slightly differently for each city based on number of sights available Daily QA meterics for each city derived in same way as temp to result in a single city mean value	Cause of death • All cause • CVD Age • All ages • <85 • >85 AQ • PM10 • Ozone Each pollutant was banded into three groups, low, medium and high, and % increases in mortality due to all internal causes of death and CVD specifically assessed to determine what effect AQ might have on mortality during heatwave events. **Confounding** Long term trends/seasonality; day of the week; influenza; relative humidity; barometric pressure	Percent increase in mortality — for heat, 99th percentile relative to 90th percentile for each city. Relative risks were calculated but not reported. Munich 90^{th} – 20.3C 99^{th} – 25.2C Nuremburg 90^{th} – 20.0C 99^{th} – 25.2C Augsburg 90^{th} – 19.4C 99^{th} – 24.1C Combined 90^{th} – 20.0C 99^{th} - 24.8C	0-1 lag for heat Munich • All-cause = 8.1 (5.3 to 10.9) • CVD = 9.1 (5.1 to 13.3) Nuremburg • All cause – 15.4 (11.4 to 19.6) • CVD – 15.7 (9.9 to 21.8) Augsburg • All cause – 14.6 (8.8 to 20.6) • CVD – 8.7 (0.8 to 17.1) ***<u>Combined (all cities)</u>*** **All ages** All cause – 11.4 (7.6 15.3) CVD – 10.6 (7.3 to 14.0) Resp – 23.3 (13.6 to 33.8) **<85**	Respiratory death data appears to only be available for all ages and not age specific, hence why it is not included. In addition, the daily counts also appear to be small, hence why its only included in the pooled analysis. Unsurprisingly, age is the main factor associated with increased risk here. Overall, temperature effects on mortality varied slightly across the different levels of air pollution. Associations between high temperatures and mortality were strongest when PM10 and ozone levels were higher **Main messages**: 1. Older age associated with bigger risk in mortality

All cause - 9.0 (4.8 to 13.3) CVD – 7.9 (4.0 to 12.1) **≥85** All causes – 17.3 (11.0 to 24.0) CVD – 14.8 (9.4 to 20.3) **All internal deaths & AQ** ***Low*** PM – 9.7 (5.6 to 13.9) Ozone – 7.4 (3.2 to 11.7) ***Medium*** PM – 8.4 (6.2 to 10.6) Ozone – 6.3 (4.4 to 8.3) **High** PM – 10.5 (8.5 to 12.5) Ozone – 11.5 (8.5 14.6) **CVD deaths & AQ** ***Low*** PM – 7.2 (2.6 to 11.9) Ozone – 8.2 (2.8 to 13.8) ***Medium***	2. Risk of respiratory deaths higher in Bavaria overall than CVD. 3. Risk of mortality higher on high AQ days than compared to moderate or low.

						PM – 6.4 (4.3 to 8.6) Ozone – 5.9 (3.2 to 8.7) ***High*** PM – 9.4 (6.1 to 12.9) Ozone – 9.6 (4.1 to 15.5)	
The effect of high temperatures on cause-specific mortality in England and Wales **Gasparrini et al**	2012	ONS mortality data – data of death age and underlying cause of death Covering years 1993-2006 focusing on June to September. Timeseries analysis	Regional, population weighted daily maximum dry bulb temperature derived by calculating mean value across monitoring statins which contained 75%+ daily values across study period weighted by the proportion of population living near each station.	Age • 0-64 • 65-74 • 75-84 • 85+ Cause of death • 32 specific causes of death (CVD, Resp, and non- cardioresp deaths Analysis is by region in England **Confounding** Day of the week; day of the year; long term trends; humidity	The results are reported as pooled relative risk exp(b^m) or percentage increase (exp(b^m-1)x100 related to a 1C increase above the region-specific heat thresholds, and as numbers and fractions of deaths attributable to days with temperatures exceeding such thresholds.	32 specific causes of death across five age groups – See paper for sull results – % increase in deaths for every 1C increase in temp above regional threshold for each cause of death is assessed in the five age bands.	Overall, risk increases for almost all specific causes of death with age, however the increase in risk is not uniform across all causes – with some observing higher risk within the younger age groups (e.g. arterial fibrillation risk is highest among the 0-64 years group) **Main messages:** 1. Cardiovascular deaths in general show a strong relationship of increasing risk with increasing age 2. The effect of respiratory causes are more

							consistent across the age groups 3. CVD death risk is smaller than Resp, however the values of estimated CVD deaths are higher than Resp deaths 4. Other non-cardiorespiratory deaths show a similar relationship with age as CVD deaths
Cause-Specific Mortality and the Extended Effects of Particulate Pollution and Temperature Exposure **Goodman et al**	2004	Study period was 1 April 1980 to 31 Dec 1996 Daily death counts for Dublin obtained Timeseries, regression analysis assuming a Poisson distribution in a generalised additive model	Black smoke (BS) and daily minimum temperatures BS average of 6 monitoring sites across the city Min temps from Dublin Airport	Cause of deaths • All cause (non-trauma) • CVD • Respiratory • Non-cardio/resp Age • 0-64 • 65-74 • ≥75 **Confounding** Seasonal trends; day of week; influenza activity	Percent change in mortality associated with each increase of 1C in daily minimum temperature	Cause of death • All = 0.4 (0.1 to 0.6) • CVD = 0.0 (-0.4 to 0.4) • Resp = 0.8 (0.1 to 1.5) • Other = 0.5 (0.5 to 0.6) Age (all cause) • 0-64 = -0.1 (-0.7 to 0.5) • 65-74 = 0.7 (0.2 to 1.3) • ≥75 = 0.3 (-0.1 to 0.7)	Study does not explore the effect of the combination of temp and AQ, but rather temp and AQ separately. I have only reported the temp effect sizes here. **Main messages:** 1. Mortality peaks as temps increase 2. All-causes except CVD increase on hot days 3. The effect was only observed in

							the older ages, i.e. 65+
Heat-related and cold-related deaths in England and Wales: who is at risk? **Hajat et al**	2007	All-cause deaths obtained from ONS between 1993 and 2003 for all regions of England and Wales linked to database of care and nursing homes. Postcodes for each death registration was linked to 2001 census data – Proportion of over 65's living alone, proportion of people living in flats/ maisonette/ apartments, proportion of all people of all ages who describe themselves as Asian, and Black; classified as urban or rural; indices of multiple deprivation. AQ (PM10 and Ozone) and lab	Mean daily temperature by region using one monitoring station and CET	Individual level factors assessed region where they lived, gender and age (in 4 age bands) • 0-64 • 65-74 • 75-84 • ≥85 **Confounding** AQ from a single monitoring station within a region (PM10 & O3); influenza activity	Temperature mortality relative risk plots for each class of factor and the relationships described, rather than values provided.	Cause of death • Strongest effect for Respiratory deaths followed by external deaths • Effect on CVD and "other" deaths also increased as temp increased Age & gender • Overall risk increased with age • Effect strongest for Resp deaths in the ≥85 group, and external deaths in the 0-64 group • Effect on female death risk higher than males for all age groups – most pronounced as age increased Place of death • Effect highest in nursing homes,	**Main messages** (relevant to my project): 1. Age groups, gender and place of death may all modify the risk profile during heat episodes 2. Census deprivation measures didn't really affect risk of mortality for heat 3. Little evidence of other census factors modifying the risk profile.

		confirmed flu data included as potential confounding Poisson generalised linear models. 95th percentile assumed heat threshold. Sensitivity analysis by repeating analysis by death which occurred at home.				followed by care homes then those with no care	
The Impact of Heat Waves and Cold Spells on Mortality Rates in the Dutch Population **Huynen et al**	2001	The Netherlands Central Bureau of Statistic daily deaths, with age and specific causes of death Between 1 Jan 1988 to 31 Dec 1997 Methods are a little unclear. Appears to be concerned with both calculating excess mortality and looking at the temperature	Daily average temperature temp taken from single monitoring station in central Netherlands.	Age only • 0-64 • ≥65 **Confounding** Long term trends (population structure, socioeconomic conditions, provision of health care) and season	Excess deaths and… "*Percental effects estimated from regression analysis of the temperature–mortality relationship; different adjusted models were used for the different causes of death*" I think they report relative increase in risk…. But this	Relative increase in risk for lag of 0 for heat above optimum temperature within models Age – for all causes 0-64 = 0.98 ≥65 = 1.51 Cause specific for all ages - Malignant neoplasms = 1.34 - CVD = 1.42 - Resp = 2.43	This study is not clear in what its reporting. I don't think I would include this in any meta-analysis as I don't have confidence in what's being reported and there are no confidence intervals reported either. **Main Messages** 1. Older age is associated with increased risk 2. Risk of resp death higher

		mortality relationship, and exploring the effect of lag within the model			is not clear from the text or the included tables. Its actually quite difficult to make out what they have done		than CVD and malignant neoplasm
Relation between Temperature and Mortality in Thirteen Spanish Cities **Iñiguez et al**	2010	City specific all cause mortality by all ages and 70+, and CVD cause of death Poisson generalised additive model (GAM)	Daily mean temp for each city taken from the nearest airport	Age and CVD and all natural deaths **Confounding** AQ (PM10, black smoke, total suspended particulates); influenza incidence; day of the week; bank holiday Unclear where confounding data (AQ) came from or how derived	Percentage change in mortality for a temperature change of 1C from minimum mortality temperature	% Increase in mortality for every 1C increase in minimum mortality temperature varied across all cities for heat All cause all age - 0.93% to 2.88% All ages CVD – 0.87% to 7.72% All cause 70+ - 0.99% to 6.50%	No confidence intervals reported. **Main messages** 1. Temperature mortality relationships vary by city 2. Risk higher for 70+ group than for all ages, suggesting the risk in all ages is mostly a result of higher risk for higher ages
Air temperature-related human health outcomes: Current impact and estimations of future risks in Central Italy **Morabito et al**	2012	Poisson generalised additive model (GAM) Daily mortality data -all cause and by age Daily hospitalisation data by age	Daily average air temperature derived through model which includes potential effects of latitude on exposures and is calibrated using observational data	Age **Confounders** Year; season; day of the week; fluctuation in summer population	Heat effect: % change in mortality/morbidity events per 1 °C increase in temperature above the identified threshold Reported as cumulative effect over lag period of 30 days	Mortality Inland plain - <65 = 2.08 (-0.40 to 4.57) - 65-74 = 2.10 (0.07 to 4.13) - ≥75 = 6.22 (0.11 to 12.71) Coastal plain - <65 = 1.51 (-3.93 to 6.95)	**Main messages:** 1. Risk of mortality increases with age 2. Estimated % increase in mortality by age varies by climate zone type i.e. coastal plain vrs inland areas

		10 Tuscan cities between 1999 and 2008 City specific climate and altitude assessment – inland plains, coastal plains, inland hill				- 65-74 = 7.76 (1.41 to 14.11) - ≥75 = 15.97 (7.43 to 24.51) Inland hill - <65 = -0.01 (-3.94 to 3.92) - 65-74 = -0.61 (-8.93 to 7.72) - ≥75 = 5.49 (1.39 to 9.58)	
High Summer Temperatures and Mortality in Estonia **Åström et al**	2016	Daily mortality data for Estonia for Jun to Sep between 1997 and 2013 Distributed lag non-linear model	Mean daily maximum temperature for two regions calculated using three monitoring statins per region.	Age, gender place of death (to assign as either coastal or inland **Confounders/controlled variables** Weekday; day of the week; public holidays	Cumulative relative risk over 0-2 days lag with 95% confidence intervals for the 75th and 99th percentiles per region	**All age all cause** Estonia = 1.18 (1.13 to 1.24) Coastal = 1.12 (1.05 to1.21) Inland = 1.28 (1.20 to 1.37) **Male all cause** Estonia = 1.17 (1.09 to 1.24) Coastal = 1.16 (1.05 to 1.27) Inland = 1.24 (1.13 to 1.36) **Female all cause** Estonia = 1.20 (1.13 to 1.28) Coastal = 1.10 (0.99 to 1.21) Inland = 1.33 (1.20 to 1.46) **0-74 group**	**Main messages**: 1. Higher ages lead to higher mortality risk from hot weather 2. In Estonia risk of death during high temperatures is higher in males 3. Risk is modified by location

						Estonia = 1.14 (1.08 to 1.22) Coastal = 1.13 (1.03 to 1.25) Inland = 1.15 (1.04 to 1.26) **75+ group** Estonia = 1.15 (1.10 to 1.21) Coastal = 1.11 (1.01 to 1.23) Inland = 1.45 (1.31 to 1.60)	
Ozone, heat and mortality: acute effects in 15 British conurbations **Pattenden et al**	2010	Individual health records from ONS matched to conurbation by postcode – all cause (excluding external deaths) and cause specific deaths - CVD, resp & other deaths Age - 0-64, 65-74, 75-84 & 85+	Average temperature and O3 – average daily series used to create representative data set. Only urban background sites were used for AQ metrics as this better represents average exposures **Confounders** Season, day of the week and PM10	Age in four bands	Rate ratio adjusted for confounding by Conurbation and mean rate ratio by age Effect of heat and Ozone individually assessed and heat with extra effect of Ozone	Mean rate ratios *heat only (refer to paper for heat and ozone rate ratios, table 3. In addition rate ratios for tmax also available within paper))*: All deaths = 1.071 (1.050 to 1.093) CVD = 1.055 (1.025 to1.087) Resp = 1.139 (1.079 to 1.202) Other = 1.083 (1.061 to 1.106) 0-64 = 1.019 (0.978 to 1.062) 65-74 = 1.067 (1.031 to 1.103) 75-84 = 1.092 (1.063 to 1.121)	Additive effect of ozone suggestive from this analysis. **Main messages:** 1. Mean effects of heat were higher for respiratory mortality, 2. non-significant in those aged under 65 3. rose considerably with age. 4. Overall ozone effects showed no age- related pattern, but ozone effects on hot days and ozone-heat interactions were greatest in those aged <75.

						85+ = 1.124 (1.093 to 1.155)	
Exploring the association between heat and mortality in Switzerland between 1995 and 2013 **Ragetti et al**	2017	Daily all-cause (excluding external deaths) mortality data from eight cities in Switzerland, which were added into one data series for analysis 1995 to 2003 distributed lag non-linear models with lag 0-6 days Case-crossover study design	Tmax Tmin Tmean Average apparent temp Data for each city included in study taken from a "representative" monitoring station	Age and gender Timing of heat events (early and late summer) **Confounding** Study design controls for long term trends and a number of potential confounding factors. Unclear if AQ was controlled for in	Relative risk for each exposure metric when temps reach the 98th percentile and 95% CI	Tmean reported here (*see other metric of temp in paper*) **Total population** = 1.16 (1.09 to 1.23) Male = 1.09 (1.00 to 1.19) Female = 1.21 (1.12 to 1.32) **≤74 years** = 1.07 (0.96 to 1.20) Males = 0.99 (0.84 to 1.16) Females = 1.08 (0.88 to 1.32) **>74 years** = 1.19 (1.11 to 1.28) Males = 1.16 (1.02 to 1.32) Females = 1.22 (1.11 to 1.35) Early summer total pop = 1.31 (1.17 to 1.47) Late summer = 1.09 (1.00 to 1.19)	**Main messages:** 1. Risk increases with age 2. Risk appears to be higher for females in Switzerland and for males 3. Risk is higher for early season heat events

The effect of temperature on mortality in Stockholm 1998–2003: A study of lag structures and heatwave effects **Rocklöv and Forsberg**	2008	1998 – 2003 time series analysis of daily deaths due to non-external causes, by age and specific causes of death CVD and Respiratory deaths Generalized additive Poisson regression models	Unclear what measure of temperature has been used. Methods simply state that temperature data was obtained from air quality monitoring station in city centre of Stockholm.	**Age** <65 65-74 >74 **Confounding** Influenza activity; seasonality; holiday; day of the week	Relative risk of mortality in summer for every degree increase above mean temp	**All age** Summer with *Lag 0-6,* 1.4% increase for each degree increase above mean temps (RR=1.014?) *Lag 0-1* = 1.3% increase in risk for every degree increase above mean temps (RR=1.013) **<65s @ lag 0-1** 0.5% increase (RR=1.005) **65-74 @ lag 0-1** 1.5% increase (RR-1.015) **>74 @ lag 0-1** 1.6% increase (RR 1.016)	**Main messages** • Risk increases with age for heat
Susceptibility to mortality related to temperature and heat and cold wave duration in the population of Stockholm County, Sweden **Rocklöv et al**	2014	Mortality and hospitalisation data for all residents in Stockholm County from the Swedish National Cause of Death Register *and* matched with the data on the same individual	Daily maximum temperature, lag 0-1taken from AQ monitoring station in the centre of Stockholm,	**Age** 0-44 45-64 65-79 80+ **Gender** **Country of birth** Nordic or elsewhere **Pre-existing indicators** - Out of hospital	Odds ratio (OR) in each of the groups studied per degree increase of temperature	See Tables 1 & 2 within study for full results. **Significant effect modification observed:** 85+ alone = 1.010 (1.002 to 1.017) Men >65 = 1.009 (1.001 to 1.017)	Interesting that this paper suggests that heatwave length plays a role in risk but makes sense. **Main messages:** 1. Risk of mortality increases with age 2. Risk for those under 65 appear to increase more

	in the National Hospital Discharge Register. Indicators created to denote if hospitalisation occurred on same day of death. Indicators of pre-existing conditions were based on underlying or secondary contributing causes in the hospital discharge register. Hospitalisations 0-28 days before death were not considered pre-existing – due to potential deterioration of patient over that period 1990 – 2002 Indicator for income also included, but	- Hospitalised - *Hospitalised with*: 1. Diabetes Mellitus ever 2. COPD ever 3. Mental disorder ever 4. Substance abuse ever 5. Acute myocardial infraction (AMI) and recurrent myocardial infraction 6. CVD 7. Cerebrovascular disease 8. Respiratory disease 5,6,7 & 8 all within 28 days 2-years of hospitalisation **Confounding** Study design controls for a rang of individual level factors which could cause confounding in other study designs. In addition, analysis controlled for holidays, NOx and O3		Women >65 = 1.007 (1.000 to 1.015) Out of hospital >65 = 1.009 (1.001 to 1.016) Hospitalisation COPD ever <65 = 1.030 (1.003 to 1.056) Hospitalisation Mental disorder ever >65 = 1.011 (1.002 to 1.019) Hosp AMI within 28 days and 2-years <65 = 1.041 (1.003 to 1.082) Hosp Resp within 28 days and 2-years >65 = 1.009 (1.000 to 1.018)	than 65+ with longer HW episode 3. Heat intensity is not associated with mortality in hospitalized patients, but is associated with mortality in non-hospitalized individuals over 65 (Table 3). 4. Mortality in non-hospitalized patients younger than 65 significantly increases with heat wave duration, and even more so in hospitalized patients. 5. Persons above 65 years with a pre-existing mental disorder are at dramatically increased odds of death with heat wave duration, and also with increasing temperatures in general

		derived from census data so not considered in this assessment Associations were examined using the **time stratified case-crossover design**					6. Having pre-existing COPD or myocardial infarction is associated with large increases in death rates when temperature increases in persons younger than 65 7. In persons older than 65, mortality increases when temperature increases among persons with a pre-existing respiratory disease, and for heat wave duration among persons with a pre-existing cardiovascular disease
How to estimate exposure when studying the temperature-mortality relationship? A case study	2016	All non-accidental mortality for three age groups – all ages, under 75 and 75+ Temperature (max, min, mean) and other	Assessment of what met factors best predict mortality, and then 4 different approaches for assigning exposure	**Age only** -All ages -Under 75 -75+ **Confounding** Long term trends; seasonal variations days of the week; public	Relative Risk of mortality at the 99th percentile compared to the 90th percentile for heat. All cause used to determine what temperature	All ages RR = 1.07 (1.04 to 1.10) <75 RR = 1.02 (0.98 to 1.05) >75 RR = 1.10 (1.07 to 1.14)	Choice of indictor did not really have a significant effect of the risk estimates, average temp used as that is what has been used in previous assessments of

of the Paris area **Schaeffer et al**		meteorology factors from range of monitoring stations (humidity, AQ)	1. Temperature in Paris (not clear if this is mean) 2. Average temp from all monitoring station in study area 3. Average of all stations in study area, but weighted based on population density of each station – i.e. more weight given if more people lived there 4. Average of monitoring stations based on their land use category	holidays; air pollution (not specified)	metric to use in rest of study. Average temp selected as best indicator for this study. So that is what is reported here. Other results available within paper		temp-mortality relationships. Main messages: 1. Risk of mortality increases with age
Vulnerability to Heat-Related Mortality: *A Multicity, Population-Based, Case-Crossover Analysis* **Stafoggia et al**	2006	*Case-Crossover Analysis* Residents at 35 or older who died in 4 Italian cities over slightly different periods for all	Daily mean apparent Temp from nearest city airport	Age Gender Marital status Income (by area mean) Hospital admissions within 2-years (see paper for full list) Place of death -out of hospital	All results are expressed as pooled odds ratios (ORs), with 95% confidence intervals (CIs), of dying on a day with 30°C apparent temperature	• Overall OR of 1.34 (CI = 1.27–1.42) at 30°C relative to 20°C. • The odds ratio increased with age and was higher among women (OR =	Study design useful for exploring further details about mortality at the individual level. However, use of hospitalisation is proxy for chronic conditions.

	non-accidental deaths. 2000-2003 in Bologna 1999-2003 in Milan 1998-2001 in Rome 1997-2003 in Turin Deaths data on cause of death by age and gender. Deaths data linked with city specific population registers providing info on: -marital status (Milan and Turin) -Census block of residence -Median pop income for each census block used as area socioeconomic status indicator. Deaths data then linked to hospital discharge files,		-in hospital (two different time periods) -nursing home **Confounding** Population changes in summer; holidays; influenza epidemics; linear terms for PM10; barometric pressure; summer ozone (as sensitivity analysis)	relative to 20°C. Effect modification was tested and results are reported as the relative effect modification (REM) index calculated as the ratio between the specific odds ratio and the odds ratio from the reference category.	1.45; 1.37–1.52) • and among widows and widowers (1.50; 1.33–1.69). • Low area-based income modestly increased the effect. Among the pre-existing medical conditions investigated, effect modification was detected for: • previous psychiatric disorders (1.69; 1.39 –2.07), • depression (1.72; 1.24 – 2.39), • heart conduction disorders (1.77; 1.38 –2.27), and • circulatory disorders of the brain (1.47; 1.34–1.62). • Temperature-related mortality was higher among people	Main messages: • Risk increases with increasing age, with risk higher for females across all ages • Living along seemed to increase risk • psychiatric disorders, depression, heart conduction disorders, circulatory disorders of the brain all increased risk of death on hot days • Being in care facilities (care homes and hospitals) increased risk of death on hot days

	with all hospital admissions between 28 days and 2 years were selected for inclusion. Primary causes and secondary contributing diagnoses. 28 groups of diagnoses included Use of 28 data 0-28 days before deaths used to ascribe place of death: -out of hospital -Discharged 2-28 days before death -in hospital -in nursing home Step 1: temp relationship for each city to find the best model Step 2: analysis of each effect modifying variable Step 3: City specific results combined for meta-analysis		residing in nursing homes, and a large effect was also detected for hospitalized subjects. See paper for full results	

Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis **Stafoggia et al**	2008	*Case-Crossover Analysis* 2000-2003 in Bologna 1999-2003 in Milan 1998-2004 in Rome 1997-2003 in Turin Only 65+ included in analysis Mortality data linked with hospitalisation data for previous 2-years before death. Gender, age, marital status, median income of the census block of residence and location of death Type of hospital (public or privet) Type of ward (general medicine or medical or surgical wards	Apparent temp from nearest city airport	Age Gender Marital status Income (by area mean) Place of death Hospital admissions within 2-years (see paper for full list) Hospital type Ward type Primary and secondary causes of hospitalisation Acute clinical conditions for admissions up to 28 days before deaths PoD = in hospital <2 days AND in hospital >2 days) **Confounding** PM10; influenza epidemics; population decrees in summer; holidays; barometric pressure. Ozone had no effect in exploratory analysis so not included.	The results are expressed as pooled odds ratios (OR), with 95% confidence intervals (95% CI), of dying on a day with a 30uC apparent temperature relative to a day with a 20uC apparent temperature Effect modification was tested and results are reported as the relative effect modification (REM) index calculated as the ratio between the specific odds ratio and the odds ratio from the reference category.	See paper for full results, but key results (based on REM index) Age = 1.22 (85+) Gender = 1.06 (F) Marital status = 1.19 (single) ***Chronic*** Other ischemic heart disease = 1.00 Conduction disorders = 1.26 Cardiac dysrhythmia = 1.06 Cerebrovascular disease = 1.15 Diseases of arteries, arterioles, and capillaries = 1.08 Disorders of the thyroid gland = 1.04 Diabetes = 1.10 Obesity = 1.63 Psychoses = 1.56 Paralysis = 1.11 Other disorders of CNS = 1.20 Hip fracture = 1.25 ***Acute*** Diseases of pulmonary circulation = 1.55 Heart failure = 1.28	**Main messages**: • Risk highest in care homes and for those in hospital for less than 2-days. Risk also significant for those in hospital more than 2-days • Temperature-related mortality is high among patients already hospitalised at the time of a heatwave. • Elderly subjects and those hospitalised in general medicine wards are especially vulnerable. • Acute heart failure, stroke and exacerbation of chronic pulmonary diseases are associated with an increased heat- related mortality among hospitalised patients.

		with low-moderate care, with med-high care, ICU) Deaths data then linked to hospital discharge files, with all hospital admissions between 28 days and 2 years were selected for inclusion. Primary causes and secondary contributing diagnoses. 28 groups of diagnoses included 0-28 days before death used to categorise sudden deterioration weeks or days before death. Acute clinical conditions assessed during last 28 days before death				Cerebrovascular diseases = 1.22 Pneumonia = 1.10 Chronic pulmonary disease = 1.87	
Apparent Temperature and Cause-	2011	Linked mortality and hospital admissions data.	Maximum apparent temperature	**Individual level**: Age Gender	Percentage increase in risk (%) and 95%	Heat only reported here	No sure what this is telling us. There was no significant

Specific Mortality in Copenhagen, Denmark: A Case-Crossover Analysis **Wichmann et al**	Specifically looking at Respiratory disease, CVD, and cerebrovascular disease The time-stratified case-crossover design controls for individual confounding by design	Place of death Cause of death Hospital admissions and discharge **Community factors**: SES – area indicator based on home address of each death record **Confounding** AQ, influenza	confidence intervals per inter-quartile increase in the 6-day cumulative average of $Tapp_{max}$ (in °C) during warm period of 1 January 1999– 31 December 2006 in Copenhagen.	No significant increases in mortality risk were observed for the three causes of death by the individual or community level factors. Overall warm season was associated with decrease on CVD. All modifying factors investigated appear to have a protective effect on CVD However, some patterns were seen: **Age** increasing age appeared to increase risk of respiratory disease deaths, increasing age lowered risk of cerebrovascular disease deaths and no pattern observed for CVD deaths **Gender** Change in risk appears to be greater for males	increase observed – the only significant finding was for a decrease in risk for CVD amongst the 80+ years group, which doesn't really make a lot of sense. Uncertain about the CVD results, and potentially will not include. This study is not particularly useful. **Main messages** (excluding CVD): • Risk increases with age • Risk for males appears to be higher • Risk of death increases out-of-hospitals for Resp and cerebrovascular disease •

						across all three CoD categories (with CVD % change being protective) **Place of death** % increase is risk of death occurring out of hospital for Resp disase and cerebrovascular disease. CVD risk in hospital decreases the most.	
Extreme Temperatures and Mortality: Assessing Effect Modification by Personal Characteristics and Specific Cause of Death in a Multi-City Case-Only Analysis **Medina-Ramón et al**	2006	Case only study design USA city-specific logistic regression model was fitted, and an overall estimate calculated in a subsequent meta-analysis.	Minimum temperature in the warm season	Primary and secondary causes of death (do we need to think about this?) Place of death Age Gender Race Educational attainment Chronic conditions listed on death certificate: -diabetes -COPD	Odds ratios for someone dying with a particular characteristic or not, e.g. female v male	Results reported for subject level modification factors, as below: Age ≥65 = **1.020** (1.005 to 1.034) Female = 1.011 (0.997 to 1.024) Black = 1**.037** (1.016 to 1.059) Low education = 1.016 (0.999 to 1.033) Out of hospital = **1.066** (1.036 to 1.098) Diabetes = **1.035** (1.010 to 1.062) COPD = 1.004 (0.979 to 1.030)	Order of relative strength if risk: 1. Out of hospital 2. Race 3. Diabetes 4. Age Gender, education level and suffering from COPD not significant but still increased odds 5. Low education 6. Gender 7. Suffering from COPD Would the risk profile of those dying in hospital be the same as those dying in the

							community (out of hospital)?
Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome **Schifano et al**	2009	Cohort study of 651,195 residents in Rome 65 years or older Linked regional hospital discharge files and municipality population register using individual social security numbers to link individual records. 13 diagnostic groups considered for record of hospitalisation or hospital visit. Subjects were classed as having attended for that reason or not. Diagnostic groups needed to be either the primary or contributing factor in admission	Daily maximum apparent temperature taken at Ciampino airport (Rome)	Age - 65-75 - 75-84 - 85-94 - 95+ Gender Marital status SES Hospital diagnosed disease – see paper for full list **Confounders** No mention of controlling for potential confounding factors, other than conducting analysis by age groups		Relative risks (RR) of dying on heat wave and non-heat wave days for each of the modalities of the covariates and the REM (Relative Effect Modification) index.	**For the 65-74 group:** Results suggested that the excess in mortality during heat waves was higher among those who were previously hospitalised for a *chronic pulmonary disease*, and to a smaller extent, for *psychiatric disorders* (not considered significant) and among those *not married, widowed or divorced* **For the 75+ group:** Relative risks showed that the excess in mortality during heat waves is significantly higher in *females*, and for all those not *married, widowed or divorced* In addition, having been admitted 4+ times was also	This is a Rome specific study, and therefore may not be generalisable. Interesting that there is a difference in the age groups Overall risk is clearly increases as age increases, however of the diagnostic conditions investigated none appeared to be associated with increased risk for the older group.

		Poisson regression model to estimate the adjusted relative risk of mortality during heat wave days versus not heat wave days.				associated with increased risk, although this is not considered significant.	
The influence of pre-existing health conditions on short-term mortality risks of temperature: Evidence from a prospective Chinese elderly cohort in Hong Kong **Son et al**	2016	66K+ participants all 65+ years old, registered at elderly centres in Hong Kong between 1998 and 2001, and followed up until 2011, with vital stat measurements made and reported chronic conditions (which we validated by clinical records) All natural deaths then assessed against the identified chronic conditions using distributed lag,	Not clear	No chronic disease reported Diabetes Circulatory system diseases COPD All subjects	Cumulative relative risk at different temperature percentiles and at different lags. Heat assessed at 0-1 lag and 0-3 lag – not much difference between reported RR, therefore have used lag of 0-1 for this as that fits with most other studies.	Non-disease group: RR = 0.97 (0.95 to 1.04) Diabetes: RR = 1.03 (0.96 to 1.11) Circulatory system diseases: RR = 1.04 (1.00 to 1.08) COPD: RR = 1.09 (0.99 to 1.20) All persons: RR = 1.01 (0.98 to 1.04)	Only circulatory system diseases appears to be significant. However its unclear if the analysis is is adjusted for seasonal effects, i.e. only using summer months to look at heat or if the analysis is using year-round data to investigate the relative risks – which may dilute the overall effect observed. Main messages: • Increased risk of mortality associated with presence of chronic condition

		non-linear methods Presuming that it those included were those who had died within a certain period, but this is not stated. ***Prospective cohort study design in Hong Kong***					• Circulatory system deaths only chronic disease category showing significant effect • While not significant, largest effect observed for COPD, followed by circulatory systems diseases followed by diabetes and finally all persons. • Not having a chronic disease appears to be protective
The mortality burden of hourly temperature variability in five capital cities, Australia: Time-series and meta-regression analysis	2017	Daily all-cause mortality records from each of the 5 Australian cities. Prevalence data of each city (not linked data, but rather prevalence data from unknown stats source) of	Mean temperature – daily?	Self-assessed health status Long term health condition Cancer COPD Diabetes	Unclear what is being reported, but appears to be reporting the estimated slope of the risk curve for each potential modifying factor (beta)	No values actually provided, but chart included which suggests that: Prevalence of Cancer, COPD and Diabetes all significantly increase risk of mortality across all 5 cities.	The pre-existing conditions investigated here is based on prevalence data and not linked mortality data. Therefore, the same limitations likely exist in that the data used for assessing the effect modification of each chronic disease

Cheng et al. 2017	chronic conditions	Heart/stroke/vascular diseases Hypertension Overweight/obese Smoking Alcohol Fruit/veg intake Physical exercise		may not adequately represent the reality of the population investigated. Therefore this is not relevant to this review.

Appendix 2 – Data Management

CPRD data access

Health data for this study was obtained from Clinical Practise Research Datalink (CPRD) which provide real-world data for retrospective and prospective public health and clinical studies. CPRD provide anonymised individual level primary care data for patients which opt-in to their data being used for research purposes, and who's primary care practise contribute.

In addition to primary care data, CPRD also offer linked data sets for included individuals. Of relevance to this study are linked ONS mortality data, NHS Hospitalisation Episode Statistics (HES) Admitted Patient Care (APC) data set and HES Accident and Emergency (A&E) data. The data linkage is carried out by NHS digital and administered by CPRD, following standardised approach to ensure data quality, which has been described elsewhere.

Access to CPRD linked data sets are subject to protocol approval by CPRD's Research Data Governance Process, which includes review by an independent Expert Review Committee and Central Advisory Committee. Within the application, detail about the research question, methodology and specific data requirements of the study are required, and where questions arise, applicants are required to address these before protocols are approved. The protocol for study titled *Temperature extremes and clinical vulnerability in England; development of a risk stratification tool for primary care use* (ID number 21_000621) was approved on 14 February 2022. Once approved, there is a two-step process involved to obtain the data as described in the study population definition section of the approved protocol, in this case all deaths between 2016 and 2020.

The first step is termed “Type 1” data request and involves defining the study cohort. To do this, CPRD provide a minimum level of data required to identify individuals within the CPRD data base who fall into the study population. For this study, all deaths between 2016 and 2020 were provided as per the approved protocol. This limited data set included patient IDs, date of death, and cause(s) of death as recorded in ONS mortality data. Because this study considered all-cause mortality, cause of death information was not required to define the study cohort, and therefore, all individuals who died within the study period were included.

This resulted in 1,143,518 unique patient IDs within CPRD database which have valid ONS mortality data over the whole study period of 1st Jan 2016 to 31st December 2020. These unique patient IDs were used to extract all primary care data available. The extracted primary care data is then saved on a secure drive where data cleaning and processing must take place, as per LSHTM data agreement and licence.

The second step is termed “Type 2” data request, and this is the stage at which linked data sets are requested. Using the list of patient IDs used to extract primary care data from CPRD systems, as outlined above, the same list is sent to CPRD to request the relevant data (ONS mortality, HES APC and HES A&E) for each individual patient ID. The additional data sets are linked via the unique patient IDs which run across all data sets, allowing users to link primary care records to other health data sets. Figure A2.1 below illustrates the type 1 and type 2 data request process from CPRD, and indicates how the study cohort population is defined, and how that definition is used to identify the included individuals linked data.

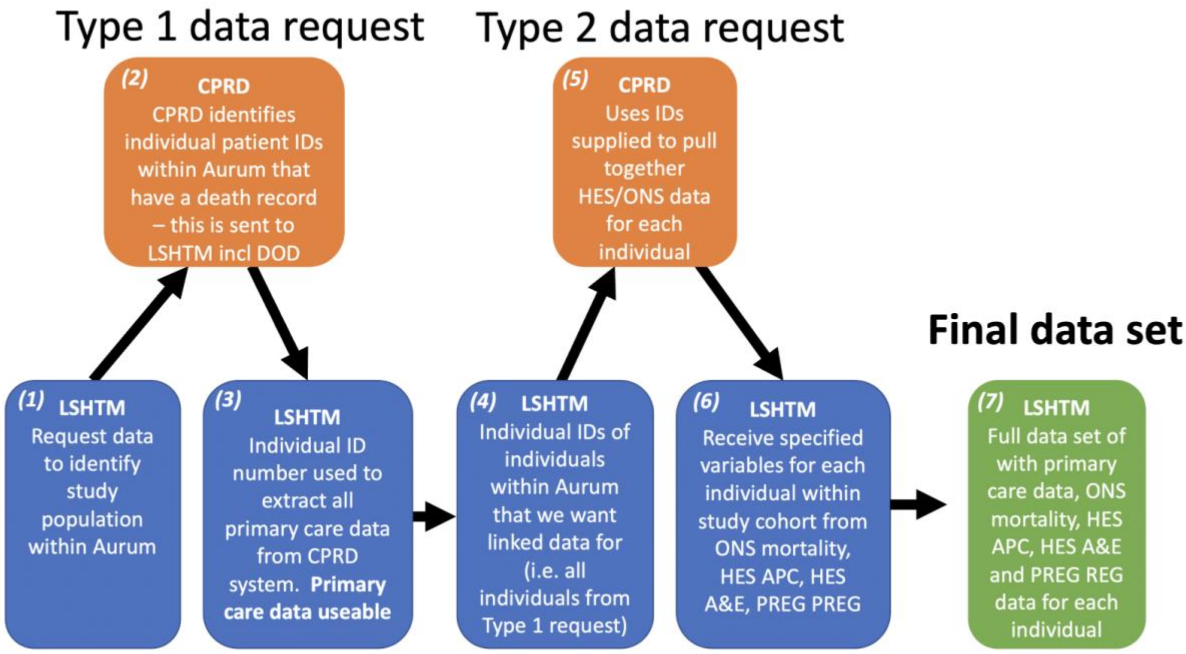


Fig A2. 1 *illustrates the steps required in both type 1 and type 2 data requests. Blue boxes (1,3,4 and 6)indicate actions required by LSHTM. The orange boxes (2 & 5) indicate actions required by CPRD. The green box (7) indicates the full set of linked primary care data. Arrows and numbers indicate the sequence of actions once the study protocol has been approved by CPRD expert review committee.*

CPRD policy states that any study with a study population greater than 600K individuals, study authors must identify variables within the linked datasets which are essential and only those variables identified will be supplied. This is due to the large volume of information for studies over 600K individuals and to cut down on processing requirements when transferring the linked data sets. To this end, study authors must supply a Data Minimisation Workbook detailing what variables will be required along with justification for that variables use, as indicated in the approved protocol.

Structure of data

CPRD Aurum

CPRD primary care data is available from two separate data bases, GOLD and Aurum, with Aurum being the larger of the two, and which is representative if the English population. For this study, Aurum was selected as the most relevant database to use. Aurum data comes in nine different data files, each containing primary care information related to each individual within the study sample. First the patient file contains basic patient information such as gender and registration details. The practise file contains details of each practise including which region within England each practise is. The staff file contains details for each staff member in each practise. The consultation file contains information on the type of consultation (e.g. home visit, phone call or practise visit): The observation file contains the medical history data, including symptoms, clinical measurements, lab test results and diagnosis, as well as demographic info. Observation file also contains two sub-files: the referral file which contains details of referrals by the GP, and the problem file which contain details of medical history that have been defined as a “problem” by the GP, all associated with observations. And finally, the drug issue file contains details of all prescriptions on the GP system.

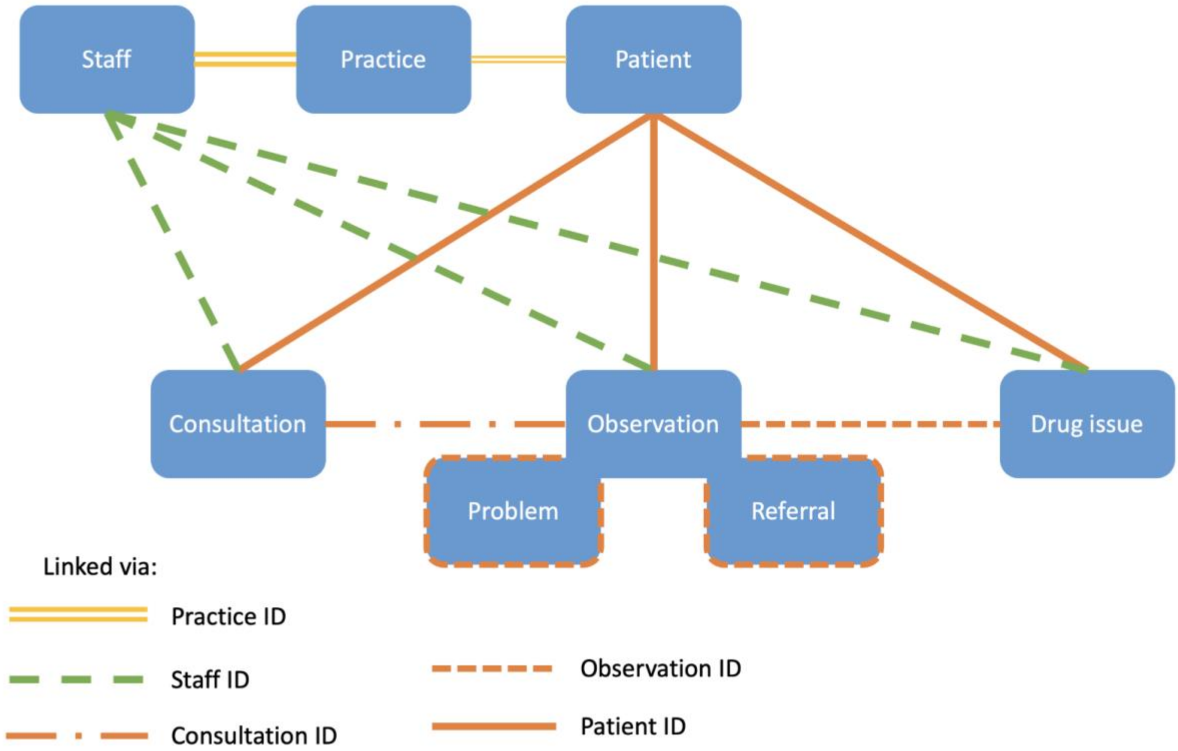


Fig A2. 2 ***CPRD Aurum Data Structure showing the different data files per individual and the linkages between (adapted from CPRD Aurum data guide)***

A full overview of the variables available within each data file type is available directly from CPRD (here) however, not all were required for this study. Table A2.1 below indicates the variables which were used when building the data for analysis for each of the primary care files. Data contained within the other data files were beyond the scope of this project, however, could potentially be used in further studies asking more specific questions.

Table A2. 1 Indicating key variables and brief description of each within the different CPRD Aurum data file types which will be used to build data for analysis.

File type	Variable	Description

Patient file – contains personal details of each individual included in the study	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	pracid	Encrypted unique identifier given to each participating practise. This is required to identify geographic region.
	gender	Patients gender
	yob	Year of birth
	mob	Mont of birth
	regstartdate	Date that the patient registered with participating practise. This is used to ensure patients with at least one year's wort of primary care data are included in the study.
	regenddate	Date registration ended. This may not be aligned with date of death, and could be used to further refine the study population
Practise file – contains limited info about each participating practise	*pracid*	Encrypted unique identifier given to each participating practise. This is required to identify geographic region.
	region	indicates where in the UK the practise is located by Strategic Health Authority boundaries (use of lookup table required)
Observation file – contains all clinical details of consultations – including diagnosis, vistits measurements etc.	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	obsdate	Date associated with the event (visit/consultation etc)
	medcodeid	Codes associated with medical terms recorded within primary care . Code is unique to Aurum and is the primary variable for defining chronic conditions and identifying any clinical results for measurements/tests.
	value	Provides the numerical value of any test results or measurements associated with the *medcodeid*
	numunitid	provides the units of measurements/test results (use of look up file to de-code)
Drug issue file – provides details of all prescribed medications	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	issuedate	Date on which the prescription was issued
	prodcodeid	Unique CPRD code used to describe treatment issued. This is the primary variable for defining the type of medications issued to the individual

ONS mortality

Deaths in England by law must be registered with the General Registry Office (GRO) within 7 days following a death. These registered deaths are then loaded onto a database operated by the Office for National Statistics (ONS). Mortality data is provided by ONS which is linked to individual primary care records by NHS Digital using an 8 step linkage algorithm which matches individual records by NHS number, gender, date of birth and postcode. Full details of this linkage procedure is available from CPRD. Unlike Aurum data, ONS mortality data is provided in a single data file. Table A2.2 below indicates the key variables which will be used within this analysis, and which were requested as part of the CPRD data minimisation policy for linked data requested.

Table A2. 2 Indicates the variables of interest contained within the ONS mortality dataset, along with a brief description of what each variable relates to.

File type	Variable	Description
Mortality file	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	dod	Date of death – date the death occurred
	nhs_indicator	Indicates if the death occurred in an NHS establishment
	pod_catagory	Indicated place of death category
	cause	Underlying cause of death as recorded on death certificate
	cause1	Primary contributing cause of death (may be different to underlying cause in some circumstances)
	cause2	Secondary contributing cause of death

HES Admitted Patient Care (APC)

The HES APC data contains details of all admissions to English NHS health care providers, including both elective and emergency admissions. The patients include private patients and residents outside of England, who were treated by NHS health care providers, including treatment by the independent sector, if funded by the NHS. All NHS health care providers in England, including acute hospital trusts, primary care trusts and mental health trusts provide data. There are several data files which link to individual primary care records that contain a wealth of information.

Within HES APC, there are a number of files which relate to care provided to admitted patients over the whole period of their stay. Data for one complete stay in hospital is considered a spell, while each time a patient is seen by a consultant, this is considered an event. Therefore, some patients may have a record of several events within one spell of hospitalisation. In addition, each spell and or event may be associated with a number of identified diagnoses. Figure A2.2 below illustrates the differences between events and spells.

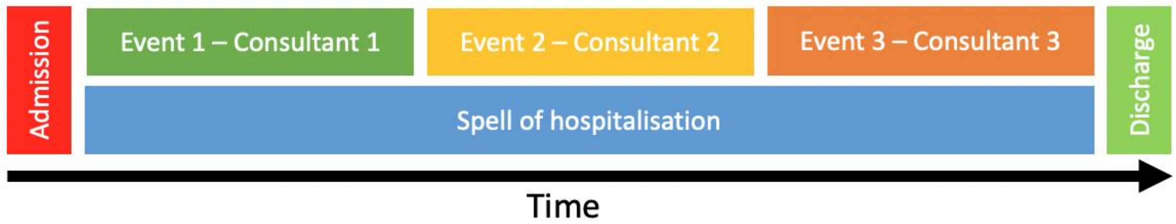


Fig A2. 3 ***indicates how spells and events are defined within HES APC data. Each event is associated with the period in which the individual is transferred to the care of a new consultant. This can be following admission or can be an internal referral within the same trust. Each time a patient is transferred to the care of a new consultant, a new event is***

recorded. A spell however is the whole duration of that individuals stay within the hospital, from admission to discharge.

Table A2. 3 Indicates the key variables which will be used within this analysis, and which were requested as part of the CPRD data minimisation policy for linked data requested.

File type	Variable	Description
Patient	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	gen_ethnicity	Patient's ethnicity derived from all HES data
Hospitalisation	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	spno	Spell number uniquely identifying a hospitalisation
	admindate	Date of admission
	discharged	Date of discharge
	adminmeth	Method of admission (A&E, elective etc)
Episode	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	spno	Spell number uniquely identifying a hospitalisation
	epikey	Unique key identifying an episode of care
	epistart	Start date of episode
	epiend	End date of episode
	eorder	Order of episode within spell
	epidur	Duration of episode
	epitype	Type of episode (e.g. general, delivery, birth, psychiatric etc)

Diagnosis	*patid*	Encrypted unique identifier for each individual included within the database. Critical for linking data at the individual level.
	spno	Spell number uniquely identifying a hospitalisation
	epikey	Unique key identifying an episode of care
	ICDx	An ICD10 diagnosis code in XXX or XXX.X format (for both episodes and hospitalisations

Building data set for analysis

Before any data analysis can begin, the data as described above will need to be cleaned and formatted so that it is in an analysable format. This transformation of raw data into a suitable format for analysis involves a number of separate, but related processes. This includes receiving the raw data and converting to a suitable format (e.g. .txt to .dta), generation of clinical code lists which are used to identify individuals with specific records within the data, using the code lists to define the study population for each factor of interest and finally merging data for all factors of interest into one final dataset for subsequent analysis. Each of the stages are briefly described below.

Clinical code lists

Obtaining clinical code lists for the relevant outcomes and factors of interest is the first step required building data for analysis using primary care data. Primary care records are entered manually by primary care professionals directly into patient records software. The software allows clinicians to entre data as both drop down box options but more commonly as free text. The free text is then coded using the international standard coding SnoMedCT, which maps the data entered into clinically relevant groups based on the terms entered into the system. In addition, older records may not be mapped to SnomedCT concepts and will use the older ReadTerm coding system. The mixture of coding systems and coding based on free text terms makes primary care data notoriously messy and difficult to navigate.

In an attempt to make this easier for researchers, CPRD generate unique codes which are mapped to both SnomedCT and Read terms. The resultant MedcodeIDs can then be used to define conditions/personal info/prescribed medication etc within the data. CPRD Aurum MedcodeID are specific to this data base and therefore code lists generated for primary care studies using different databases are not appropriate.

What code lists will be required

Upon review of the literate a list of conditions was identified that had either previously been shown to be associated with increased risk of heat related mortality or physiologically have a plausible mechanism to cause ill-health at elevated temperatures. These include conditions of the circulatory system, respiratory system, diabetes, mental health conditions, Alzheimer's and Dementia, Parkinson's disease, conditions of the kidney and thyroid, and obesity. There are some factors which may be recorded within primary care records which may also indicate increased risk such as high BMI (linked to obesity) and blood pressure observations (both diastolic and systolic blood pressure).

In addition to clinical conditions, a number of other factors have also been identified within the literature which suggest an association between high temperatures and mortality, including isolation (living alone), ethnicity, gender and age. Furthermore, while evidence is limited, certain types of medications are also suspected to modify the risk associated with elevated temperature. These types of medicine are listed within several heat-health action plans across the globe.

Having consulted with academic experts, EMIS (patient records software providers) and public health consultants (UKHSA), the decision was taken to ensure that the code lists developed for this project are aligned to ICD-10 categories where at all possible. The rational for this is that ICD-10 categorisation is a standardised approach at classifying conditions which is static. One of the issues researchers have when dealing with primary care data is that it is messy and constantly evolving. For example, and as previously mentioned, in the UK two coding and classification methods are used, Read and SnomedCT. These two methods have both different coding structures but also different hierarchy structures. By ensuring that this analysis is aligned to ICD-10 classifications will ensure that the results are interpretable in other locations around the world which may not use the same coding systems and structures. However, corresponding Read and SnomedCT codes will be kept for full transparency of how each code list was developed, and to allow any future development of this work.

The Electronic Health Records Group within LSHTM have generated a number of Aurum specific clinical code lists. In addition, there is a concerted effort amongst researchers to share code lists which have been created to increase both transparency of research results and consistency across research groups in terms of how variables analysed are defined. After a comprehensive search of both LSHTM data compass network and CPRD data base of approved studies, a number of clinical code lists were identified which are of relevance to this project.

Tables A3.1 and A4.1 within Annex 2 and 3 outline the specific hypothesis being tested for each variable, the likely mechanisms which may lead to ill health during heatwaves, a comment on how that variable is defined within the clinical code lists and information about the source of the code, i.e. reference to existing code lists or outlines how the bespoke lists were generated and clinically validated.

Bespoke clinical code list development

There is no standardised approach to developing clinical code lists. However, after extensive discussions with other researchers who work in this area and searching for relevant literature, the following approach was adapted. Using the CPRD medical term dictionary which maps all terms entered in to EMIS systems to unique CPRD specific medical code, a systematic search of relevant clinical terms was conducted. This included selection of suitable terms of relevance to the condition/factor the code list was being generated for (e.g. Alzheimer's; dementia etc) and selection of any exclusion terms which would otherwise return a large number of unwanted terms (e.g. family history). Specifics of each code list developed varied slightly and was an iterative process of addition of exclusion terms to ensure the search of the CPRD medial term dictionary was as specific as possible, while also ensure that all potentially relevant terms were included. The systematic approach employed in this study to generate bespoke clinical code lists followed three stages as described below.

<u>Step 1 – automated search of relevant terms</u>

First, the selection of inclusion and exclusion terms are defined which are then used when searching the CPRD Aurum dictionary in Stata. Inclusion terms should be any word/series of words which are of relevance to the condition/factor of interest. For example, inclusion terms for factor *living* *arrangement* may include “living alone, cohabitation, homeless, accommodation” etc. Exclusion terms should be selected to help limit search results that are not of relevance to the condition/factor. For example, again for living arrangement, exclusion terms may be “stand, tibia, fibula, aortic, revision, reaction, test” etc. Choosing exclusion terms is an iterative process and guided by the results of the search of inclusion terms. This may involve scanning the search results of terms returned to identify terms which may not be relevant, then going back to add any additional exclusion terms identified. This can reduce the number of potentially relevant terms identified significantly. To check for that the resultant number of terms is roughly in the right region, clinical code lists from other databases were inspected for total number of included terms. Comparing the number of total included terms provided a level of confidence that the automatic process of code development was adequate before full validation by clinician.

Step 2 – manual inspection

The next step is to manually inspect the returned clinical terms and make a judgment on their inclusion or exclusion. This is required in addition to the automatic exclusion process as some included terms may still not be of relevance to the specific condition/factor of interest. For example, the returned search terms for the condition *depression* included a large number of other clinical terms which included the word “depression” in it, like “depression in the skull” which is not of relevance to the mental disorder. In addition, where a condition/factor was to be categorised, manual inspection and indicating which category each included term was to be assigned was also undertaken. For example, eFI score is recorded directly into primary care records, and therefore, these can be obtained directly from CPRD data. However, this needed to be categorised into mildly frail, moderately frail, and severely frail categories. It is within tis manual step that this categorisation is carried out. The manual inspection (and categorisation) was carried out by two individuals independently.

Step 3 – clinical validation

Upon completion of manual inspection of the initial code lists (indicating what codes will be included, and where appropriate categorised, and which will be further excluded), these were then shared with a clinician for clinical validation. The clinician reviewed the lists and indicated whether or not each term should be included or excluded. Where differences occurred between the clinical expert and the initial assessment in step 2, the clinician provided rational for exclusion and the clinical judgment was accepted. The final validated clinical code list was then generated and saved as .txt file format.

Prescribed medication code list generation

Generating code lists for prescribed medications which have been identified as potential risk modifiers will require a slightly different approach to that of medical conditions or other personal information as described above. Rather than using the Aurum Medical Term Dictionary, the EMIS Product Dictionary is used. Within this, every prescribed drug, dosage and formulation is coded to a unique CPRD product code (*prodcodid*) which is the equivalent of the *medcodeid* for conditions. In addition, because the drug substances to be included in this study are broad drug categories based on British National Formula (BNF) chapter, a slightly altered approach was required.

Step 1 – initial search

A search of the EMIS Product dictionary was conducted using BNF chapter numbers only of the relevant drug class (e.g. 2050501, 2050502 and 2050503 for ACE-inhibitors). This returned only drugs which matched the relevant BNF chapters. However, not all drugs within the dictionary have a corresponding BNF chapter recorded.

Step 2 – refinement of search criteria

Therefore, to increase the sensitivity of the search, a second search was carried out using the BNF chapter again, but also including a range of chemical substances associated with the relevant BNF chapter. For example, for ACE-inhibitors this included *aliskiren azilsartan medoxomil* amongst other chemical terms. This search would return a number of other potentially relevant terms.

Step 3 – manual inspection

As with the condition/factor clinical code lists the next step involved manual inspection of the potential drug product terms/BNF chapters returned by the automated aspect of the search described in steps 1 and 2. Here RT and DO individually inspect the returned detail and made a judgment on inclusion or exclusion.

Step 4 – clinical validation

The final list of potential drug product code lists was shared with a clinician who specialises in clinical therapeutics to validate. Where differences occurred between the clinical expert and the initial assessment in step 2, the clinician provided rational for exclusion and the clinical judgment was accepted. The final validated clinical code list was then generated and saved as .txt file format.

CPRD data formatting

Converting all txt files to stata .dta files

As previously outlined, CPRD primary care data comes in several different file types (patient data, practise data, observation data etc), and when extracted from CPRD system, each of those files is further broken down into sub-file. For example, in the Patient file may be made up of two different .txt data files, the observation file will be split into 102 .txt data files etc. The first step is to convert these all into .dta files.

This was done by converting all individual txt files per file type (patient, practise, observation etc) into a temporary *dta* file and then appending all temporary *dta* files into one master file for that file type. However, for the larger data files such as drug issue and observation files, the number of observations were extremely large. Therefore, for these two file types, these were converted into a number of *dta* files.

Table A2. 4 Indicates the number of txt files and the converted dta files for each file type.

File type	Number of raw txt files	Number of raw dta file
Drug issue	26	1
Observation	102	4
Patient	88	4
Practice	2	1
Problem	2	1
Referral	2	1
Staff	2	1

Generating variables

Primary care records

For all binary variables, code lists described above will be used to identify relevant entries within individual primary care records by merging the specific code lists with the CPRD observation files using the many to 1 merge function in Stata. The merge will be carried out by *medcodeid* variable. Upon identification of relevant observation records, only *obsdate*, *variable of interest*, *patid* and *obsid* will be kept and saved. This will be done by condition groups, for example cardiovascular diseases, respiratory diseases, mental health outcomes etc. Rational for this is that there are a large number of variables to be defined and the large number of observations within the *observation* data file. In addition, by defining the variables for the whole data extract (i.e. not just deaths which occur in the summer months) these defined variables will be available for use in other studies, for example looking at the impact of cold temperatures in winter months.

For variables which require numerical values (e.g. diastolic blood pressure and BMI) there are a number of Stata programmes which have been developed by LSHTM which standardise the approach to deriving these variables.

HES APC records

Within this study, HES APC data will be used as part of sensitivity analysis to assess the use of different data sources to define chronic disease, i.e. the use of primary care data versus the use of hospitalisation data as has been used in previous studies. Therefore, not all variables which have been identified in tables 1,2 and 3 above will be required. Variables will be selected based on the results within the main analysis, where there is strong evidence of an association between the exposure and outcome by population sub-group (i.e. chronic condition group).

Unlike primary care data, hospitalisation data is coded using the International Classification of Disease, 10th Revision (ICD-10) coding. ICD-10 coding uses standardised codes to define disease groups, and as such ICD-10 codes are available directly from ICD-10 dictionaries available from multiple sources, including World Health Organisation. The codes can simply be applied to the linked data to observations based on specific ICD-10 codes within the diagnosis file for the HES APC.

A similar process of merging clinical code lists with the diagnosis file will be carried out to identify all relevant HES records. Once these have been identified, the hospitalisation file will be merged to allow dates of admission/discharge to be determined, ensure that each hospitalisation is unique (to minimise double counting) and to determine what type of admission it was (i.e. emergency or elective). As with primary care records this process will be carried out for each condition of interest for the whole population (2016 to 2020), to allow other researchers to use the pre-defined variables in additional analysis, i.e. cold weather.

ONS Mortality data

Only mortality records for those who died between 2016 and 2020 have been provided. Date of death is the outcome of interest and as such this will be the main point of reference for follow up period to define chronic conditions.   Apart from date of death, all other variables will require the use of either lookup tables (place of death category, NHS establishment etc) or the use of ICD-10 codes (underlaying cause of death, primary cause of death etc) to define them. A similar approach to defining /classifying each variable will be used as previously described for primary care records. In addition, date of death in combination with date of birth in the patient file will be used to generate a new variable *age,* which will indicate the age at death.

Deprivation data

As part of the request for liked data from CPRD, index for multiple deprivation (IMD) quintiles at the patient GP practise post code level was also requested. IMD was linked by patid.

Temperature data

Population-weighted daily-mean ambient temperature, taken from a dataset generated for work contributing to UKHSAs Health Effects of Climate Change report. These are based on the most up to date datasets for the UK, including HadUk-grid, which has been widely used and is generally regarded as the go-to for gridded historical observations. This data is available at the regional level and therefore will be merged by region and date of death.

Air quality data

Geographic resolution of CPRD data is a significant limitation of this study in terms of exposure assessment and adjusting for any potential contributing effect of other exposures not of primary interest, for example air pollutants. The correlation between temperature and the concentration of common air pollutants which are known to cause harm to health is well established. And it is common practise to adjust for the presence of these co-exposures. However, while temperature at the regional level is well correlated, concentrations of air pollutants disperse quickly with distance from the pollution source. Therefore, within this analysis adjustment for air quality will only take place within sensitivity analysis within London region.

A single daily series of mean concentrations of ozone (O_3) nitrogen dioxide (NO_2) and PM_{10} were derived using data downloaded from London Air historical data portal (https://www.londonair.org.uk/london/asp/datadownload.asp). Mean daily values were calculated using six background monitoring stations within London, with monitoring sites chosen to provide representation from across the London region. Table A2.5 below indicated which monitoring sites were used when calculating mean values for each pollutant.

Table A2. 5 Indicates which monitoring sights were used to calculate daily mean concentration of O_3, NO_2 and PM_{10} in the London region between 2016 and 2020.

Ozone monitoring sites	**NO_2 monitoring sites**	**PM_{10} monitoring sites**
Haringey - Priory Park South	*Haringey - Priory Park South*	*Islington - Arsenal*
Kensington and Chelsea - North Ken	*Kensington and Chelsea - North Ken*	*Camden - Bloomsbury*
Redbridge - Ley Street	*Newham - Wren Close*	*Kensington and Chelsea - North Ken FIDAS*
Lewisham - Honor Oak Park	*Lewisham - Deptford*	*Southwark - Elephant and Castle*
Wandsworth - Wandsworth Town Hall	*Lambeth - Streatham Green*	*Lewisham - Honor Oak Park*
Camden - Bloomsbury	*Southwark - Elephant and Castle*	*Lambeth - Streatham Green*

Building master data for analysis

As indicated, individuals who died between 1st May and 30th September 2016 to 2020 will be included in this study. As such these individuals will be identified within the study cohort based on their date of death (from ONS data). The resultant, and reduced, list of included individuals (removing those who died outside the months of interest) will then be used as the base to which all variables will be merged. Once all health-related data has been merged into the master data set for analysts, daily exposure data will be merged by date of death and governmental region of residence. Table A2.6 provides an overview of how the analysis data file will be developed and the key variables used to append and merge the additional data sets.

Table A2. 6 Indicates which file and variables will be used to merge variables of interest into this base patient list.

Source file	Variable used to match	Variable(s) included	Comments
Base file	n/a	Patient ID Date of death Age at death Place of death Cause of death 1 Cause of death 2	This will be the base to which all other variables will be added
Patient file	patid	Gender Ethnicity Practise ID	Adding gender and ethnicity to the data set and to allow the addition of regional information
Practise file	pracid	Region	To allow appropriate exposures to be assigned to each individual.
Chronic conditions and clinical measurements	patid	All processed chronic disease/clinical	Covariates of interest

(e.g. cvd_observation_file)		measurement variables within two years of date of death	
Drug issue (e.g. hypertension drugs)	patid	All processed drug issue variables within two years of date of death	Covariates of interest
Hospitalisation data	patid	All processed chronic disease/clinical measurement variables within two years of date of death	For sensitivity analysis
Temperature	region	Population weighted daily temperature data	Exposure of interest
AQ data	Region – London only	London daily mean values for PM_{10}, NO_2 and O_3	Potential confounder – sensitivity analysis

Within the merging process of chronic conditions and any clinical measurements, drug issue and HES hospitalisation records will be restricted to a two-year window preceding date of death, in line with the definition of chronic disease being used within this study. Where multiple records are present for a single variable, the most recent record will be used. Table A2.7 below illustrates the format of the complete data for analysis.

Table A2. 7 is an example of how the final master data for analysis will be formatted. The master data for analysis will be saved on the secure server in accordance with LSHTM policy, and according to the CPRD user agreement.

patid	dod	Region	temp	age	gender	efi	diabet	ihd	asthma	…
19726372	20 June 2018	LON	23.6	72	1	2	1	1	0	…
18726351	20 June 2018	LON	23.6	66	0	1	0	0	1	…

26830283	20 June 2018 SE	23.6	45	1	0	1	1	0	…
16203840	20 June 2018 SE	23.6	59	1	0	1	0	0	…
82793610	20 June 2018 SW	23.6	67	1	1	0	1	1	…

Data format requirements for time-stratified case-crossover study design

Transforming data into case-crossover format

In order to analyse the data, timeseries format data needs to be transformed into individual matched case-control format. The process followed to make this data transformation is briefly described below.

First, a *day* variable was generated for each day for each individual within the study population based on the date of death. For example, a death which occurred on 12th June 2019 was given a value of 12. Second, a new identification number was given to each individual for use when identifying groups for analysis (e.g. group 1; group2; group3 etc). Next the *fillin* command was used to create one row for each day of the month in which death occurred for each individual. This resulted in 30 to 31 rows per individual depending on month, after removing the additional days for months which only have 30 days. The *fillin* command autogenerates a new variable (*_fillin*) to indicate where the row has been generated or if it is the original where a 1 indicates a new row and 0 indicates the original. Next a new variable is generated called *casecontrol* using the *if* statement in regard to the *_fillin* variable to within Stata define the original row as case (1) and all new rows as control (0).

To identify the relevant control days of the week of the same month of death, a day of the week (*dow*) variable was then generated for all 30/31 rows per individual. For each individual (*id*) only days of the week (*dow*) were kept that matched the day on which the death occurred. This resulted in 4/5 rows per individual, equating to at least 3 case days per control day (see figure A2.4). Once the data was transformed into the case crossover format as described, exposure variables (average temperature and PM_{10}, NO_2 and 0_3 for London), and their lags were merged so that each row per person had the appropriate exposure variables based on date of death.

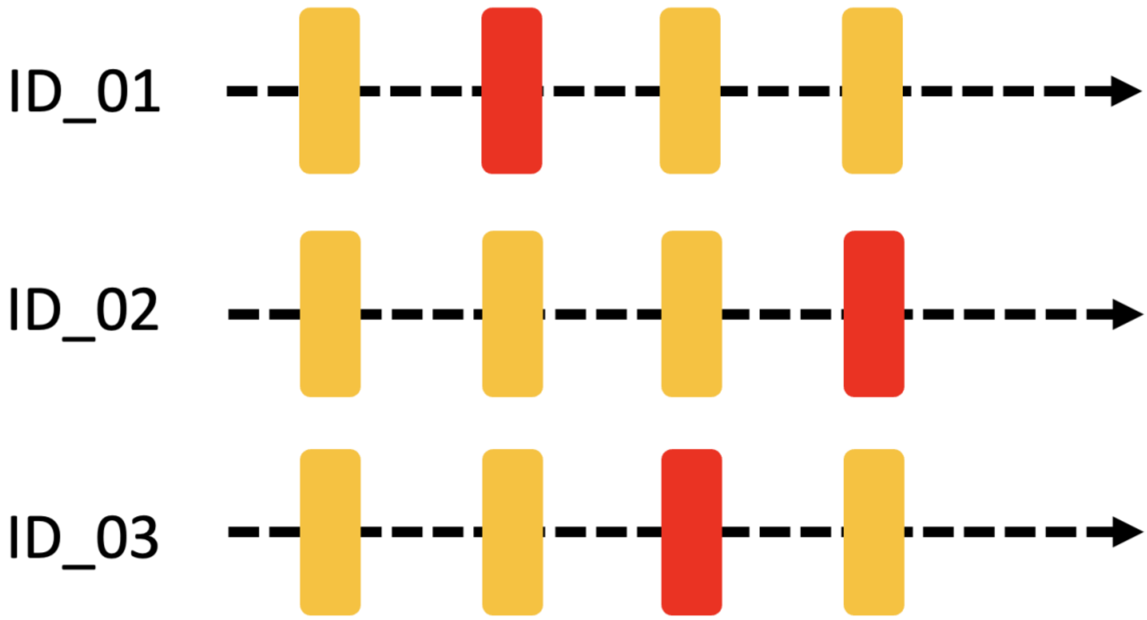


Fig A2. 4 ***depicting the time stratified approach for selecting control periods for use in case-crossover study. The dashed arrow represents the movement of time over one calendar month; the red blocks indicate the case exposure associated with the event of interest; the yellow blocks represent the control periods of exposure on the same day within the same month of the case period; ID_XX indicates one individual within the study (Adapted from Figueiras et al, 2010)***

Data format requirements for Random Forest analysis

No further processing of data is required for the use of random forest classification. One of the strengths of this novel approach is that it is able to deal with relativity unprocessed data. As such, the master data for analysis previously described will be suitable for use.

Appendix 3 – Supplementary materials for paper “*Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records*”

Using individual-level clinical factors and prescribed medicines to identify those at risk of death during heatwaves – a time-stratified case-crossover study using national primary care records – <u>Supplementary materials</u>

Authors:

Ross Thompson[1,2], Sari Kovats[1], Shakoor Hajat[1], Helen Macintyre[3,4], Emer O’Connell[1,5]

Affiliations:

[1] NIHR Health Protection Research Unit in Environmental Change and Health, London School of Hygiene and Tropical Medicine, London

[2]Extreme Events and Health Protection Team, UK Health Security Agency, London

[3]Climate and Health Assessment Team, Centre for Climate and Health Security, UK Health Security Agency

[4]School of Geography, Earth and Environmental Sciences, University of Birmingham

[5] Public Health Unit, Greater London Authority

Contents

Table A3.4 - Results for national and sub-national level unadjusted odds ratios for individuals with pre-existing conditions and prescribed medication with very strong to moderate evidence of increased risk of death during periods of heat using the low impact threshold compared to the minimum mortality temperature and the relative effect modification index (REM).

Table A3.5 – Full results across both low and high impact temperature thresholds

Table A3. 1 - Table outlines each variable included in the analysis, justification for inclusion and information about how each was defined and link to clinical code lists used to extract the data.

Variable for analysis	**Hypothesis to be tested and likely mechanisms**	**Definition of variable**	**Reference for CPRD Aurum code list(s) used to define variables of interest**
Diabetes	Diabetes increases an individual's risk of death during periods of heat, either directly or indirectly. *Likely mechanism*: diabetes is a recognised risk factor for a range of conditions which are also known to increase risk of mortality during heatwaves, e.g. cardiovascular diseases. Therefore, diabetes may indirectly increase risk.	Diagnosis of type 1 and type 2 diabetes within the two years before death. Binary variable (yes or no).	Davidson, J, Warren-Gash, C, Mcdonald, H, Banerjee, A, Smeeth, L, Evans, D and Clay, S. Clinical codelist - Diabetes [Internet]. London School of Hygiene & Tropical Medicine; 2021. Available from: https://doi.org/10.17037/DATA.00002407.
Myocardial infarction	Experiencing a heart attack increases an individual's risk of death during periods of heat *Likely mechanism:* increased strain on heart to maintain blood pressure due to increased viscosity of blood	Diagnosis/record of myocardial infarction (including subsequent MI) within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available.

			To capture all MI cases the following code lists were used in combination: • *Myocardial infarction (ICD-10: I21)* • *Subsequent myocardial infarction (ICD-10: I22)*
Cardiomyopathy	High temperatures increase risk of death for individuals who suffer from conditions that affect the hearts' ability to pump blood. *Likely mechanism:* increased strain on heart to maintain blood pressure due to increased viscosity of blood	Diagnosis/record of cardiomyopathy within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture all cardiomyopathy cases the following code lists will be used in combination: • *Ischaemic cardiomyopathy (ICD-10: I25)* • *Cardiomyopathy (ICD-10: I42*
Cardiac arrest	High temperatures can increase risk of death for those who have previous experienced cardiac arrest *Likely mechanism:* increased strain on heart to maintain blood pressure due to increased viscosity of blood	Diagnosis/record of cardiac arrest within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture cardiac arrest cases the following code lists will be used: • *Cardiac arrest (ICD-10: I46)*
Heart failure	High temperatures increase the risk of mortality for individuals with history of heart failure	Diagnosis/record of heart failure within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ

	Likely mechanism: increased strain on heart to maintain blood pressure due to increased viscosity of blood		Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture heart failure cases the following code lists will be used: • *Heart failure (ICD-10: I50)*
Haemorrhage	High temperatures increase the risk of death for individuals who have previously suffered from bleeding in or around the brain. *Likely mechanism:* blood flow redirected to where heat can escape, increasing potential for burst vessels and bleeding in or around the brain.	Diagnosis/record of haemorrhage within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture all haemorrhage cases the following code lists will be used in combination: • *Subarachnoid haemorrhage (ICD-10: I60)* • *Intracerebral haemorrhage (ICD-10: I61)* • *Other nontraumatic intracranial haemorrhage (ICD-10: I62)*
Stroke	Individuals with a history of stroke are at increased risk of mortality during periods of heat. *Likely mechanism:* disruption in supply of blood/oxygen to brain cells for range of reasons that the	Diagnosis/record of stroke within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available.

	CV system may be affected during heat events.		To capture all stroke cases the following code lists will be used in combination: • *Cerebral infarction (ICD-10: I63)* • *Stroke, not specified as haemorrhage or infarction (ICD-10: I64)*
Other cerebrovascular diseases	Individuals who suffer from blockages or narrowing of blood vessels/arteries are at increased risk of death during periods of high temperature. *Likely mechanism:* disruption in supply of blood/oxygen to brain cells for range of reasons that the CV system may be affected during heat events.	Diagnosis/record of other cerebrovascular diseases within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture all other cerebrovascular disease cases the following code lists will be used in combination: • *Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction (ICD-10: I65)* • *Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction (ICD-10: I66)* • *Other cerebrovascular diseases (ICD-10: 67)* • *Cerebrovascular disorders in diseases classified elsewhere (ICD-10: I68)*
Arrythmia	Individuals who suffer from irregular heart beating are at increased risk of mortality during periods of high temperature. *Likely mechanism:* additional strain on circulatory system may	Diagnosis/record of irregular heart beat within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6

	be too much for heart which is already beating irregularly.		A range cardiovascular disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. To capture all other cerebrovascular disease cases the following code lists will be used in combination: • *Atrial fibrillation and flutter (ICD-10: I48)* • *Other cardiac arrhythmias (ICD-10: I49)*
Emphysema	Individuals suffering from emphysema are at increased risk of mortality during periods of heat. *Likely mechanism:* reduced ability to get additional oxygen to cells required when the body is overheating (e.g. heart muscle working harder to heat loss).	Diagnosis of emphysema within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range respiratory disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. The following code lists will be used: • *Emphysema (ICD-10: J43)*
COPD	Individuals suffering from COPD are at increased risk of mortality during periods of heat. *Likely mechanism:* reduced ability to get additional oxygen to cells required when the body is overheating (e.g. heart muscle working harder to heat loss).	Diagnosis of COPD within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range respiratory disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. The following code lists will be used: • *Other chronic obstructive pulmonary disease (ICD-10: J44)*

Asthma	Individuals who suffer from asthma are at increased risk of death during high temperatures. *Likely mechanism:* reduced ability to get additional oxygen to cells required when the body is overheating (e.g. heart muscle working harder to heat loss).	Diagnosis of asthma within the two years before death. Binary variable (yes or no).	Nealon J, Modin D, Ghosh RE, et al. The feasibility of pragmatic influenza vaccine randomized controlled real-world trials in Denmark and England. NPJ Vaccines 2022;7(1) doi: https://doi.org/10.1038/s41541-022-00444-6 A range respiratory disease CPRD Aurum specific code lists, mapped to ICD-10 categories are available. The following code lists will be used: • *Asthma (ICD-10: J45)*
Severe mental illness	Individuals who suffer from severe mental health disorders are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk; combination of all three.	Definition used includes range of specific conditions, where the diagnosis suggests it is severe (e.g. severe psychosis, moderate and severe schizophrenic episode, severe depression, bipolar – severe manic episode etc). record of the above within two years prior to death. Binary variable (yes or no).	Davidson, J and Strongman, H. Clinical codelist - CPRD Aurum - severe mental illness [Internet]. London School of Hygiene & Tropical Medicine; 2022. Available from: https://doi.org/10.17037/DATA.00002826.
Learning disability	Individuals with learning disabilities have a higher risk of mortality during periods of heat.	Definition includes a wide range of learning disabilities, ranging from mild to severe general learning disabilities, and more specific	Davidson, J, Warren-Gash, C and Cadogan, S. Clinical codelist - learning disabilities [Internet]. London School of Hygiene & Tropical Medicine; 2021. Available from: https://doi.org/10.17037/DATA.00002401.

	Likely mechanism: Unclear, however it may be partly due to inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk in the first place; or a combination of both.	syndromes which are not categorised. Binary variable (yes or no).	
Chronic Kidney disease	Individuals suffering from chronic kidney disease are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* dehydration as a mechanism to serious kidney injury – reduced kidney function due to reduced water contend in blood.	Definition includes a range of terms linked to chronic kidney disease (e.g. all renal failure terms, hemofiltration therapy, glomerulonephritis etc). Binary variable (yes or no)	Davidson, J, Warren-Gash, C, Mcdonald, H, Evans, D and Clay, S. Clinical codelist - chronic kidney disease [Internet]. London School of Hygiene & Tropical Medicine; 2021. Available from: https://doi.org/10.17037/DATA.00002406.
Psychosis	Individuals who suffer from psychosis are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to	Psychosis is quite a broad term, and as such this definition includes a wide range of mental health conditions which include delirium, psychosis etc. Due to the terms included in the list generation, it is not possible to comment on severity of psychosis/psychotic episodes. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

	perceive a risk; combination of all three.		
Anxiety	Individuals who suffer from anxiety are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk; combination of all three.	Defining anxiety (and most mental health diseases) is challenging, due to the nature of primary care data. Terms included within the definition for anxiety were exhaustive and include any mention of anxiety conditions regardless of severity, where anxiety is mentioned. Refusal of questionnaires and family history were removed. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Depression	Individuals who suffer from depression are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk; combination of all three.	Defining depression (and most mental health diseases) is challenging, due to the nature of primary care data. Terms included within the definition for depression were exhaustive and include any mention of depression (excluding bi-polar) conditions regardless of severity. Refusal of questionnaires and family history were also removed. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

Bipolar disorder	Individuals who suffer from bipolar disorder are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive a risk; combination of all three.	Defining bipolar (and most mental health diseases) is challenging, due to the nature of primary care data. Terms included within the definition for bipolar disorder were exhaustive and include any mention of bipolar disorder regardless of severity. Refusal of questionnaires and family history among other routinely used terms not associated with a diagnosis were also removed. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Schizophrenia	Individuals who suffer from schizophrenia are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive risk; combination of all three.	Variable is defined as any terms as recorded within primary care observations which include schizophrenia. From terms included within primary care data it is unclear of severity of symptoms is recorded in primary care data, therefore this will be a binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

Dementia and Alzheimer's disease	Individuals who suffer from Alzheimer's and dementia are at increased risk of mortality during periods of high temperatures. *Likely mechanism:* Unclear, however it may be partly due to medication prescribed to control symptoms; inability of the individual to adapt their own behaviours and or environments; an inability of the individual to perceive risk on hot days; or a combination of all the above.	All terms which indicate dementia and Alzheimer's diagnosis (all types and in association with other diseases). Recording severity of diagnosis unlikely to be well recorded in primary care data. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Parkinson's disease	Individuals who suffer from Parkinson's disease are at increased risk of death during periods of high temperatures. *Likely mechanism:* Unclear, but likely dehydration due to medication taken to control symptoms, as anti-Parkinson's medication are known to have a side effect of dehydration. Link this to Anti-Parkinson's medication variable (number 50)	All terms which indicate Parkinson's disease (including where it is in association with another disease)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Hypothyroidism	Individuals with an underactive thyroid gland are at increased risk	All terms related to hyperthyroidism/overactive thyroid.	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro

	of death during periods of elevated temperatures. *Likely mechanism:* body temperature rises because the basal metabolic rate is raised as there is increased oxygen consumption and the patients hypoactive adrenal function is globally reduced. While heat intolerance is most associated with hyperthyroidism, any thyroid disease, and particularly those related from autoimmune thyroid disfunction, can experience heat intolerance as the body struggles to maintain body temperature.	Binary variable (yes or no)	List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Hyperthyroidism	Individuals with an overactive thyroid gland are at increased risk of death during periods of elevated temperatures. *Likely mechanism:* body temperature rises because the basal metabolic rate is raised as there is increased oxygen consumption and the patients hyperactive adrenal function is globally enhanced. The body will therefore have to work harder to	All terms related to hyperthyroidism/overactive thyroid. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

	lose excess heat potentially increasing strain on other organs sensitive to high temperatures		
Systolic blood pressure	Individuals with high SBP are at increased risk of death during heatwaves *Likely mechanism*: increased strain on heart to maintain blood pressure due to increased viscosity of blood among other potential mechanisms related to CVD.	Terms included of relevance for extracting SBP measurements. Variable will then be classified according to NHS guidance. Categorical variable: • Low = <80 mmHg • Normal = 80 to 120 mmHg • Prehypertension = 120 to 139 mmHg • HT stage 1 = 140 to 159 mmHg HT Stage 2 = 160 mmHg <	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Diastolic blood pressure	Individuals with high DBP are at increased risk of death during heatwaves *Likely mechanism*: increased strain on heart to maintain blood pressure due to increased viscosity of blood among other potential mechanisms related to CVD.	Terms included of relevance for extracting SBP measurements. Variable will then be classified according to NHS guidance. Categorical variable: • Low = <60 mmHg • Normal = 60 to 80 mmHg • PreHT = 80-89 mmHg • HT Stage 1 = 90 to 99 mmHg • HT Stage 2 = 100 mmHg <	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Cardiac Glycosides	Individuals taking cardiac glycosides to treat circulatory	Defined initially by BNF chapter code (020101), then secondly by drug names of those identified within the	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro

	system conditions are at increased risk of death during heatwaves. Likely mechanism: Unclear however the drug causes a more forceful heartbeat, therefore may increase workload of already strained muscle during periods of heat.	BNF chapter. Formulation and concentration of prescribed medication is out of the scope of this project. Binary variable (yes or no)	List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Diuretics	Individuals taking diuretics are at increased risk of death during heatwaves. Likely mechanism: Dehydration, as diuretics are designed to rid the body of sodium and water.	Defined initially by BNF chapter codes (020201, 020202, 020203, 020204, 020205, 020206), the secondly by drug names of those identified within the BNF chapters. Formulation and concentration of prescribed medication is out of the scope of this project. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Beta blockers	Individuals taking beta blockers are at increased risk of mortality during heatwaves. Likely mechanism: interference of thermoregulatory mechanisms, making the heart work harder to get blood to the surface for the body to radiate excess heat.	Defined initially by BNF chapter codes (020400), the secondly by drug names of those identified within the BNF chapter. Formulation and concentration of prescribed medication is out of the scope of this project. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

ACE inhibitors	As above, but specifically for ACE inhibitors. Ace inhibitors have been singled out specifically as they are mentioned within the literature in addition to the wider anti-hypertension drugs.	Defined initially by BNF chapter codes (020505), the secondly by drug names of those identified within the BNF chapter. Formulation and concentration of prescribed medication is out of the scope of this project. NOTE: this variable is ACE inhibitors only Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Vasoconstrictor sympathomimetics	Individuals who are using vasoconstrictor medications to address low blood pressure or severe allergic reactions are at increased risk of death during heatwaves. *Likely mechanism*: Vasoconstrictors increase blood pressure which can increase risk of stroke and other cardiovascular issues.	Defined initially by BNF chapter codes (020700), the secondly by drug names of those identified within the BNF chapter. Formulation and concentration of prescribed medication is out of the scope of this project. Binary variable (yes or no)	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Non-steroidal anti-inflammatory drugs	Individuals taking NSAIDs during heatwaves are at increased risk *Likely mechanism*: NSAIDs are known to interfere with hormones involved in thermoregulation.	Defined initially by BNF chapter codes (100101), then secondly by drug names of those identified within the BNF chapter. Formulation and concentration of prescribed medication is out of the scope of this project.	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

		Binary variable (yes or no)	
Anticholinergic drugs	Individuals taking Anticholinergics during heatwaves are at increased risk *Likely mechanism*: Anticholinergics are known to interfere with thermoregulation and potentially blood pressure	Presence of Anticholinergics as identified within clinical code list (previously available). Formulation and concentration of prescribed medication is out of the scope of this project. Binary variable (yes or no)	Archer L, Koshiaris C, Lay-Flurrie S, et al. Development and external validation of a risk prediction model for falls in patients with an indication for antihypertensive treatment: retrospective cohort study. *BMJ* 2022;379:e070918. doi: 10.1136/bmj-2022-070918 CPRD Aurum specific code list generated for *the STRAtifying Treatments In the multi-morbid Frail elderlY (STRATIFY) study.*

Table A3. 2 - Description of exposure data series in analysis.

Variable	Observations	Proportion	Mean	Std. dev.	Min	Max
Temperature (°C) *(England)*	430,682	100.00%	15.7	3.3	5.7	27.7
London	65,145	26.33%	17.1	3.5	7.7	27.7
The South *(SW & SE)*	149,502	23.83%	15.9	3.1	6.7	25.6
Midlands and East *(WM, EM, EoE)*	102,630	15.13%	15.6	3.3	6.1	26.4
The North *(NE, NW & Y&H)*	113,405	34.71%	14.9	3.0	5.7	24.8
Ozone (ug/m^3) *(London)*	65,145	15.13%	48.9	16.1	11	126.7
NO_2 (ug/m^3) *(London)*	65,145	15.13%	21.9	9.8	3.5	61.4
PM_{10} (ug/m^3) *(London)*	65,145	15.13%	16.7	7.4	6	47.7

NW = Northwest; NE = Northeast; Y&H = Yorkshire and the Humber; WM = West Midlands; EM = East Midlands; EoE = East of England; Lon = London; SE = Southeast; SW = Southwest

Table A3. 3 - Temperature thresholds used in sub-national analysis derived from the temperature mortality relationships.

Sub-national area	MMT = RR=1	Low =RR-1.1
London	17	22
The South *(SW & SE)*	17	21.5
Midlands and East *(WM, EM, EoE)*	17	22
The North *(NE, NW & Y&H)*	16	21.5

NW = Northwest; NE = Northeast; Y&H = Yorkshire and the Humber; WM = West Midlands; EM = East Midlands; EoE = East of England; Lon = London; SE = Southeast; SW = Southwest

Table A3. 4 - Results for national and sub-national level unadjusted odds ratios for individuals with pre-existing conditions and prescribed medication with very strong to moderate evidence of increased risk of death during periods of heat using the low impact threshold compared to the minimum mortality temperature and the relative effect modification index (REM).

Variable	National		London		The South		Mids and East		The North	
	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM
Whole population*	**1.09 (1.08 to 1.11)**	**1.00**	**1.09 (1.07 to 1.11)**	**1.00**	**1.09 (1.07 to 1.11)**	**1.00**	**1.07 (1.03 to 1.11)**	**1.00**	**1.17 (1.09 to 1.25)**	**1.00**
Diabetes	**1.12 (1.07 to 1.17)**	**1.02**	**1.09 (1.02 to 1.16)**	**1.00**	**1.14 (1.07 to 1.21)**	**1.05**	1.11 (0.99 to 1.25)	1.04	1.13 (0.93 to 1.37)	0.97
Heart Failure	**1.11 (1.04 to 1.19)**	**1.02**	**1.12 (1.01 to 1.23**	**1.02**	1.06 (0.96 to 1.16)	0.97	1.10 (0.94 to 1.28)	1.03	**1.41 (1.07 to 1.87)**	**1.21**
Haemorrhage	**1.25 (1.06 to 1.49)**	**1.15**	1.21 (0.95 to 1.55)	1.11	1.05 (0.83 to 1.33)	0.97	1.54 (0.93 to 2.48)	1.42	1.57 (0.78 to 3.16)	1.34
Stroke	**1.14 (1.03 to 1.25)**	**1.04**	1.11 (0.97 to 1.27)	1.02	1.11 (0.98 to 1.26)	1.02	1.16 (0.91 to 1.48)	1.09	1.41 (0.98 to 2.01)	1.20
Arrythmia	**1.09 (1.03 to 1.16)**	**1.00**	1.06 (0.97 to 1.16)	0.97	**1.09 (1.01 to 1.18)**	**1.00**	1.11 (0.97 to 1.28)	1.04	1.15 (0.91 to 1.44)	0.98
Occultation	1.35 (0.90 to 2.04)	1.24	0.99 (0.78 to 1.25)	0.91	1.06 (0.90 to 1.27)	0.98	**1.14 (1.00 to 2.00)**	**1.32**	0.88 (0.48 to 1.62)	0.76
COPD	**1.14 (1.02 to 1.27)**	**1.04**	**1.22 (1.04 to 1.43)**	**1.12**	1.04 (0.90 to 1.21)	0.96	1.18 (0.93 to 1.50)	1.10	0.98 (0.69 to 1.39)	0.84
Asthma	1.11 (0.99 to 1.24)	1.01	**1.17 (1.01 to 1.37)**	**1.08**	0.98 (0.84 to 1.14)	0.90	1.12 (0.86 to 1.45)	1.05	0.81 (0.50 to 1.31)	0.69
Severe Mental Health	**1.21 (1.01 to 1.45)**	**1.11**	1.12 (0.90 to 1.40)	1.03	**1.79 (1.30 to 2.47)**	**1.64**	1.24 (0.73 to 2.11)	1.16	0.35 (0.09 to 1.29)	0.30
Psychosis	**1.17 (1.03 to 1.32)**	**1.07**	1.20 (0.99 to 1.45)	1.10	1.06 (0.91 to 1.23)	0.97	1.23 (0.86 to 1.76)	1.15	1.39 (0.82 to 2.36)	1.19
Depression	**1.25 (1.09 to 1.44)**	**1.15**	**1.24 (1.04 to 1.49)**	**1.14**	1.14 (0.94 to 1.38)	1.05	1.27 (0.88 to 1.83)	1.19	1.36 (0.75 to 2.46)	1.16
Bipolar Disorder	1.03 (0.69 to 1.55)	0.94	0.54 (0.21 to 1.37)	0.49	**3.02 (1.63 to 5.60)**	**2.78**	-	-	-	-
Alzheimer's & Dementia	**1.09 (1.03 to 1.16)**	**1.00**	1.06 (0.98 to 1.15)	0.97	**1.11 (1.03 to 1.19)**	**1.02**	1.11 (0.98 to 1.27)	1.04	1.23 (0.96 to 1.58)	1.05
Parkinson's Disease	**1.22 (1.05 to 1.41)**	**1.11**	1.12 (0.91 to 1.39)	1.03	**1.27 (1.05 to 1.53)**	**1.16**	1.09 (0.74 to 1.60)	1.02	1.74 (0.78 to 3.91)	1.49
Chronic Kidney Disease	1.07 (0.99 to 1.15)	0.98	1.01 (0.89 to 1.09)	0.93	**1.11 (1.01 to 1.22)**	**1.02**	1.09 (0.92 to 1.28)	1.02	1.18 (0.91 to 1.53)	1.01
Cardiac Glycosides	**1.14 (1.03 to 1.26)**	**1.04**	1.07 (0.93 to 1.24)	0.98	**1.14 (1.02 to 1.29)**	**1.05**	**1.26 (1.00 to 1.60)**	**1.18**	1.49 (0.94 to 2.34)	1.27
Diuretics	**1.09 (1.04 to 1.13)**	**0.99**	**1.07 (1.01 to 1.14)**	**0.98**	**1.11 (1.06 to 1.17)**	**1.02**	1.06 (0.96 to 1.16)	0.99	1.16 (0.98 to 1.37)	0.99
Beta Blockers	**1.09 (1.05 to 1.14)**	**1.00**	1.05 (0.99 to 1.12)	0.97	**1.13 (1.07 to 1.19)**	**1.03**	1.08 (0.98 to 1.20)	1.02	**1.27 (1.06 to 1.51)**	**1.08**
Ace Inhibitors	**1.08 (1.04 to 1.12)**	**0.99**	**1.07 (1.01 to 1.13)**	**0.98**	**1.09 (1.03 to 1.15)**	**1.00**	1.05 (0.95 to 1.17)	0.99	**1.22 (1.03 to 1.45)**	**1.05**
Vasoconstrictors	**1.83 (1.19 to 2.80)**	**1.67**	-	-	**1.76 (1.00 to 3.09)**	**1.62**	-	-	-	-
NSAIDs	**1.13 (1.08 to 1.19)**	**1.03**	**1.11 (1.04 to 1.19)**	**1.02**	**1.14 (1.07 to 1.21)**	**1.04**	**1.16 (1.03 to 1.29)**	**1.08**	**1.31 (1.09 to 1.59)**	**1.12**

Bold values indicate estimates with moderate to very strong evidence ($p<0.05$) that individuals with a valid primary care record have an increased odds of death on hot days when a Low impact HHA is likely to be issued by UKHSA

REM is the relative effect modification index which is calculated by dividing the OR of the specific factor of interest by the reference factor, here taken as whole study sample population

Minimum mortality temperatures and low Impact temperature thresholds used in each sub-national areas analysis are provided in table S3 in supplemental materials

Table A3. 5 - OR estimates and 95% CI and p-values for all clinical individual-level risk factors investigated, for the whole population using both the Low Impact threshold (temperature associated with RR of 1.1) and Medium Impact threshold (RR of 1.2). In addition, the Relative Effect Modification index (REM) us also reported for each variable investigated.

Variable	Low Impact threshold			Medium Impact threshold		
	OR (95% CI)	p-value	REM	OR (95% CI)	p-value	REM
Whole population	1.09 (1.08 to 1.11)	<0.001	n/a	1.20 (1.16 to 1.25)	<0.001	n/a
Diastolic blood pressure						
Low DPB	1.07 (1.00 to 1.15)	0.053	0.96	1.12 (0.94 to 1.33)	0.195	0.91
Normal DBP	1.11 (1.08 to 1.14)	<0.001	1.00	1.23 (1.15 to 1.32)	<0.001	1.00
Prehypertensive	1.15 (1.1 0to 1.21)	<0.001	1.04	1.35 (1.21 to 1.51)	<0.001	1.10
Hypertension Stage 1	1.06 (0.96 to 1.16)	0.270	0.95	1.09 (0.87 to 1.38)	0.444	0.89
Hypertension Stage 2	1.09 (0.91 to 1.31)	0.349	0.98	1.07 (0.68 to 1.69)	0.766	0.87
Hypertension (all)	1.06 (0.98 to 1.15)	0.166	0.96	1.08 (0.88 to 1.33)	0.458	0.88
Systolic blood pressure						
Low SPB	1.21 (0.90 to 1.63)	0.208	1.07	1.20 (0.55 to 2.59)	0.652	0.92
Normal SBP	1.13 (1.09 to 1.17)	<0.001	1.00	1.30 (1.30 to 1.43)	<0.001	1.00
Prehypertensive	1.11 (1.03 to 3.16)	<0.001	0.98	1.24 (1.14 to 1.34)	<0.001	0.95
Hypertension Stage 1	1.08 (1.03 to 0.13)	0.003	0.95	1.14 (1.01 to 1.29)	0.035	0.88
Hypertension Stage 2	1.08 (0.97 to 1.20)	0.169	0.95	1.15 (0.87 to 1.52)	0.318	0.88
Hypertension (all)	1.08 (1.03 to 1.13)	0.001	0.95	1.14 (1.02 to 1.28)	0.021	0.88
Chronic conditions						
Diabetes	1.12 (1.07 to 1.17)	<0.001	1.02	1.25 (1.12 to 1.40)	<0.001	1.04
Myocardial infarction	1.04 (0.91 to 1.19)	0.593	0.95	1.15 (0.82 to 1.62)	0.412	0.96
Cardiomyopathy	0.87 (0.63 to 1.21)	0.414	0.80	0.65 (0.29 to 1.47)	0.302	0.54
Cardiac Arrest	1.01 (0.79 to 1.29)	0.919	0.93	0.84 (0.44 to 1.60)	0.597	0.70
Heart Failure	1.11 (1.04 to 1.19)	0.003	1.02	1.24 (1.04 to 1.47)	0.015	1.03
Haemorrhage	1.25 (1.06 to 1.49)	0.010	1.15	1.65 (1.08 to 2.54)	0.021	1.38
Stroke	1.14 (1.03 to 1.25)	0.010	1.04	1.20 (0.93 to 1.54)	0.166	0.99
Other CVD	0.96 (0.72 to 1.29)	0.809	0.88	0.71 (0.31 to 1.64)	0.427	0.59
Arrythmia	1.09 (1.03 to 1.16)	0.004	1.00	1.17 (1.00 to 1.36)	0.043	0.97
Occultation	1.35 (0.90 to 2.04)	0.147	1.24	1.13 (0.97 to 1.31)	0.118	0.94
Emphysema	1.00 (0.76 to 1.31)	0.983	0.91	1.25 (0.66 to 2.35)	0.492	1.04
CPOD	1.14 (1.02 to 1.27)	0.018	1.04	1.28 (0.97 to 1.69)	0.077	1.07
Asthma	1.11 (0.99 to 1.24)	0.070	1.01	1.13 (0.87 to 1.47)	0.353	0.94
Severe Mental Health	1.21 (1.01 to 1.45)	0.041	1.11	1.15 (0.73 to 1.81)	0.541	0.96
Learning disability	0.85 (0.62 to 1.17)	0.312	0.78	0.31 (0.10 to 1.03)	0.055	0.26
Psychosis	1.17 (1.03 to 1.32)	0.014	1.07	1.46 (1.08 to 1.98)	0.015	1.21

Anxiety	1.08 (0.97 to 1.20)	0.155	0.99	1.30 (0.98 to 1.73)	0.073	1.08
Depression	1.25 (1.09 to 1.44)	0.001	1.15	1.25 (0.89 to 1.75)	0.203	1.04
Bipolar Disorder	1.03 (0.69 to 1.55)	0.877	0.94	0.81 (0.21 to 3.17)	0.766	0.68
Schizophrenia	1.18 (0.91 to 1.54)	0.208	1.08	1.22 (0.64 to 2.34)	0.540	1.02
Alzheimer's & Dementia	1.09 (1.03 to 1.16)	0.002	1.00	1.13 (0.97 to 1.30)	0.113	0.94
Parkinson's Disease	1.22 (1.05 to 1.41)	0.008	1.11	1.46 (1.02 to 2.08)	0.040	1.21
Hypothyroidism	0.99 (0.87 to 1.11)	0.812	0.90	1.09 (0.80 to 1.48)	0.575	0.91
Hyperthyroidism	1.01 (0.68 to 1.49)	0.965	0.92	1.69 (0.60 to 4.74)	0.319	1.41
Chronic Kidney Disease	1.07 (0.99 to 1.15)	0.071	0.98	1.08 (0.91 to 1.28)	0.387	0.90
Prescribed medications						
Cardiac Glycosides	1.14 (1.03 to 1.26)	0.011	1.04	1.14 (0.87 to 1.50)	0.348	0.95
Diuretics	1.09 (1.04 to 1.13)	<0.001	0.99	1.19 (1.07 to 1.32)	0.001	0.99
Beta Blockers	1.09 (1.05 to 1.14)	<0.001	1.00	1.21 (1.09 to 1.34)	<0.001	1.00
Ace Inhibitors	1.08 (1.04 to 1.12)	<0.001	0.99	1.17 (1.06 to 1.30)	0.002	0.97
Vasoconstrictors	1.83 (1.19 to 2.80)	0.006	1.67	6.94 (1.93 to 24.98)	0.003	5.77
NSAIDs	1.13 (1.08 to 1.19)	<0.001	1.03	1.27 (1.13 to 1.43)	<0.001	1.06
Anticholinergic drugs	1.07 (0.94 to 1.21)	0.316	0.97	0.94 (0.68 to 1.31)	0.723	0.78

Append 4 – Supplemental materials for paper "*Social determinants of heat related mortality in England – a time-stratified case-crossover study using primary care records*"

Social determinants of heat related mortality in England – a time-stratified case-crossover study using primary care records – supplemental materials

Authors:

Ross Thompson[1,2], Sari Kovats[1], Shakoor Hajat[1], Emer O'Connell[1,3]

Affiliations:

[1] NIHR Health Protection Research Unit in Environmental Change and Health, London School of Hygiene and Tropical Medicine, London

[2]Extreme Events and Health Protection Team, UK Health Security Agency, London

[3] Public Health Unit, Greater London Authority

Contents

Table A4. 1 - Table outlines each variable included in the analysis, justification for inclusion and information about how each was defined and link to clinical code lists used to extract the data.

Variable for analysis	Hypothesis to be tested and likely mechanisms	Definition of variable	Reference for CPRD Aurum code list(s) used to define variables of interest
Age	Risk of death on a hot day increases with age. *Likely mechanism:* linked to reduced thermoregulatory responses; increased likelihood of comorbidities and polypharmacy etc	Individuals categorised in the following groups • <65 • 65+ • 45-65 • 65-75 • 75-85 • 85+	Individuals age at death provided within CPRD data.
Sex	There is no difference in risk of death by sex. *Likely mechanism:* Unclear, but potential for physiological, social and contextual influences	Categorised as male or female.	Individuals sex provided within CPRD data.
Ethnicity	Risk of death during a heat episode differs by ethnic group. *Likely mechanism:* Unclear but if a difference is observed, it may be more to do with social drivers of inequality rather than a physiological difference.	Due to low number of records only broad categories are possible. • White • Asian • Black • Other ethnic groups	Mathur, R (2021). Risk factor codelist - Ethnicity. [Data Collection]. London School of Hygiene & Tropical Medicine, London, United Kingdom. https://doi.org/10.17037/DATA.00002414.

	There is little evidence of this in the UK.		
Marital Status	Individuals who are unmarried/widowed/divorced are at higher risk of death during heatwaves. *Likely mechanism:* This is likely to be a proxy for more social vulnerabilities which increased risk of an individual, such as social isolation, but has been identified as a risk factor in Italy.	All terms related to marital status have been included, searching specifically for divorced, widow partner and married. Categorical variable • Unmarried/divorced/widowed • Married	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Living arrangement	People living alone or who are homeless are at increased risk of death during heatwaves. *Likely mechanism*: several potential mechanisms to consider depending on which group are considered. Homeless individuals are potentially at increased risk due to a range of social vulnerabilities and other underlying issues which may contribute to those individuals living arrangements. Individuals living along may also lack social connections which increase risk during heatwaves.	All terms related to living arrangements have been included. E.g. information recorded about partner, living in temporary accommodation/hostel etc. Categorical variable: • Homeless • Living alone • Co-habitation	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Electronic frailty index	1) Individuals who are assessed as being mildly, moderately, and severely frail during routine frailty assessments are also at increased risk of death during heatwaves.	eFI was not calculated for each individual within the study, but they were included and used where a relevant record, therefore this variable is restricted to over 65s only. Some individuals were already classed as fit, mildly frail, moderately frail or severely	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists

	2) Heat risk increases with increasing eFI category *Likely mechanism*: frailty assessment is a ratio of the number of "deficits" present in a patients records, and is a general measure of vulnerability. Therefore it is plausible that eFI may also indicate heat risk	frail, while others only had an eFI value. These were categorised and combined into one group as outlined below. Categorical variable: • Fit - 0-0.12 • Mildly Frail = 0.12 to 0.24 • Moderately Frail = 0.24 to 0.36 • Severely Frail = 0.36+	
Alcohol abuse and misuse	Risk of death on hot days are higher for individuals who drink a lot compared to individuals who do not. *Likely mechanism*: dehydration; for heavy drinkers lack of ability to adapt ones own behaviour or environment	Within GP records either the amount of alcohol consumed regularly is recorded, or an individual is categorised as a non-drinker, light drinker, moderate drinker or heavy drinker. Previous studies suggest that this categorisation is complete enough within CPRD to justify using this information in analysis. Categorical data • non-drinker • Light drinker • Moderate drinker • Heavy drinker	Bespoke code list generated for this project by Daniel Omoyeni and Ross Thompson, clinically validated by Luis Baptista Mieiro List available on GitHub here: https://github.com/Rossdud/Clinical-code-lists
Body mass index (BMI)	Risk of death is higher for individuals with higher BMI during periods of heat. *Likely mechanism:* due to higher body mass, the individuals organs are likely to need to work harder (higher strain on the heart for example) than for those with lower body mass	Code list identifies all terms within Aurum to extract either direct measurements of BMI recorded in patient clinical records or extract weight and height measurements to allow BMI to be easily calculated. BMI value measurements for individuals have been categorised in line with NHS guidelines.	Forbes, H and Carreira, H (2021). Risk factor codelist - Body Mass Index (BMI). [Data Collection]. London School of Hygiene & Tropical Medicine, London, United Kingdom. https://doi.org/10.17037/DATA.00002413.

		- Underweight - <18.5 kg/m2 - Healthy 18.5 to 24.9 kg/m2 - Overweight - 25 to 29.9 kg/m2 - Obese 1 - 30 to 34.9 kg/m2 - Obese 2 - 35 39.9 kg/m2 - Obese 3 - 40-80 kg/m2	
Deprivation	Risk of death on a hot day is higher for those considered most deprived *Likely mechanism*: Unclear, but likely related to all domains of deprivation (income, employment, education, health, crime, barriers to housing and services and living environment)	Index of Multiple Deprivation (IMD) Deciles: 1 = least deprived 10 = most deprived	IMD provided by CPRD based on the individuals GP practise address

Table A4. 2 - Description of exposure data series in analysis.

Variable	Observations	Proportion	Mean	Std. dev.	Min	Max
Temperature (°C) *(England)*	430,682	100.00%	15.7	3.3	5.7	27.7
London	65,145	26.33%	17.1	3.5	7.7	27.7
The South *(SW & SE)*	149,502	23.83%	15.9	3.1	6.7	25.6
Midlands and East *(WM, EM, EoE)*	102,630	15.13%	15.6	3.3	6.1	26.4
The North *(NE, NW & Y&H)*	113,405	34.71%	14.9	3.0	5.7	24.8
Ozone (ug/m^3) *(London)*	65,145	15.13%	48.9	16.1	11	126.7
NO_2 (ug/m^3) *(London)*	65,145	15.13%	21.9	9.8	3.5	61.4
PM_{10} (ug/m^3) *(London)*	65,145	15.13%	16.7	7.4	6	47.7

NW = Northwest; NE = Northeast; Y&H = Yorkshire and the Humber; WM = West Midlands; EM = East Midlands; EoE = East of England; Lon = London; SE = Southeast; SW = Southwest

Table A4. 3 - Temperature thresholds used in sub-national analysis derived from the temperature mortality relationships.

Sub-national area	MMT = RR=1	Low =RR-1.1
London	17	22
The South *(SW & SE)*	17	21.5
Midlands and East *(WM, EM, EoE)*	17	22
The North *(NE, NW & Y&H)*	16	21.5

NW = Northwest; NE = Northeast; Y&H = Yorkshire and the Humber; WM = West Midlands; EM = East Midlands; EoE = East of England; Lon = London; SE = Southeast; SW = Southwest

Table A4. 4 - OR estimates and 95% CI and p-values for all clinical individual-level risk factors investigated, for the whole population using both the Low Impact threshold (temperature associated with RR of 1.1) and Medium Impact threshold (RR of 1.2). In addition, the Relative Effect Modification index (REM) us also reported for each variable investigated.

Variable	Low Impact Threshold			Medium Impact Threshold		
	OR (95%CI)	p-value	REM	OR (95%CI)	p-value	REM

All people	1.09 (1.08 to 1.11)	<0.001	1.00	1.2 (1.16 to 1.25)	<0.001	1.00
Age						
0-65	1.07 (1.03 to 1.1)	<0.001	0.98	1.14 (1.05 to 1.24)	0.001	0.95
65+	1.1 (1.08 to 1.12)	<0.001	1.01	1.22 (1.17 to 1.27)	<0.001	1.01
45-65	1.05 (1.01 to 1.1)	0.007	0.96	1.13 (1.04 to 1.24)	0.006	0.94
65-75	1.09 (1.05 to 1.14)	<0.001	1.00	1.24 (1.13 to 1.36)	<0.001	1.03
75-85	1.08 (1.05 to 1.11)	<0.001	0.99	1.17 (1.09 to 1.26)	<0.001	0.97
85+	1.1 (1.08 to 1.12)	<0.001	1.01	1.21 (1.16 to 1.27)	<0.001	1.01
Sex						
Male	1.08 (1.06 to 1.1)	<0.001	1.00	1.18 (1.12 to 1.24)	<0.001	1.00
Female	1.1 (1.08 to 1.13)	<0.001	1.02	1.23 (1.17 to 1.29)	<0.001	1.04
Ethnicity						
White	1.14 (1.08 to 1.2)	<0.001	1.00	1.42 (1.26 to 1.61)	<0.001	1.00
Black	1.44 (1.18 to 1.76)	<0.001	1.27	1.67 (1.1 to 2.53)	0.017	1.17
Asian	1.25 (1.02 to 1.52)	0.030	1.10	1.64 (1.02 to 2.63)	0.042	1.15
Other ethnicities	0.96 (0.69 to 1.34)	0.827	0.85	0.62 (0.25 to 1.53)	0.299	0.44
Marital status						
Single/ Divorced/ Widowed	1.17 (1.03 to 1.34)	0.020	1.00	1.46 (1.06 to 2.01)	0.022	1.00
Married/ Has Partner	1.16 (1.05 to 1.27)	0.003	0.99	1.35 (1.07 to 1.71)	0.011	0.93
Living arrangement						
Living alone	1.17 (1.07 to 1.29)	0.001	1.00	1.29 (1.04 to 1.62)	0.023	1.00
Cohabiting	1.05 (0.92 to 1.19)	0.467	0.89	1.07 (0.8 to 1.44)	0.647	0.83
Homeless	0.83 (0.48 to 1.45)	0.509	0.71	0.51 (0.06 to 4.15)	0.525	0.39
Electronic Frailty Index (eFI)						
Fit	1.13 (0.97 to 1.31)	0.128	1.00	1.3 (0.92 to 1.85)	0.142	1.00
Mildly Frail	1.11 (1.03 to 1.2)	0.007	0.99	1.24 (1.04 to 1.47)	0.017	0.95
Moderately Frail	1.13 (1.07 to 1.2)	<0.001	1.01	1.37 (1.2 to 1.56)	<0.001	1.05
Severely Frail	1.13 (1.07 to 1.19)	<0.001	1.00	1.25 (1.11 to 1.42)	<0.001	0.96
Alcohol Intake						
Non-Drinker	1.05 (0.9 to 1.22)	0.518	1.00	1.18 (0.84 to 1.64)	0.337	1.00
Light Drinker	1.02 (0.87 to 1.21)	0.803	0.97	1 (0.66 to 1.5)	0.993	0.85
Moderate Drinker	1.14 (1.1 to 1.18)	<0.001	1.08	1.3 (1.19 to 1.42)	<0.001	1.10
Heavy Drinker	1.2 (1.02 to 1.41)	0.032	1.14	1.5 (0.99 to 2.26)	0.055	1.27
BMI						
Underweight	1.13 (1.04 to 1.22)	0.003	1.05	1.28 (1.07 to 1.54)	0.008	1.15
Normal weight	1.07 (1.03 to 1.12)	0.002	1.00	1.11 (0.99 to 1.24)	0.064	1.00
Overweight	1.09 (1.04 to 1.15)	0.001	1.02	1.21 (1.06 to 1.37)	0.005	1.08
Obese 1	1.14 (1.06 to 1.27)	0.001	1.07	1.51 (1.31 to 1.75)	<0.001	1.36
Obese 2	1.32 (1.17 to 1.5)	<0.001	1.23	1.35 (1.12 to 1.62)	0.002	1.21
Obese 3	1.19 (1.02 to 1.39)	0.023	1.11	1.99 (1.47 to 2.71)	<0.001	1.79
Obese (all)	1.19 (1.12 to 1.27)	<0.001	1.11	1.5 (1.01 to 2.22)	0.046	1.35
Index of Multiple Deprivation (IMD)						

1 (Least Deprived)	1.05 (1 to 1.11)	0.048	1.00	1.19 (1.03 to 1.37)	0.018	1.00
2	1.03 (0.98 to 1.08)	0.221	0.98	1.02 (0.9 to 1.16)	0.733	0.86
3	1.12 (1.06 to 1.17)	<0.001	1.06	1.26 (1.12 to 1.43)	<0.001	1.06
4	1.1 (1.04 to 1.15)	<0.001	1.04	1.17 (1.04 to 1.32)	0.01	0.98
5	1.07 (1.02 to 1.12)	0.008	1.01	1.18 (1.05 to 1.32)	0.004	0.99
6	1.09 (1.04 to 1.14)	<0.001	1.03	1.16 (1.04 to 1.3)	0.008	0.98
7	1.08 (1.03 to 1.12)	0.002	1.02	1.2 (1.08 to 1.33)	0.001	1.01
8	1.09 (1.05 to 1.14)	<0.001	1.04	1.17 (1.07 to 1.28)	0.001	0.98
9	1.17 (1.12 to 1.22)	<0.001	1.11	1.33 (1.21 to 1.47)	<0.001	1.12
10 (Most Deprived)	1.19 (1.11 to 1.27)	<0.001	1.13	1.53 (1.28 to 1.82)	<0.001	1.29

Table A4. 5 - National and sub-national level results for individual level factors with very strong to moderate evidence of increased risk of death during periods of heat equating to at least the low impact threshold and the relative effect modification index (REM).

Variable	National		London		The South		Mids and East		The North	
	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM	OR (95% CI)	REM
All people*	**1.09 (1.08 to 1.11)**	**1.00**	**1.09 (1.07 to 1.11)**	**1.00**	**1.09 (1.07 to 1.11)**	**1.00**	**1.07 (1.03 to 1.11)**	**1.00**	**1.17 (1.09 to 1.25)**	**1.00**
0-65	**1.07 (1.03 to 1.10)**	**1.00**	**1.09 (1.04 to 1.13)**	**1.00**	0.98 (0.93 to 1.04)	0.90	**1.12 (1.01 to 1.24)**	**1.00**	0.99 (0.85 to 1.16)	0.85
65+	**1.10 (1.08 to 1.12)**	**1.03**	**1.09 (1.07 to 1.11)**	**1.00**	**1.11 (1.08 to 1.13)**	**1.13**	**1.06 (1.01 to 1.11)**	**0.95**	**1.21 (1.12 to 1.30)**	**1.03**
45-65	**1.05 (1.01 to 1.10)**	**1.00**	**1.09 (1.04 to 1.14)**	**1.00**	0.97 (0.91 to 1.03)	1.00	1.08 (0.96 to 1.21)	1.00	0.98 (0.83 to 1.17)	0.84
65-75	**1.09 (1.05 to 1.14)**	**1.04**	**1.09 (1.04 to 1.14)**	**1.00**	**1.08 (1.02 to 1.14)**	**1.11**	**1.13 (1.02 to 1.25)**	**1.05**	1.14 (0.96 to 1.34)	0.97
75-85	**1.08 (1.05 to 1.11)**	**1.02**	**1.08 (1.04 to 1.12)**	**0.99**	**1.07 (1.03 to 1.12)**	**1.11**	1.03 (0.96 to 1.11)	0.96	**1.26 (1.11 to 1.42)**	**1.08**
85+	**1.10 (1.08 to 1.12)**	**1.04**	**1.09 (1.06 to 1.12)**	**1.00**	**1.11 (1.09 to 1.14)**	**1.15**	**1.04 (1.00 to 1.10)**	**0.97**	**1.23 (1.13 to 1.34)**	**1.05**
Sex										
*Male**	**1.08 (1.06 to 1.10)**	**1.00**	**1.08 (1.05 to 1.12)**	**1.00**	**1.07 (1.04 to 1.10)**	**1.00**	1.05 (0.99 to 1.11)	1.00	**1.12 (1.01 to 1.23)**	**0.96**
Female	**1.10 (1.08 to 1.13)**	**1.02**	**1.09 (1.06 to 1.13)**	**1.01**	**1.11 (1.08 to 1.14)**	**1.04**	**1.08 (1.03 to 1.14)**	**1.03**	**1.22 (1.11 to 1.34)**	**1.04**
Ethnicity										
*White**	**1.14 (1.08 to 1.20)**	**1.00**	**1.14 (1.06 to 1.22)**	**1.00**	-		-		-	
Black	**1.44 (1.18 to 1.76)**	**1.27**	**1.45 (1.19 to 1.78)**	**1.28**	-		-		-	
Asian	**1.25 (1.02 to 1.52)**	**1.10**	**1.25 (1.01 to 1.55)**	**1.10**	-		-		-	
Other	0.96 (0.69 to 1.34)	0.85	0.84 (0.57 to 1.25)	0.74	-		-		-	
IMD (deciles)										
1 (least)*	**1.05 (1.00 to 1.11)**	**1.00**	0.99 (0.88 to 1.10)	1.00	**1.08 (1.03 to 1.14)**	**1.00**	1.11 (0.99 to 1.24)	1.00	1.17 (0.91 to 1.49)	1.00
2	1.03 (0.98 to 1.08)	0.98	1.06 (0.98 to 1.14)	1.07	**1.08 (1.02 to 1.14)**	**0.99**	0.96 (0.85 to 1.09)	0.87	1.18 (0.91 to 1.51)	1.01
3	**1.12 (1.06 to 1.17)**	**1.06**	**1.09 (1.02 to 1.17)**	**1.11**	1.05 (0.99 to 1.12)	0.97	**1.19 (1.06 to 1.33)**	**1.07**	**1.28 (1.02 to 1.61)**	**1.10**
4	**1.10 (1.04 to 1.15)**	**1.04**	**1.08 (1.01 to 1.16)**	**1.10**	**1.11 (1.05 to 1.19)**	**1.03**	1.03 (0.92 to 1.16)	0.93	**1.53 (1.23 to 1.89)**	**1.31**
5	**1.07 (1.02 to 1.12)**	**1.01**	1.04 (0.97 to 1.10)	1.05	**1.10 (1.02 to 1.17)**	**1.01**	1.04 (0.92 to 1.17)	0.94	1.02 (0.80 to 1.31)	0.88
6	**1.09 (1.04 to 1.14)**	**1.03**	**1.09 (1.02 to 1.15)**	**1.10**	**1.10 (1.02 to 1.17)**	**1.01**	1.07 (0.94 to 1.20)	0.96	1.12 (0.87 to 1.44)	0.96
7	**1.08 (1.03 to 1.12)**	**1.02**	**1.07 (1.02 to 1.13)**	**1.09**	**1.10 (1.03 to 1.18)**	**1.02**	0.94 (0.82 to 1.08)	0.85	**1.35 (1.07 to 1.71)**	**1.16**
8	**1.09 (1.05 to 1.14)**	**1.04**	**1.09 (1.04 to 1.14)**	**1.10**	**1.15 (1.05 to 1.25)**	**1.06**	1.07 (0.92 to 1.24)	0.97	0.99 (0.80 to 1.24)	0.85
9	**1.17 (1.11 to 1.22)**	**1.11**	**1.17 (1.11 to 1.23)**	**1.19**	**1.13 (1.04 to 1.23)**	**1.04**	1.07 (0.92 to 1.24)	0.96	**1.30 (1.06 to 1.58)**	**1.11**
10 (most)	**1.19 (1.11 to 1.27)**	**1.13**	**1.20 (1.09 to 1.33)**	**1.22**	1.02 (0.91 to 1.14)	0.94	**1.18 (1.02 to 1.35)**	**1.06**	1.03 (0.89 to 1.20)	0.89
Marital Status										
*Single/ Divorced/ Widowed**	**1.17 (1.03 to 1.34)**	**1.00**	**1.22 (1.03 to 1.44)**	**1.00**	1.10 (0.92 to 1.32)	1.00	0.95 (0.61 to 1.48)	1.02	1.23 (0.70 to 2.16)	0.92
Married/ Has Partner	**1.16 (1.05 to 1.27)**	**0.99**	**1.15 (1.02 to 1.30)**	**0.94**	**1.14 (1.01 to 1.28)**	**1.03**	0.93 (0.69 to 1.26)	1.00	1.35 (0.82 to 2.20)	1.00
Living Arrangement										

*Living alone**	**1.17 (1.07 to 1.29)**	**1.00**	**1.13 (1.00 to 1.28)**	**1.00**	**1.20 (1.06 to 1.36)**	**1.00**	1.05 (0.82 to 1.35)	1.00	**1.43 (1.01 to 2.02)**	**1.00**
Cohabiting	1.05 (0.92 to 1.19)	0.89	1.07 (0.92 to 1.24)	0.94	0.82 (0.63 to 1.05)	0.68	1.08 (0.70 to 1.68)	1.03	1.17 (0.60 to 2.26)	0.82
Homeless	0.83 (0.48 to 1.45)	0.71	0.93 (0.53 to 1.64)	0.82	-	-	-	-	-	-
BMI Category										
Underweight	**1.13 (1.04 to 1.22)**	**1.05**	1.09 (0.98 to 1.21)	1.01	1.09 (0.97 to 1.23)	0.98	**1.23 (1.02 to 1.48)**	**1.26**	1.38 (0.99 to 1.94)	1.38
*Normal weight**	**1.07 (1.03 to 1.12)**	**1.00**	**1.08 (1.02 to 1.15)**	**1.00**	**1.11 (1.05 to 1.18)**	**1.00**	0.97 (0.97 to 1.09)	1.00	1.01 (0.84 to 1.21)	1.00
Overweight	**1.09 (1.04 to 1.15)**	**1.02**	**1.09 (1.01 to 1.17)**	**1.01**	**1.09 (1.01 to 1.17)**	**0.98**	1.05 (0.92 to 1.21)	1.08	1.11 (0.90 to 1.38)	1.11
Obese 1	**1.14 (1.06 to 1.27)**	**1.07**	**1.14 (1.02 to 1.27)**	**1.05**	1.02 (0.91 to 1.14)	0.92	**1.35 (1.14 to 1.61)**	**1.39**	1.25 (0.92 to 1.71)	1.24
Obese 2	**1.32 (1.17 to 1.50)**	**1.23**	**1.31 (1.10 to 1.57)**	**1.21**	**1.17 (1.00 to 1.37)**	**1.05**	1.26 (0.93 to 1.71)	1.30	**1.79 (1.07 to 3.00)**	**1.78**
Obese 3	**1.19 (1.02 to 1.39)**	**1.11**	**1.30 (1.06 to 1.59)**	**1.20**	1.10 (0.87 to 1.40)	0.99	0.97 (0.66 to 1.44)	1.00	0.74 (0.35 to 1.53)	0.73
Obese (all)	**1.19 (1.12 to 1.27)**	**1.11**	**1.21 (1.11 to 1.31)**	**1.12**	1.07 (0.98 to 1.16)	0.96	**1.26 (1.10 to 1.45)**	**1.29**	1.26 (0.99 to 1.62)	1.26

Bold values indicate estimates with moderate to very strong evidence (p<0.05) that individuals with a valid primary care record have an increased odds of death on hot days when a Low impact HHA is likely to be issued by UKHSA

* Indicates the reference group for calculation of the REM index value

A dash (-) indicates where numbers were too small to provide an OR estimate for specific sub-groups

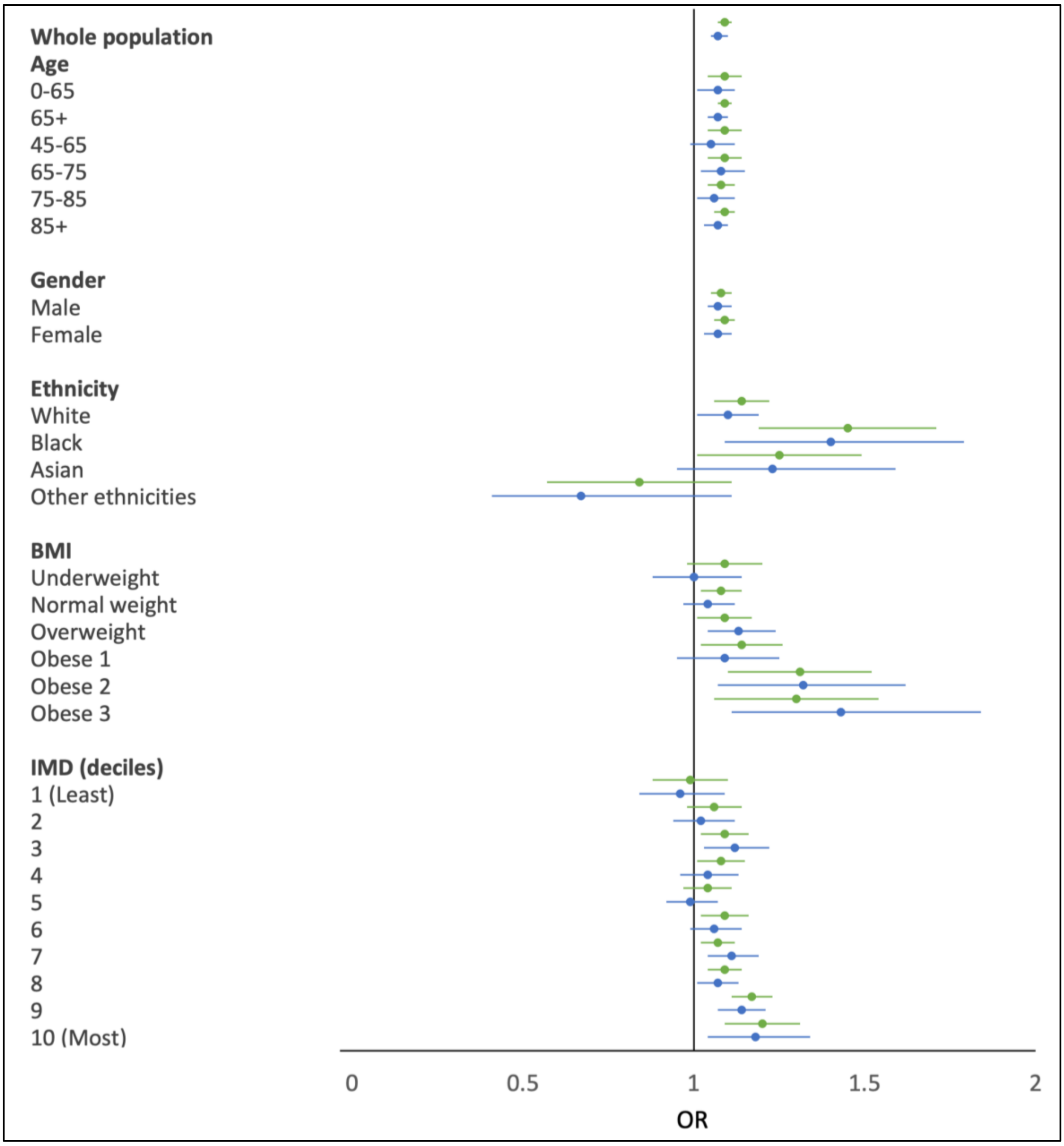


Fig A4. 1 ***Forest plot of OR estimates and 95%CIs by age, sex, ethnicity, BMI and IMD categories in London. Green points represent the crude estimates and blue points represent the adjusted estimates for mean daily concentrations of O_3, PM_{10} and NO_2.***