



# Comparison of various heat waves definitions and the burden of heat-related mortality in France: implications for existing early warning systems

Fleur Serge Kanti, Anna Alari, Basile Chaix, Tarik Benmarhnia

## ► To cite this version:

Fleur Serge Kanti, Anna Alari, Basile Chaix, Tarik Benmarhnia. Comparison of various heat waves definitions and the burden of heat-related mortality in France: implications for existing early warning systems. *Environmental Research*, 2022, 215, pp.114359. 10.1016/j.envres.2022.114359 . hal-03973107

**HAL Id: hal-03973107**

**<https://hal.science/hal-03973107>**

Submitted on 3 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Comparison of various heat waves definitions and the burden of heat-related mortality in France: implications for existing early warning systems.**

Fleur Serge Kanti<sup>a1</sup>

Anna Alari<sup>a</sup>

Basile Chaix<sup>a</sup>

Tarik Benmarhnia<sup>b</sup>

- a. Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, IPLESP, Nemesis team, Faculté de Médecine Saint-Antoine, 27 rue Chaligny, 75012 Paris, France
- b. Scripps Institution of Oceanography University of California, San Diego, La Jolla, San Diego, CA, USA

<sup>1</sup>Corresponding author:

Fleur Serge Kanti

E-mail: [sergekanti@gmail.com](mailto:sergekanti@gmail.com)

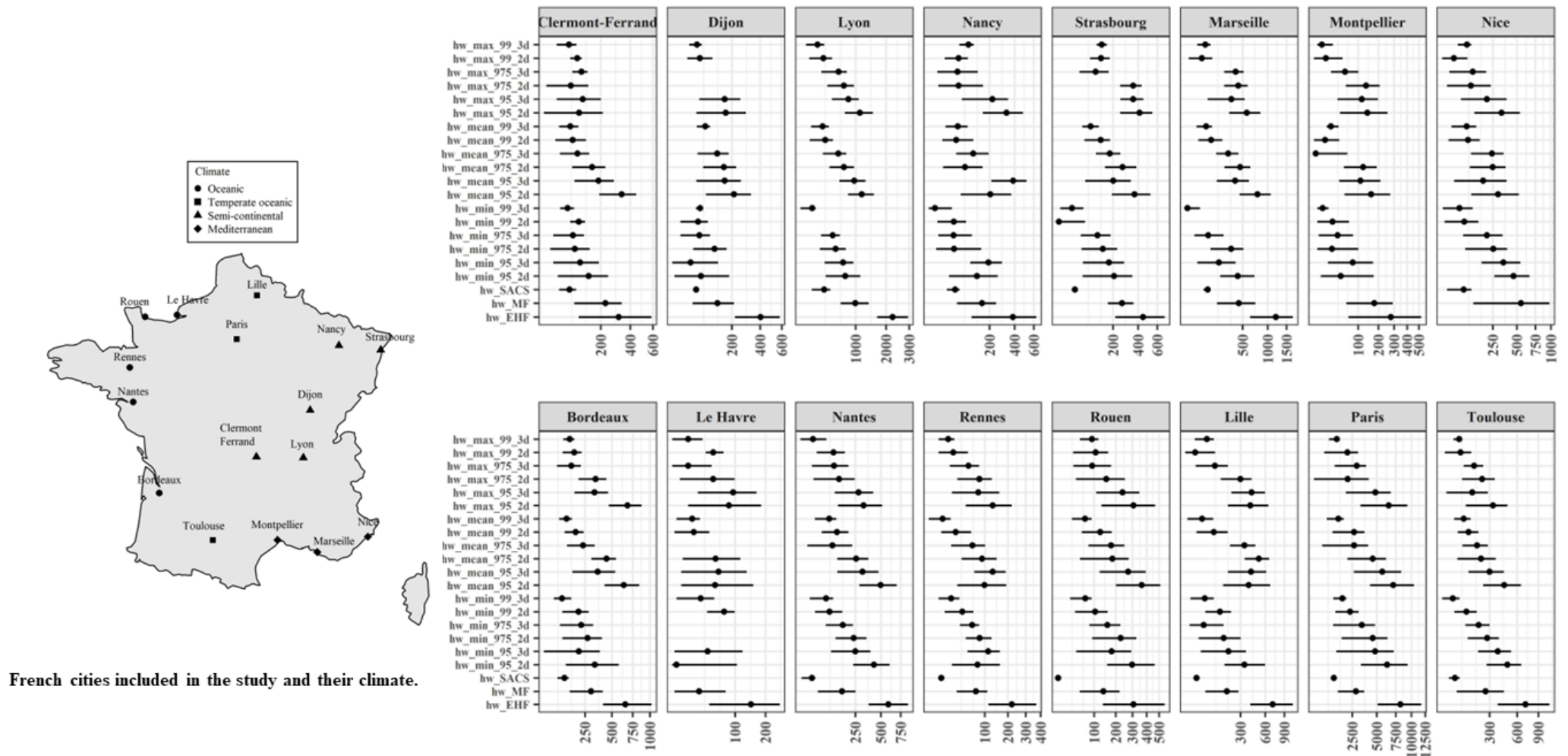
## **Sources of financial support**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Highlights

- Heat-related mortality was assessed on relative and absolute scales using a large set of heat waves definitions.
- Associations between heat waves and mortality differed greatly depending on the definition.
- According to the French heatwave warning system indicator, 129 heat-related deaths per summer were estimated.
- Excess heat factor was the most impactful definition for the temperate oceanic and Mediterranean climate types.

## Graphical abstract



Number of heat wave deaths in France estimated with a large set of heat wave definitions (May-September, 2000-2015).

The definitions are the Excess heat factor of the Australian Bureau of Meteorology (hw\_EHF), the climatological indicator of the French national weather service (hw\_MF), the heat wave indicator of the French national heatwave warning system (hw\_SACS), and the others based on the temperature parameters, which are presented in the format as follow: (heatwave)\_(temperature)\_(percentile)\_(minimum number of consecutive days). For example, hw\_min\_95\_2d defines a heat wave if the minimum temperature was above the 95<sup>th</sup> percentile of the minimum temperature for at least two days.

## Abstract

**Introduction:** In France, a heat warning system (HWS) has been implemented almost two decades ago and rely on some official heat wave (HW) definitions. However, no study has compared the burden associated with a large set of alternative HW definitions to the official definitions. Such comparison could be particularly helpful to identify HW conditions for which effective HWS would minimize the health burden across various geographical contexts and possibly update thresholds to trigger HWS. The aim of this study is to identify (and rank) definitions that drive the highest health burden in terms of mortality to inform future HWS across multiple cities in France.

**Methods:** Based on weather data for 16 French cities, we compared the two official definitions used in France to: i) the Excess Heat Factor (EHF) used in Australia, and ii) 18 alternative hypothetical HW definitions based on various combinations of temperature metrics, intensity, and duration. Propensity score matching and Poisson regressions were used to estimate the effect of each HW exposure on non-accidental mortality for the May-September period from 2000 to 2015.

**Results:** The associations between HW and mortality differed greatly depending on the definition. The greatest burden of heat was 1,055 (95% confidence interval "CI": [856; 1,302]) deaths per summer and was obtained with the EHF. The EHF identified HW with 2.46 (95% CI: [1.92; 3.58]) or 8.18 (95% CI: [6.63; 10.61]) times the global burden at the national level obtained with the climatological indicator of the French national weather service and the HW indicator of the French national HWS, respectively and was the most impactful definition pattern for both temperate oceanic and Mediterranean climate types.

**Conclusion:** Identifying the set of extreme heat conditions that drive the highest health burden in a given geographical context is particularly helpful when designing or updating heat early warning systems.

**Keywords:** climate change; heat waves; temperature; mortality; heat warning system; evidence-based decision making.

# 1. Introduction

Climate change has been repeatedly identified as a major issue for public health for contemporary and future generations (Watts et al., 2018). The most emblematic consequence of climate change is the increase in the average temperature of the globe (Pascal, 2021; Pascal et al., 2019). Between 1959 and 2015, on a ten-year rolling average, France has seen the annual number of heat wave (HW) days, identified by the climatological indicator of Météo-France (French national weather service), increasing from one to five (Soubeyroux et al., 2016). In a scenario of no regulation of greenhouse gas emissions, the average temperature in France could rise sharply until the end of the century (Soubeyroux et al., 2020). Increased mortality is the most extreme of a wide range health impacts associated to extreme temperatures. These impacts are considered preventable with the implementation of individual and structural measures, intended to reduce heat exposure (McElroy et al., 2020; Pascal, 2021; Vicedo-Cabrera et al., 2018).

HWs can considerably increase mortality, as observed during the historical HW of summer 2003 in Europe (Le Tertre et al., 2006). However, there is no universally accepted definition for HWs (Perkins and Alexander, 2013; Robinson, 2001). The general definition remains very broad, describing a period of unusually hot and dry or hot and humid weather, lasting at least two to three days, and having a discernible impact on health, quality of life, and infrastructure (McGregor and Menne, 2015). Most of several approaches found in the epidemiological literature defining HWs use a combination of temperature intensity and duration criteria (e.g. temperature above a historical temperature percentile, or an absolute value, generally for at least two to six days) (Hansen Alana et al., 2008; Nitschke et al., 2007; Robinson, 2001; Tong et al., 2010). HW indicators can be based on one or more temperature parameters (maximum, average or minimum temperature) on average values over several days. Temperature percentiles can be



calculated on the temperature distribution of the month, the summer, or the year. Finally, some indicators combine several criteria and several temperature parameters (Chen et al., 2015; D'Ippoliti et al., 2010; Pascal et al., 2019; Perkins, 2015).

A small change in the definition of HW has a significant effect on the estimated health impact (Pascal, 2021; Tong et al., 2010). On the one hand, a greater number of identified HW days leads to more important associated excess number of deaths (Anderson and Bell, 2011; Corso et al., 2017; Pascal et al., 2019; Gasparrini et al., 2015b). On the other hand, some HW definitions have shown similar attributable mortalities despite differences in the number of HW days identified (McElroy et al., 2020). Previous work reported that both the attributable burden and the relative risk of HW mortality are important to report and may not consistently coincide (Benmarhnia et al., 2014; McElroy et al., 2020; Nori-Sarma et al., 2019). It is important to identify HWs definitions that are driving the highest health burden when designing or updating public health intervention strategies to prevent and mitigate the impact of HWs in order to maximize the potential benefits. Following the 2003 HW, the French authorities designed a HW warning system based essentially on one definition chosen for its ability to provide a compromise between an acceptable number of alerts and an acceptable number of false alerts (Antics et al., 2012; Laaidi and Pascal, 2004). Yet, no previous study has compared in France a large set of HWs definitions to identify HWs conditions that would minimize the health burden across various geographical contexts.

Our hypothesis is that the absolute risk of mortality is a relevant measure for defining the set of criteria for activating HW warning systems. The aim of this study is to identify (and rank) definitions that drive the highest health burden in terms of mortality to inform future early warning systems across multiple cities in France. Specifically, we conducted an epidemiological study based on the data from 16 French cities, and after identifying HW days

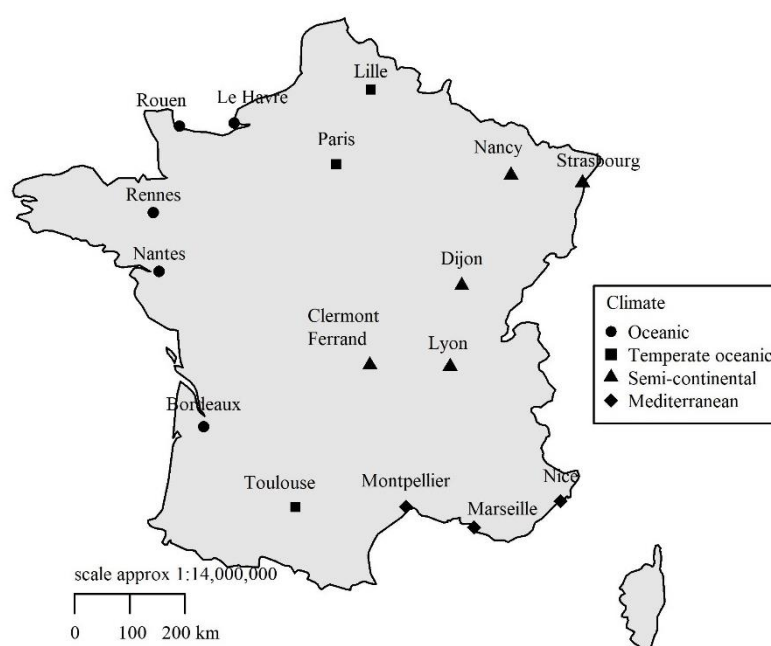
according to 21 definitions, we estimated the relative risk of mortality and mortality attributable to HW defined by each of these 21 definitions.

## 2. Materials and methods

### 2.1. Study period and areas

The study covered the period from 2000 to 2015. This period includes a phase when no heat prevention was available in France (2000-2003), and a phase when the national HW prevention plan was implemented (2004-2015).

Sixteen cities in metropolitan France were selected: Bordeaux, Clermont-Ferrand, Dijon, Le Havre, Lille, Lyon, Marseille, Montpellier, Nancy, Nantes, Nice, Paris, Rennes, Rouen, Strasbourg, and Toulouse. These cities are representative of the diversity of climates (namely the oceanic, the temperate-oceanic, the semi-continental and the Mediterranean climate) and of urban environments in France (Figure 1).



**Figure 1.** Cities included in the study and their climate.

## **2.2. Data**

### *2.2.1. Weather data*

For all 16 cities, the French national weather service (Météo-France) provided the daily observed temperature data (maximum, mean and minimum) in degree Celsius (°C), precipitation height (in millimeter, *mm*), average wind speed (in meter per second, *m/s*) and average relative humidity (in percent, %) from one reference meteorological station per city (see list of stations in Table S1).

### *2.2.2. Mortality data*

In each city, daily deaths for the entire population and for non-accidental causes (ICD10: A00-R99) were provided by the Center for Epidemiology on medical Causes of Death (CépiDC) of the National Institute of Health and Medical Research (Inserm) for the period 2000-2015.

## **2.3. Heat waves definitions**

We used a total of 21 definitions of HWs. Three of them are officially used in France and Australia and the others were developed. The Australian indicator was included for its potential for international applicability (Nairn and Fawcett, 2015).

The step of identification of HW episodes was focused on the days in the periods going from May to September. HWs were identified in each city using each definition. A binary variable was determined for each day, having value 1 for HW days identified according to a specific definition and 0 otherwise. For each indicator, at the national level (all cities combined), the HW days over the entire period 2000-2015 were cumulated without considering whether they were HWs in other cities. For example, a HW day in one city and in two other cities was counted three times.

### 2.3.1. *Heat wave indicator of the French national heatwave warning system*

We retrospectively applied the official SACS (*Système d'Alerte Canicule et Santé*) definition, which is based on comparing minimum and maximum biometeorological indicators (in the following, IBM, in reference to the corresponding expression in French) to the defined alert thresholds. The IBM are the three-days rolling average of minimum (*IBM min*) and maximum (*IBM max*) temperatures provided by a reference weather station representative of the city's climate (Laaidi and Pascal, 2004; Pascal et al., 2013):

A HW is identified by a period when the *IBM min* and *IBM max* simultaneously reached or exceeded the alert thresholds for at least three consecutive days. The minimum duration of a HW for a given city is thus, by definition, three nights and three consecutive days. The thresholds correspond to the 99.5<sup>th</sup> percentile of the rolling average temperature distribution for the months of June to August from 1973 to 2002 (Laaidi et al., 2005; Laaidi and Pascal, 2004; Pascal et al., 2013; Ung et al., 2019). The thresholds (in °C) used in this study are those used in 2005 (see Table S2), defined at the departmental level (Laaidi et al., 2005) and used according to the department in which each city is located.

### 2.3.2. *Climatological indicator of the French national weather service (Météo-France)*

A HW has been defined as a period of at least three consecutive days with the daily mean temperature above the initial threshold (97.5<sup>th</sup> percentile of the temperature distribution), with at least one day with the daily mean temperature exceeding the peak heat threshold (99.5<sup>th</sup> percentile) (Soubeyroux et al., 2016). The HW is terminated if the mean temperature falls below the 95<sup>th</sup> percentile (interruption threshold) for even a single day. The thresholds are defined at the departmental level. The percentiles used to define these thresholds in our study are those from the work of Pascal et al. (Pascal et al., 2019) and calculated over the entire year and the period 1981-2010 (Table S2).

### 2.3.3. *The Excess Heat Factor (Australian Bureau of Meteorology)*

The Excess Heat Factor (EHF) is comprised into a "historical" indicator ( $EHI_{sig}$ ) defining a significant deviation from a historical threshold (95<sup>th</sup> percentile of the mean temperature distribution over the last 30 years), and an "acclimation" indicator ( $EHI_{accl}$ ), providing a deviation from the temperature of the previous 30 days (Nairn and Fawcett, 2015). Both deviations are calculated from the three-days moving average of daily mean temperatures.

The 95<sup>th</sup> percentiles are calculated at the departmental level on the daily mean temperatures, full year, over the period 1981-2010 (Pascal et al., 2019) (Table S2). The overall indicator (EHF) is calculated as the product of these two indicators:

$$EHF = \max(0, EHI_{sig}) * \max(1, EHI_{accl})$$

There is a HW if EHF is strictly greater than zero. In other words, a HW is identified if the three-days moving average of daily mean temperatures is both strictly above, first, the historical normal temperature (defined as the 95<sup>th</sup> percentile of the distribution of daily mean temperatures during the whole year and over the period 1981-2010), and second, the average value over the previous 30 days of daily mean temperatures.

### 2.3.4. *Alternative definitions of heat waves*

To assess heat-related health impacts more broadly, we used 18 alternative definitions based on both heat intensity and duration. These definitions are based on values of a temperature parameter compared to defined thresholds. The temperature parameters were the maximum, the mean, and the minimum temperature used separately. The thresholds were the 95<sup>th</sup>, 97.5<sup>th</sup>, and 99<sup>th</sup> percentiles of the whole-year temperature distribution over the period 2000-2015.

The definition of HWs was a period where the values of the temperature parameter (maximum temperature taken as an example in Table 1) reach or exceed the corresponding thresholds during at least two or three consecutive days.

**Table 1.** Alternative definitions of heat waves based on the maximum daily temperature.

| Heat waves indicators | Maximum temperature                | Duration (days) |
|-----------------------|------------------------------------|-----------------|
| hw_max_95_2d*         | $\geq 95^{\text{th}}$ percentile   | $\geq 2$        |
| hw_max_95_3d          | $\geq 95^{\text{th}}$ percentile   | $\geq 3$        |
| hw_max_975_2d         | $\geq 97.5^{\text{th}}$ percentile | $\geq 2$        |
| hw_max_975_3d         | $\geq 97.5^{\text{th}}$ percentile | $\geq 3$        |
| hw_max_99_2d          | $\geq 99^{\text{th}}$ percentile   | $\geq 2$        |
| hw_max_99_3d          | $\geq 99^{\text{th}}$ percentile   | $\geq 3$        |

\* hw\_max\_95\_2d defines a heat wave of at least two consecutive days where the maximum temperature is above the 95<sup>th</sup> percentile of the maximum temperature.

## 2.4. Statistical analysis

### 2.4.1. Propensity score matching and Poisson regressions (main analysis)

Propensity score (PS) methods have recently been used in a wide range of epidemiological studies (their usage increased from 294 publications in the years 1998–2002 to 1,111 in 2003–2007 and 3,539 in 2008–2012) (Littnerova et al., 2013) to reduce confounding effects (Austin, 2011) and the impact of treatment-selection bias in the estimation of causal treatment effects (Austin, 2009). PS matching is frequently used in medical literature (Austin, 2008; Gayat et al., 2010; Weitzen et al., 2004) and particularly in environmental epidemiology (McElroy et al., 2020; Nori-Sarma et al., 2019), and it creates a balanced dataset, allowing separation of study design from study analysis and a simple and direct comparison of baseline covariates between treated and untreated units of analysis (Austin, 2011).

To analyze the impacts of HWs on mortality for each city, we used a PS matching approach (Austin, 2011; Rosenbaum and Rubin, 1983) for all definitions and then we proceeded to model the effect of different type of HWs on mortality observed the day the HW occurred and one day after. The PS is defined as the probability for each observed day to experience a HW, according to different types of HW definitions. PSs were generated using logistic regression via a Generalized Additive Model (GAM), in which an indicator variable for HW was used as the dependent variable and a list of identified meteorological and calendar confounders was used as predictors.

The probability of each day to experience a HW ( $P[HW = 1]$ ) was then estimated as a function of the following covariates: a smooth function of the precipitation height ( $PH$ ), a smooth function of the average wind speed ( $WS$ ), a smooth function of the relative humidity ( $RH$ ), and categorical variables for day of week ( $DOW$ ), vacations ( $V$ ), bank holidays ( $BH$ ), month and year. We also considered a smooth function of the lag variables ranging from one to three days for the variables  $PH$ ,  $WS$ ,  $RH$ . For the smooth terms, we used thin plate regression type basic spline functions with three knots ( $k$ ):  $s(covariate, bs = "tp", k = 3)$ . In practice,  $k-1$  (or  $k$ ) sets the upper limit on the degrees of freedom ( $df$ ) associated with a smooth term (one  $df$  is usually lost to the identifiability constraint on the smooth) (Wood, 2017).

The following model was then replicated for each city and for each of the 21 HW definitions:

$$\begin{aligned} \text{logit } P[HW = 1 | X] \\ = \beta_0 + \beta_1 DOW + \beta_2 BH + \beta_3 V + \beta_4 month + \beta_5 year + s(PH) + s(WS) \\ + s(RH) + \sum_{i=1}^3 [s(lag_i PH) + s(lag_i WS) + s(lag_i RH)] \end{aligned}$$

For each definition, a one-to-one matching based on the value of the PS was performed (within-city matching). This method allows to compare days identified as HW by a given definition

with very similar non-HW days according to what is predicted by the set of covariates included in the PS calculation. The matching was performed in R using the *Match()* function in the *Matching* package (Sekhon, 2011) by specifying that every case (HW day) has be matched to a control (non-HW day) having the closest PS value and within the same year. The matching was done with replacement and by randomly breaking ties. The capacity of the matching procedure to properly match similar days was graphically evaluated through love plots, which show the absolute standardized mean difference between the two groups (HW days and non-HW days) for covariates before and after matching (Greifer, 2021). The threshold of 0.1 in absolute standardized mean differences was used to objectively assess the balance of covariates between the two groups before and after matching, with values lower than 0.1 suggesting better balance (Austin and Stuart, 2015).

The panel of matched observations was analyzed through a Poisson regression model separately for each city. General linear models with a Poisson distribution and a log link function were run in R using the *pglm()* function in the package of the same name (Croissant, 2020) to assess the relationship between HWs identified according to the different definitions and, first, non-accidental deaths observed the same day ( $d$ ) for all definitions, and second, non-accidental deaths observed the following day ( $d+1$ ). The following model was estimated for each city and for each HW definition:

$$y_{i,t} \sim P(\lambda_{t,i})$$

$$\log(\lambda_{t,i}) = \beta HW_{t,i} + \varepsilon_{t,i}$$

In this model, the daily mortality  $y$  for day  $t$  in the matched cluster  $i$  was assumed to follow a Poisson distribution having the parameter  $\lambda$  as mean and variance.  $HW$  represents a binary variable indicating whether the days were HW or not.



#### 2.4.2. Generalized additive model (sensitivity analysis)

In sensitivity analysis, a multivariable GAM using a Poisson distribution was constructed, with daily mortality as the dependent variable. This model is not estimated based on the PS matched sample but based on the whole sample of days in the May-September period of 2000-2015. The explanatory variables were a smooth function of *PH*, a smooth function of *WS*, a smooth function of *RH*, and categorical variables for day of week (*DOW*), vacations (*V*), bank holidays (*BH*), month, year and a binary variable indicating whether the days were HW or not (*HW*). We also included a smooth function of the lag variables ranging from one to three days for the variables *PH*, *WS*, *RH*. We used thin plate regression type basic spline functions for the smooth terms with three knots (*k*):  $s(\text{covariate}, bs = "tp", k = 3)$ .

In environmental epidemiology, many time-series studies adjust for seasonality and long-term time-trend by adding a single time variable (Bell et al., 2004; Dominici et al., 2000) or a set of variables representing time (such as variable of calendar time of the year called day of the year or day of the season "*DOS*", and an indicator of year or equivalent) (Gasparrini et al., 2015a; Guo et al., 2017) into regression models. Natural cubic splines (*ns*) with some *df* per year are frequently used (Bhaskaran et al., 2013; Kim et al., 2021). In mortality studies, 3–8 *df* per year for the single time variable (corresponding to 3–8 *df* at the year-scale for the *DOS* variable) and one *df* per ten years for the variable of year are often used (Gasparrini et al., 2015a; Guo et al., 2017; Peng et al., 2006; Perrakis et al., 2014; Schwartz et al., 2003), but depending on whether other variables explain seasonality to some extent are included in models, the optimal *df* may decrease (e.g., influenza epidemic, heat wave) (Schwartz et al., 2003). Based on this previous work and the restriction to 5-month seasons in our study (allowing 1–3 *df* in a 5-month scale for the *DOS* variable), we introduced a combination of two temporal indicators (described below) using *ns* functions with two *df* for each of the two temporal indicators (non-integer *df* were rounded to the nearest integer):  $ns(\text{covariate}, df = 2)$ . The first indicator, to control for

seasonality, is the variable for the *DOS*, going from one at the first of May of every year and until 153 for 30<sup>th</sup> of September. The second indicator, to control for long-term trends, is the variable for the year (*YEAR*), going from one in 2000 to 16 in 2015. Sensitivity analyses were carried out to explore the variability of estimate of interest by varying the *df* (1 and 3) in the models for each of the 21 definitions.

The corresponding model was:

$$\begin{aligned}
y_t &\sim P(\lambda_t) \\
\log(\lambda_t) &= \beta_0 + \beta_1 HW + \beta_2 DOW + \beta_3 BH + \beta_4 V + \beta_5 month + \beta_6 year + s(PH) \\
&+ s(WS) + s(RH) + ns(DOS) + ns(YEAR) \\
&+ \sum_{i=1}^3 [s(lag_i PH) + s(lag_i WS) + s(lag_i RH)] + \varepsilon_t
\end{aligned}$$

In this model, the daily mortality  $y$  for day  $t$  was assumed to follow a Poisson distribution having the parameter  $\lambda$  as mean and variance. The model was then replicated for each city and for each of the 21 HW definitions.

#### 2.4.3. Estimation of the relative risk and mortality attributable to heat wave

The relative risks (RRs) of nonaccidental mortality during HW days versus non-HW days were obtained through the exponential of the coefficient  $\beta$  associated to the HW variable. The added effects of HWs on mortality on the days following the HW days ( $d+1$ ) were also estimated by the same operation.

To estimate pooled RRs at the national level, a random-effects meta-analysis was performed to pool the findings of the different cities (Viechtbauer, 2010). We assessed residual heterogeneity in the meta-analytic model by examining the results of the Cochran Q test and the  $I^2$  statistic.

At each city level, we estimated for all definitions the absolute number of deaths related to the HW days that they identified. First, when there was at least one HW day identified by a given definition in a city and the associated RR of mortality was greater than 1, the heat-attributable fraction (*HAF*) was calculated for this city and HW indicator using the formula:

$$HAF = (RR-1)/RR$$

Then, the number of deaths attributable to each definition for each city is obtained by multiplying the *HAF* by the mean daily deaths count over the May-September period of 2000-2015 and the total number of days identified by the definition as HW in the same period (Benmarhnia et al., 2014; McElroy et al., 2020; Vaidyanathan et al., 2019).

At the national level and over the entire May-September period from 2000 to 2015, for each definition, the mortality burden on day  $d$  was obtained by summing the attributable mortalities of the 16 cities. The same operation was performed to obtain the mortality on day  $d+1$  and the overall impact in terms of mortality (on both days  $d$  and  $d+1$ ) was obtained by summing the mortality attributable to the HWs on days  $d$  and  $d+1$ . We then divided the total number of attributable deaths for the period 2000–2015 by 16 to obtain an average attributable number of deaths per "summer" (May-September period). The number of attributable deaths per summer obtained with each HW definition was determined relative to that obtained with the SACS, Météo-France, and EHF definitions in terms of ratios and absolute differences. We calculated non-parametric bootstrap confidence intervals (CIs) for each point estimate using 1,000 samples with replacement and computing the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentiles (Carpenter and Bithell, 2000).

The descriptive and statistical analysis of the data were performed using R software (version 4.2.0) (R Core Team, 2022).

### 3. Results

#### 3.1. Descriptive statistics

Over the period 2000-2015, from May to September, a total of 606,503 deaths due to non-accidental causes were recorded for all cities, with an average of 16 (standard deviation "SD": 23.4) deaths per day. Between cities, daily non-accidental mortality data showed high variability with values ranging from 3.6 (SD: 1.8) to 99.8 (SD: 24.3) average daily deaths (Rennes and Paris respectively) (Table S3).

Maximum, mean, and minimum temperature density curves over all months of the 2000-2015 period are provided for all cities in the supplementary materials (Figure S1). The number of cumulative HW days identified by the indicators ranged on average from 6.6 (SD: 6.0) to 504.8 (SD: 340.7) per summer (all cities combined), obtained with the HW indicator SACS and the EHF respectively. The SACS definition and the indicators *hw\_min\_99\_2d*, *hw\_min\_99\_3d*, *hw\_mean\_99\_2d*, *hw\_mean\_99\_3d*, *hw\_max\_99\_2d* and *hw\_max\_99\_3d* had identified on average less than 30 HW days per summer (all cities combined) (Table S4).

Considering each city, the SACS indicator (a combination of rolling average of maximum and minimum temperatures) was the most conservative (for example, it identified no HW days during the whole study period in cities like Le Havre and Montpellier), followed by the indicators *hw\_max\_99\_3d* with eight days identified in Marseille and Nice and *hw\_min\_99\_3d* with nine days identified in Lyon and 12 in Clermont-Ferrand. The EHF was the indicator that by far identified the higher number of days of HWs with for example 1,665 days in Nice and 687 days in Clermont-Ferrand (results not shown).

### **3.2. Balance diagnostics of the PS matching procedure**

The PS matching procedure correctly matched similar case and control days according to the different covariates, for all definitions and in all cities, except for the indicators *hw\_max\_99\_3d*, SACS (correct matching only for Lyon), *hw\_min\_99\_3d* (correct matching only for Nantes, Nancy, and Nice), *hw\_mean\_99\_3d* (correct matching only for Nantes, Strasbourg, and Toulouse).

Indeed, the difference between the medians and the distribution of the PSs of the two groups (HW days and non-HW days) is very marked before matching, and practically null after the matching and the PS distributions of the two groups are quasi-symmetric and comparable. While the absolute standardized mean difference between the two groups (HW days and non-HW days) was important for some covariates in the unmatched sample, it was mainly below the threshold of 0.1 in the matched sample for most of the covariates (Figure S2). For definitions such as the SACS indicator, the distribution of the PS appeared identical before and after matching. The absolute standardized mean difference was important for several continuous covariates and close to zero for most of the binary variables globally in both unmatched and matched samples (Figure S3).

### **3.3. Relative risk of mortality**

At the national level, the impact of HW days on mortality on the same days compared to non-HW days in terms of RR ranged from 1.08 (95% CI: [1.05; 1.11]) for EHF to 2.06 (95% CI: [1.72; 2.47]) for the SACS indicator. The RRs were significantly greater than 1 for all definitions. The added HWs effects on the next days were also significantly greater than 1 and ranged in terms of RR from 1.07 (95% CI: [1.04; 1.10]) for EHF to 1.90 (95% CI: [1.54; 2.35]) for the SACS indicator (Table 2). Residual intra-city heterogeneity was present and significantly important (the p-value of Q statistic was below 0.05 and  $I^2$  was above 50%). Some

degree of heterogeneity within the same climatological profiles was observed graphically, with some CIs of RRs that did not overlap (results not shown).

At the national level, for all definitions, the point estimates of the RR of mortality obtained in the main analysis (Poisson model with matched data) were overall slightly higher than those obtained in the sensitivity analyses (multivariable GAM using 2 (Table S5) and 1 or 3 (Table S6) *df* per summer for the DOS natural spline functions) with, however, an overlap in their CIs. In the sensitivity analysis, using 2 *df* per summer for the DOS natural spline functions, the RR ranged from 1.06 (95% CI: [1.04; 1.08]) with EHF to 1.70 (95% CI: [1.43; 2.02]) with SACS indicator on HW days compared to non-HW days and from 1.05 (95% CI: [1.03; 1.07]) with EHF to 1.74 (95% CI: [1.48; 2.06]) with SACS indicator the days after HW (Table S5).

**Table 2.** Pooled relative risk of mortality associated with each definition and burden of mortality in national level in France (estimated over May-September period of 2000-2015) during heat wave day and the day after heat wave.

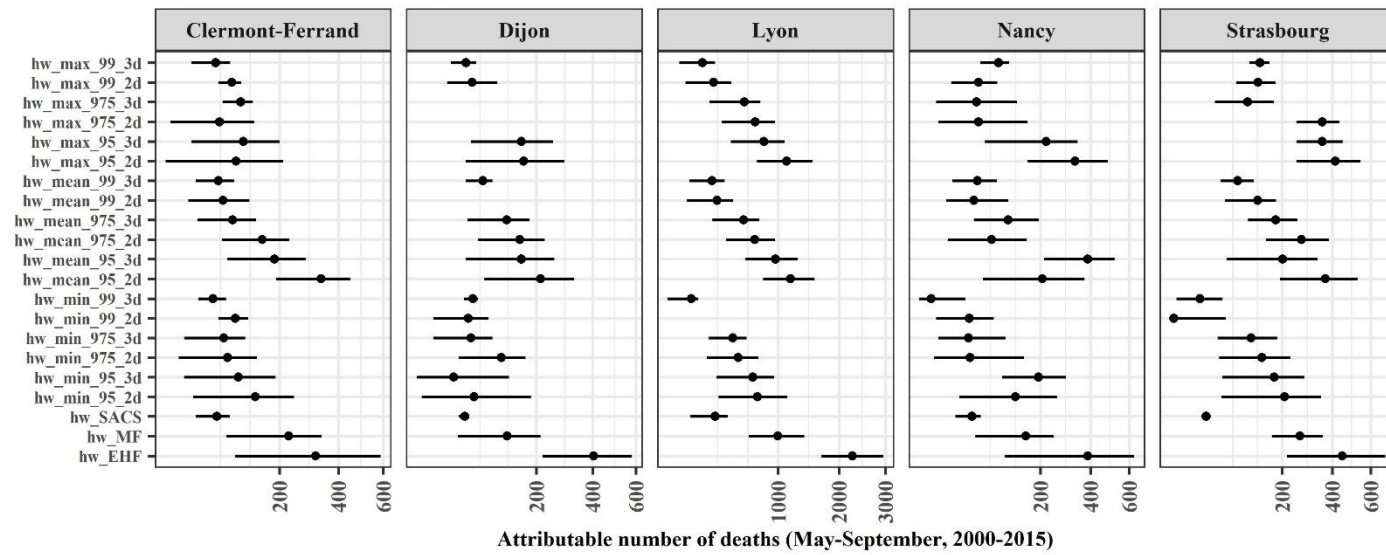
| Definitions    | During heat wave day |              |         |                 | During the day after heat wave |              |          |                 |
|----------------|----------------------|--------------|---------|-----------------|--------------------------------|--------------|----------|-----------------|
|                | RR                   | 95% CI       | Main AN | 95% CI          | RR                             | 95% CI       | Added AN | 95% CI          |
| hw_EHF         | 1.08                 | [1.05; 1.11] | 8,713   | [6,998; 10,755] | 1.07                           | [1.04; 1.10] | 8,171    | [6,375; 10,443] |
| hw_MF          | 1.22                 | [1.12; 1.32] | 3,848   | [2,879; 4,464]  | 1.17                           | [1.08; 1.27] | 3,022    | [2,107; 3,621]  |
| hw_SACS        | 2.06                 | [1.72; 2.47] | 1,090   | [938; 1,170]    | 1.90                           | [1.54; 2.35] | 973      | [803; 1,059]    |
| hw_min_95_2d   | 1.10                 | [1.06; 1.13] | 5,410   | [3,714; 7,056]  | 1.10                           | [1.07; 1.14] | 5,052    | [3,526; 6,741]  |
| hw_min_95_3d   | 1.15                 | [1.10; 1.20] | 4,287   | [2,459; 5,538]  | 1.12                           | [1.08; 1.17] | 3,654    | [1,812; 4,980]  |
| hw_min_975_2d  | 1.18                 | [1.12; 1.26] | 3,825   | [2,323; 4,869]  | 1.16                           | [1.11; 1.21] | 3,452    | [1,742; 4,369]  |
| hw_min_975_3d  | 1.27                 | [1.18; 1.36] | 2,798   | [1,590; 3,578]  | 1.20                           | [1.09; 1.33] | 2,487    | [1,387; 3,284]  |
| hw_min_99_2d   | 1.27                 | [1.11; 1.45] | 1,691   | [1,181; 2,140]  | 1.23                           | [1.09; 1.39] | 1,647    | [1,122; 2,107]  |
| hw_min_99_3d   | 1.58                 | [1.32; 1.89] | 1,313   | [982; 1,485]    | 1.36                           | [1.11; 1.65] | 1,166    | [847; 1,342]    |
| hw_mean_95_2d  | 1.13                 | [1.10; 1.17] | 6,671   | [5,083; 8,489]  | 1.13                           | [1.10; 1.17] | 6,503    | [4,778; 8,315]  |
| hw_mean_95_3d  | 1.16                 | [1.13; 1.20] | 5,244   | [3,577; 6,581]  | 1.15                           | [1.09; 1.20] | 4,844    | [3,162; 6,177]  |
| hw_mean_975_2d | 1.20                 | [1.14; 1.25] | 4,110   | [2,741; 5,069]  | 1.22                           | [1.17; 1.28] | 4,375    | [2,929; 5,230]  |
| hw_mean_975_3d | 1.26                 | [1.18; 1.35] | 2,622   | [1,452; 3,362]  | 1.21                           | [1.13; 1.30] | 2,757    | [1,641; 3,449]  |
| hw_mean_99_2d  | 1.29                 | [1.17; 1.43] | 2,141   | [1,294; 2,689]  | 1.23                           | [1.12; 1.35] | 1,996    | [1,156; 2,544]  |
| hw_mean_99_3d  | 1.45                 | [1.26; 1.67] | 1,346   | [999; 1,569]    | 1.50                           | [1.33; 1.69] | 1,300    | [885; 1,529]    |
| hw_max_95_2d   | 1.14                 | [1.10; 1.17] | 5,829   | [4,050; 7,367]  | 1.13                           | [1.10; 1.17] | 6,147    | [4,317; 7,806]  |
| hw_max_95_3d   | 1.18                 | [1.14; 1.23] | 4,372   | [2,749; 5,459]  | 1.17                           | [1.12; 1.22] | 4,383    | [2,794; 5,411]  |
| hw_max_975_2d  | 1.15                 | [1.09; 1.21] | 1,973   | [1,140; 3,161]  | 1.19                           | [1.13; 1.25] | 3,283    | [1,870; 4,201]  |
| hw_max_975_3d  | 1.23                 | [1.11; 1.36] | 2,528   | [1,686; 3,067]  | 1.25                           | [1.12; 1.40] | 2,443    | [1,546; 2,958]  |
| hw_max_99_2d   | 1.21                 | [1.08; 1.34] | 1,492   | [769; 2,037]    | 1.36                           | [1.17; 1.58] | 1,907    | [1,148; 2,345]  |
| hw_max_99_3d   | 1.65                 | [1.40; 1.93] | 1,248   | [1,001; 1,402]  | 1.68                           | [1.46; 1.93] | 1,259    | [1,027; 1,387]  |

Abbreviations: RR, Relative risk of mortality; CI, Confidence interval; AN, Attributable number of deaths (May-September period of 2000-2015).

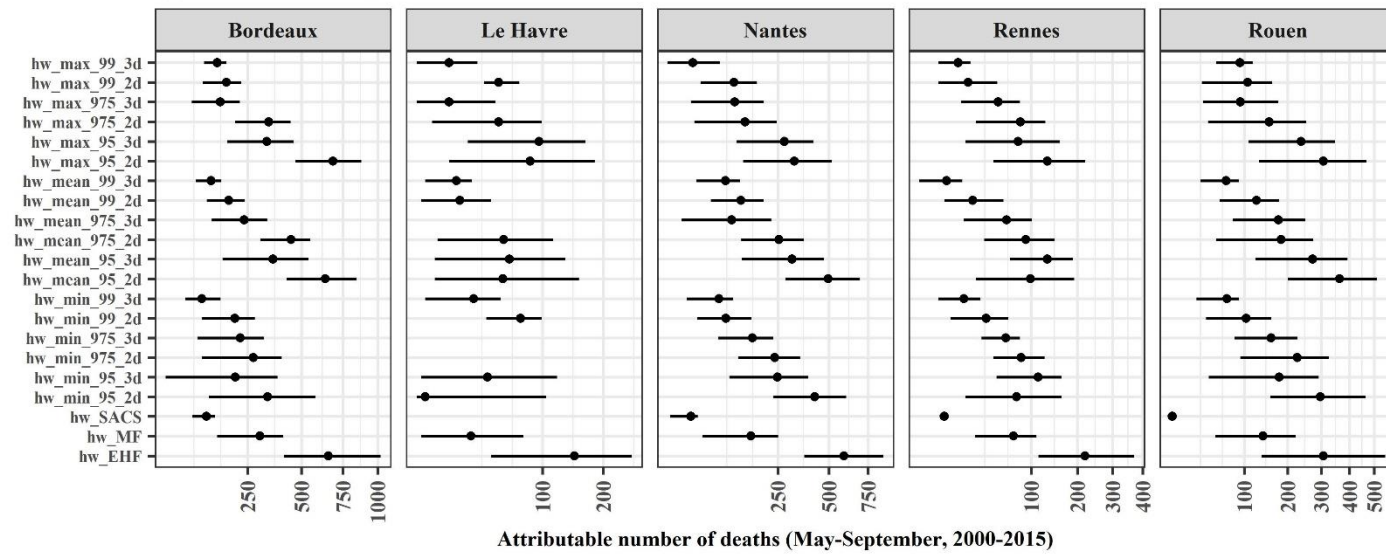
Notes: The definitions are the Excess heat factor of the Australian Bureau of Meteorology (hw\_EHF), the climatological indicator of the French national weather service (hw\_MF), the heat wave indicator of the French national heatwave warning system (hw\_SACS), and the others based on the temperature parameters, which are presented in the format as follow: (heatwave)\_(temperature)\_(percentile)\_(minimum number of consecutive days). For example, hw\_min\_95\_2d defines a heat wave if the minimum temperature was above the 95<sup>th</sup> percentile of the minimum temperature for at least two days.

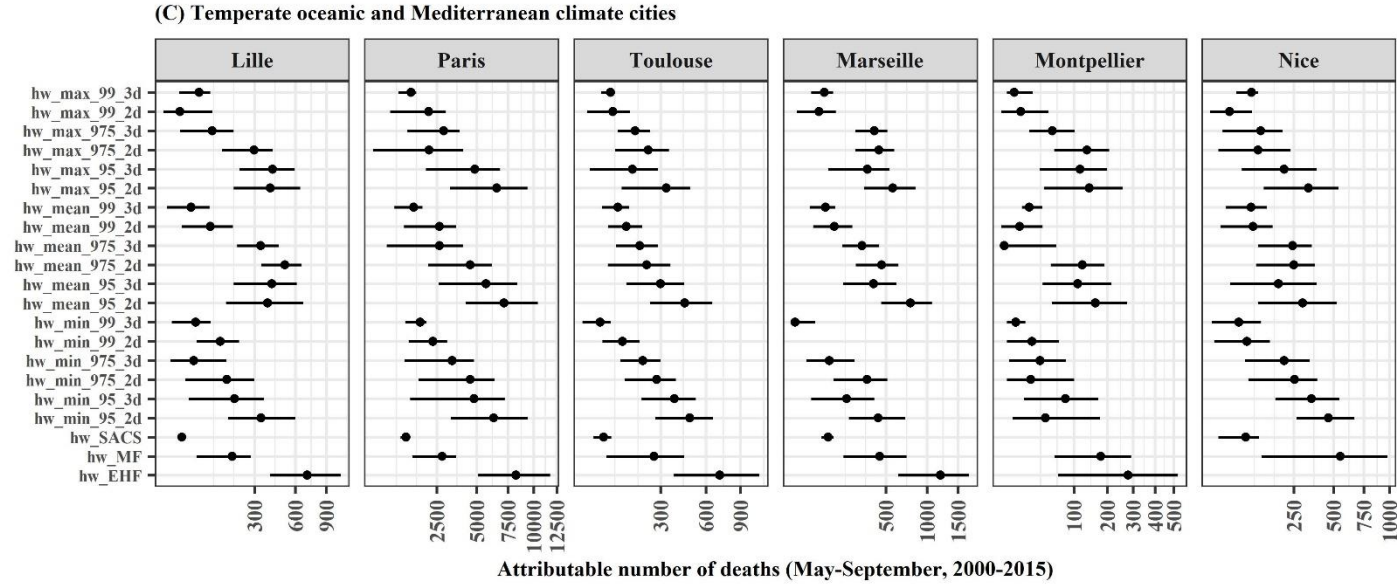


(A) Semi-continental climate cities



(B) Oceanic climate cities





**Figure 2.** Attributable number of cumulative deaths for the main and added impacts of heat wave obtained with all heat wave definitions (estimated over the May-September period of 2000-2015) for semi-continental (A), oceanic (B), temperate oceanic and Mediterranean (C) climate cities.

Notes: Temperate oceanic cities are Lille, Paris, and Toulouse. Mediterranean climate cities are Marseille, Montpellier, and Nice. The definitions are the Excess heat factor of the Australian Bureau of Meteorology (hw\_EHF), the climatological indicator of the French national weather service (hw\_MF), the heat wave indicator of the French national heatwave warning system (hw\_SACS), and the others based on the temperature parameters, which are presented in the format as follow: (heatwave)\_(temperature)\_(percentile)\_(minimum number of consecutive days). For example, hw\_min\_95\_2d defines a heat wave if the minimum temperature was above the 95<sup>th</sup> percentile of the minimum temperature for at least two days.

### 3.4. Burden of mortality attributable for heat waves definitions

The relationship between each RR (for HW days compared with non-HW days and for the day after HW days) and the corresponding number of attributable deaths differed considerably by climate zone and HW definitions. HWs with the highest RRs were not necessarily associated with the highest attributable number of deaths (Figures S4 - S7).

Through the results at the national level of the main analysis, the EHF defined the HWs associated with the highest attributable mortality burden with a main impact of 8,713 (95% CI: [6,998; 10,755]) non-accidental deaths, and an additional impact of 8,171 (95% CI: [6,375; 10,443]) non-accidental deaths. The EHF was followed by the indicators *hw\_mean\_95\_2d* (with 6,671 "95% CI: [5,083; 8,489]", and 6,503 "95% CI: [4,778; 8,315]" non-accidental deaths for the main and added impacts, respectively), and *hw\_max\_95\_2d* (with 5,829 "95% CI: [4,050; 7,367]", and 6,147 "95% CI: [4,317; 7,806]" non-accidental deaths for the main and added impacts, respectively) (Table 2). The results (heat-related mortality) of the sensitivity analyses (multivariable GAM based on all days in the period, not only PS matched days), using 2 (Table S5) and 1 or 3 (Table S6) *df* per summer for the DOS natural spline functions, were overall quite close to those of the main analysis (Table 2). In these sensitivity analyses, as in the main analysis, the same (aforementioned) HWs definitions were found to be the ones identifying the most mortality attributable to HWs (Tables 2, S5 and S6).

Focusing on the number of cumulative deaths for the main and added impacts, it appeared that HWs defined by EHF were associated with the highest burden of attributable deaths in all cities and climate zones except two out of five oceanic climate cities (Rouen with *hw\_mean\_95\_2d* and Bordeaux with *hw\_max\_95\_2d*) and two out of five semi-continental climate cities (Clermont-Ferrand with *hw\_mean\_95\_2d* and Nancy with *hw\_mean\_95\_3d*). HWs defined by the SACS indicator were consistently associated with the lowest impact on the burden of deaths

in one out of three temperate oceanic climate cities (Paris) and three out of five oceanic climate cities (Nantes, Rennes, and Rouen) (Figure 2).

The global (main and added) mortality burden related to HWs ranged from 129 (95% CI: [112; 137]) to 1,055 (95% CI: [856; 1,302]) deaths per summer obtained with the SACS and EHF definitions, respectively. Across each temperature parameter (minimum, mean, and maximum) defining the HW indicators, as percentiles decreased (resulting in many days identified as HW), corresponding heat-related mortalities increased (Table 3). The number of deaths attributable to HWs defined by the EHF corresponded to 2.46 (95% CI: [1.92; 3.58]), and 8.18 (95% CI: [6.63; 10.61]) times the number of deaths attributable to HWs defined by the Météo-France and SACS definitions, respectively. The next three definitions after EHF with the highest overall HW mortality burden were *hw\_mean\_95\_2d* (823 "95% CI: [629; 1,029]" deaths per summer), *hw\_max\_95\_2d* (748 "95% CI: [536; 930]" deaths per summer), and *hw\_min\_95\_2d* (654 "95% CI: [465; 851]" deaths per summer). Compared with the SACS definition, the other HWs definitions had identified HWs with an overall attributable mortality burden multiplied by a factor ranging from 1.20 (95% CI: [0.92; 1.46]) obtained with *hw\_min\_99\_3d* to 8.18 (95% CI: [6.63; 10.61]) obtained with EHF. This factor ranged from 0.30 (95% CI: [0.25; 0.39]) obtained with SACS definition to 2.46 (95% CI: [1.92; 3.58]) obtained with EHF relative to the Météo-France definition, and from 0.12 (95% CI: [0.09; 0.15]) obtained with SACS definition to 0.78 (95% CI: [0.56; 1.07]) obtained with *hw\_mean\_95\_2d* relative to EHF (Table 3). The overall burden of mortality per summer attributable to HWs in France obtained with each HW definition versus the SACS, Météo-France, and EHF definitions in terms of absolute differences are provided in the supplementary materials (Table S7).

**Table 3:** Global burden of mortality per summer in France (estimated over the period 2000-2015) obtained with each heat wave definition and variation relative to the SACS, Météo-France, and EHF definitions in terms of ratio.

| Definitions    | ANPS               | Ratio (95% CI) of each definition's ANPS by the one of |                   |                    |
|----------------|--------------------|--|-------------------|--------------------|
|                |                    | hw_EHF   | hw_MF             | hw_SACS            |
| hw_EHF         | 1,055 [856; 1,302] | 1  | 2.46 [1.92; 3.58] | 8.18 [6.63; 10.61] |
| hw_MF          | 429 [320; 494]     | 0.41 [0.28; 0.52]                                      | 1                 | 3.33 [2.55; 4.07]  |
| hw_SACS        | 129 [112; 137]     | 0.12 [0.09; 0.15]                                      | 0.30 [0.25; 0.39] | 1                  |
| hw_min_95_2d   | 654 [465; 851]     | 0.62 [0.42; 0.87]                                      | 1.52 [1.09; 2.23] | 5.07 [3.67; 6.87]  |
| hw_min_95_3d   | 496 [278; 643]     | 0.47 [0.25; 0.65]                                      | 1.16 [0.65; 1.66] | 3.85 [2.16; 5.25]  |
| hw_min_975_2d  | 455 [267; 571]     | 0.43 [0.24; 0.57]                                      | 1.06 [0.64; 1.50] | 3.53 [2.11; 4.61]  |
| hw_min_975_3d  | 330 [188; 421]     | 0.31 [0.17; 0.42]                                      | 0.77 [0.43; 1.12] | 2.56 [1.53; 3.45]  |
| hw_min_99_2d   | 209 [147; 263]     | 0.20 [0.13; 0.27]                                      | 0.49 [0.34; 0.72] | 1.62 [1.17; 2.14]  |
| hw_min_99_3d   | 155 [117; 176]     | 0.15 [0.10; 0.19]                                      | 0.36 [0.27; 0.48] | 1.20 [0.92; 1.46]  |
| hw_mean_95_2d  | 823 [629; 1,029]   | 0.78 [0.56; 1.07]                                      | 1.92 [1.45; 2.78] | 6.39 [4.95; 8.46]  |
| hw_mean_95_3d  | 630 [426; 787]     | 0.60 [0.38; 0.80]                                      | 1.47 [0.99; 2.10] | 4.89 [3.40; 6.49]  |
| hw_mean_975_2d | 530 [366; 635]     | 0.50 [0.32; 0.66]                                      | 1.24 [0.84; 1.72] | 4.11 [2.90; 5.22]  |
| hw_mean_975_3d | 336 [200; 420]     | 0.32 [0.18; 0.43]                                      | 0.78 [0.46; 1.10] | 2.61 [1.57; 3.44]  |
| hw_mean_99_2d  | 259 [164; 325]     | 0.25 [0.15; 0.33]                                      | 0.60 [0.38; 0.87] | 2.01 [1.31; 2.62]  |
| hw_mean_99_3d  | 165 [121; 191]     | 0.16 [0.11; 0.20]                                      | 0.39 [0.28; 0.53] | 1.28 [0.97; 1.60]  |
| hw_max_95_2d   | 748 [536; 930]     | 0.71 [0.49; 0.94]                                      | 1.74 [1.23; 2.55] | 5.81 [4.27; 7.64]  |
| hw_max_95_3d   | 547 [351; 670]     | 0.52 [0.31; 0.68]                                      | 1.27 [0.82; 1.78] | 4.24 [2.78; 5.43]  |
| hw_max_975_2d  | 328 [201; 448]     | 0.31 [0.18; 0.45]                                      | 0.77 [0.47; 1.18] | 2.55 [1.56; 3.63]  |
| hw_max_975_3d  | 311 [209; 369]     | 0.29 [0.18; 0.38]                                      | 0.72 [0.47; 1.01] | 2.41 [1.61; 3.03]  |
| hw_max_99_2d   | 212 [121; 272]     | 0.20 [0.11; 0.28]                                      | 0.49 [0.29; 0.72] | 1.65 [0.98; 2.23]  |
| hw_max_99_3d   | 157 [128; 172]     | 0.15 [0.11; 0.19]                                      | 0.36 [0.29; 0.50] | 1.22 [0.99; 1.46]  |

Abbreviations: ANPS, Attributable number of deaths per summer (May-September period) during heat wave day and the day after heat wave; CI, Confidence interval.

Notes: The definitions are the Excess heat factor of the Australian Bureau of Meteorology (hw\_EHF), the climatological indicator of the French national weather service (hw\_MF), the heat wave indicator of the French national heatwave warning system (hw\_SACS), and the others based on the temperature parameters, which are presented in the format as follow: (heatwave)\_(temperature)\_(percentile)\_(minimum number of consecutive days). For example, hw\_min\_95\_2d defines a heat wave if the minimum temperature was above the 95<sup>th</sup> percentile of the minimum temperature for at least two days.

## 4. Discussion

For the period from May to September of the years 2000 to 2015, we analyzed the impacts of HWs identified according to different definitions on mortality in 16 cities in France by using historical data of observed temperatures and non-accidental mortality provided by competent French institutions.

We adopted a PS matching approach that provided a matched sample of exposed and unexposed days who had similar PS values (Rosenbaum and Rubin, 1985, 1983). This approach includes the ability to graphically visualize the capacity of the procedure to match days that are similar (to approximate randomized experiments as closely as possible conditional on the observed covariates), to exclude days that are too different according to the covariates, and allows one to separate the design of an observational study from its analysis (Rubin, 2007). This makes easier to determine whether the model is correctly specified, comparing to a simple multivariable regression model relating the exposure variable (HW) and a set of identified confounding factors to mortality. Indeed, goodness-of-fit measures (such as model  $R^2$ ) does not provide a test of whether the multivariable model was correctly specified, and goodness-of-fit tests do not allow one to determine the degree to which the multivariable model has successfully eliminated systematic differences between HW and non-HW days (Austin, 2011). Furthermore, this method allows the HW effect to be estimated on the matched sample (as the average effect of HW on the days that ultimately experienced the HW, also known as average treatment effect on the treated "ATT") rather than on the overall sample (as average treatment effect "ATE") (Imbens, 2004). Multivariable regression may have increased flexibility when outcomes expressed as a binary variable are rare, and exposure is frequent (Braitman and Rosenbaum, 2002). Prior work has indicated that low number of events can lead to major problems and has suggested that at least 10 events should be observed for each variable included in a regression

model when the outcomes are binary (Peduzzi et al., 1996, 1995). However, in our context, for definitions such as the SACS indicator, *hw\_min\_99\_2d*, *hw\_min\_99\_3d*, *hw\_mean\_99\_2d*, *hw\_mean\_99\_3d*, *hw\_max\_99\_2d* and *hw\_max\_99\_3d* , the logistic regression model was flexible in the modeling of the PS for which the outcome (variable indicating HW or not) presented a small number of HWs (on average, less than 30 per city during the whole study period). Future studies could focus on a larger study period that would allow the identification of important HW events in a context of climate change with an increase in the occurrence of HWs.

Our results in the main (using Poisson regressions based on PS matched days) and sensitivity (using multivariable GAMs based on all days in the study period) analyses showed significant effects of HW overall comparable both for RR and heat-related deaths. It should be emphasized for RR that, similar to randomized controlled trials, PS methods (including matching) allow one to estimate population-average (or marginal) treatment effects (to estimate how the treatment/HW would change the risk of the outcome/mortality at the population of days level). In contrast, regression-based approaches allow one to obtain adjusted (or conditional) estimates of treatment effects (change in the risk of the outcome/mortality for an individual/day - at the subject/day level - when exposed to/experienced treatment/HW compared to receiving no treatment/not experiencing HW, conditional on the measured covariates) (Austin, 2011; Austin et al., 2007). Thus, depending on treatment effects researchers are interested in estimating using observational data, both approaches are appropriate analytical methods (Austin et al., 2007). However, in certain restrictive settings such as ours (count outcomes and RR as an effect measure) or linear regressions (continuous outcomes), marginal and conditional treatment effects coincide assuming that both the outcome regression model and the propensity score model were correctly specified (Austin, 2011; Austin et al., 2007).



We found that the mortality burden attributable to heat differed depending on the definitions used for HWs. Worldwide, HWs increase mortality, but there is variation in estimates of the effect depending on the definition of HWs (Anderson and Bell, 2009; Kent et al., 2014; McElroy et al., 2020; Tong et al., 2010; Xu et al., 2016). For example, McElroy et al. (McElroy et al., 2020) compared 18 HW definitions in order to identify the best indicator for targeting health impacts in San Diego County and found that small changes in HW definitions could result in substantial changes in health impacts.

Our results revealed that the choice of an appropriate definition of HWs should consider both the number of alerts it could trigger (in terms of the number of HW days identified) and the scales for HW effect measures (RR and attributable mortality). These results demonstrate that the ranking of HW definitions is not similar when using attributable mortality as compared to RRs. The definition of a HW that best considers its effects on the health of a population may not correspond to the highest RR, because this measure does not consider the frequency of HWs. For example, a severe HW episode with a stronger estimated health impact in term of RR, by occurring less frequently, may have smaller consequences in term of attributable mortality than a HW identified by a broader definition with smaller estimated RR but occurring more regularly.

We showed that the EHF indicator produced 8.18 times (95% CI: [6.63; 10.61]) the mortality attributable to heat obtained with the SACS indicator. EHF was the indicator that identified HWs with the greatest global (main and added) burden of heat attributable mortality at the national level (all cities combined) and was the most impactful definition pattern for both temperate oceanic and Mediterranean climate types. This indicator defines a HW based on the mean temperature and has advantage for considering both the historical exposure and an acclimation criterion. It followed by HWs indicators that separately use mean, maximum and minimum temperature, for the 95<sup>th</sup> percentile and a minimum duration of two consecutive days.

It may be beneficial in some cases (e.g., depending on the city or even the type of climate) to consider these definitions or a combination of mean, maximum, and minimum temperature parameters when designing a HW plan. However, it should be emphasized that the EHF indicator would produce an average of 504.8 alerts per summer in France (all 16 cities combined), compared to 6.6 for the SACS indicator.

This study found that the greater the number of HWs days identified (depending on the thresholds used), the greater the number of heat-related deaths. As temperature has an impact on mortality even in moderate heat (below the thresholds used to define HWs) (Anderson and Bell, 2011; Corso et al., 2017; Gasparrini et al., 2015b), a greater impact can be expected if the threshold window is extended, raising the delicate question of how to set thresholds for HW alert systems, or how to reach a balance between the number of alerts and the potentially associated excess mortality. Setting high thresholds only prevents a small number of heat-related deaths, but minimizes the number of alerts, as in the case of the SACS indicator, which follows an alert logic highlighting the impact of the most severe a priori events, but not the total impact of heat (Laaidi et al., 2005; Laaidi and Pascal, 2004). Setting lower thresholds avoids a large number of heat-related deaths at the cost of very frequent alerts, as is the case with the EHF indicator that highlights the existence of deaths during less extreme heat (Corso et al., 2017). In a context in which HWs will be extremely more intense, long-lasting, earlier (as of May) or later (until October) in case of increasing greenhouse gases, and of low risk perception in the population, the question of the choice of indicator thresholds could be resolved in a reasonable way by keeping alert management targeted on the most severe HWs (for which the final objectives of excess mortality explicitly defined and quantified would depend on the decision-makers), complemented by a basic prevention of heat impacts throughout the summer. This prevention could involve education of the general population, training of health professionals, improvement of housing thermal comfort, development of social links.

In France, to have an anticipatory and reliable system, the choice has been made to work with indicators and thresholds that predict (high) excess mortality (Laaidi et al., 2005; Laaidi and Pascal, 2004). These indicators and thresholds must be able to evolve according to new scientific knowledge, characteristics of future HWs which could be different from those of the last 30 years (in particular wetter, with possibly also different wind conditions), demographic or structural evolution of populations dying, evolution of cities in terms of establishment or departure of health establishments or accommodation establishments for dependent elderly people which could lead to a greater vulnerability of the population.

We also found a significant added effect of HWs on the mortality of the following HW days. In the literature, it is well documented that the effects of heat on mortality appear immediately on the same day and usually last for a few days (Guo et al., 2014; Rocklöv and Forsberg, 2008; Yu et al., 2012). To quantify the health impacts associated with heat, approaches used the temperature as a continuous metric (as in distributed lag linear or nonlinear models) or binary HW definitions based on temperature metrics. Furthermore, the choice of considering HWs as a binary exposure (HW and non-HW days) is more pertinent (Xu et al., 2016) compared to considering temperature as a continuous variable (Armstrong, 2006; Gasparrini et al., 2015b, 2010; Muggeo and Hajat, 2009; Pascal et al., 2018) when referring to the policy approach generally adopted in the HW warning systems (use of temperature thresholds to trigger warnings and risk communication).

In this study, we did not assess the determinants of the variations between or within cities in the burden associated with the different HW definitions. Indeed, several individual or contextual characteristics of subgroups marking greater vulnerability to heat effects have been documented (Benmarhnia et al., 2015; Gronlund, 2014; McElroy et al., 2020; Schwarz et al., 2021). These factors could include, for example, demographic factors (age, gender, urban design, population density, chronic diseases), socioeconomic factors (education, ethnicity,

income or social isolation), neighborhood (or ecological) community factors, and material conditions (air conditioning) (Åström et al., 2011; Basu, 2009; Xu et al., 2012). Understanding the spatial variation in these vulnerabilities to heat-related mortality is an important step in targeting specific actions to sensible subgroups to reduce heat-related effects on health (Kovats et al., 2003). In the calculation of the different temperature thresholds used in the design of alternative definitions, we used whole year data over the period 2000-2015. This does not provide enough time for an optimal estimation of the thresholds. In comparison, a period of 30 years is used for the definition of thresholds in the EHF, Météo-France, and SACS definitions (Laaïdi et al., 2005; Laaïdi and Pascal, 2004; Nairn and Fawcett, 2015).

Our study was conducted using observed temperature data, whereas in anticipation of HW events, warning systems use forecast data. We did not consider a possible mismatch between the forecasted and observed HWs. The impact of HWs in terms of mortality does not include the burden of hospitalizations, which is a more moderate signal but may reveal broader health impacts as showed by Schwarz et al. exploring spatial variation in the joint effect of extreme heat events and ozone on respiratory hospitalizations in California (Schwarz et al., 2021). The study period (2000-2015) mainly covered a phase when the national HW prevention plan was implemented (2004-2015). The mortality burden attributable to HWs may be underestimated in this study and may be explained, for example, by the preventive measures set up by the health institutions and authorities after the 2003 HW, the implementation of the heat health watch warning system as of summer 2004, an awareness of the risk associated with extreme temperatures and a reduction in the vulnerability of populations to heat.

In conclusion, differences in HWs definitions can result in substantial differences in estimated health outcomes. Identifying the set of extreme heat conditions that drive the highest health burden in a given geographical context is particularly helpful when designing or updating heat

early warning systems. Our findings may enable the implementation and reinforcement of government-, climate type-, and city-level management strategies for HWs.

## References

- Anderson, B.G., Bell, M.L., 2009. Weather-Related Mortality: How Heat, Cold, and Heat Waves Affect Mortality in the United States. *Epidemiology* 20, 205–213.  
<https://doi.org/10.1097/EDE.0b013e318190ee08>
- Anderson, G.B., Bell, M.L., 2011. Heat Waves in the United States: Mortality Risk during Heat Waves and Effect Modification by Heat Wave Characteristics in 43 U.S. Communities. *Environ. Health Perspect.* 119, 210–218.  
<https://doi.org/10.1289/ehp.1002313>
- Antics, A., Pascal, M., Laaidi, K., Wagner, V., Corso, M., Declercq, C., Beaudeau, P., 2012. A simple indicator to rapidly assess the short-term impact of heat waves on mortality within the French heat warning system. *Int. J. Biometeorol.* 57.  
<https://doi.org/10.1007/s00484-012-0535-9>
- Armstrong, B., 2006. Models for the Relationship Between Ambient Temperature and Daily Mortality. *Epidemiology* 17, 624–631.  
<https://doi.org/10.1097/01.ede.0000239732.50999.8f>
- Åström, D.O., Bertil, F., Joacim, R., 2011. Heat wave impact on morbidity and mortality in the elderly population: A review of recent studies. *Maturitas* 69, 99–105.  
<https://doi.org/10.1016/j.maturitas.2011.03.008>
- Austin, P.C., 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivar. Behav. Res.* 46, 399–424.  
<https://doi.org/10.1080/00273171.2011.568786>
- Austin, P.C., 2009. Some Methods of Propensity-Score Matching had Superior Performance to Others: Results of an Empirical Investigation and Monte Carlo simulations. *Biom. J.* 51, 171–184. <https://doi.org/10.1002/bimj.200810488>
- Austin, P.C., 2008. The performance of different propensity-score methods for estimating relative risks. *J. Clin. Epidemiol.* 61, 537–545.  
<https://doi.org/10.1016/j.jclinepi.2007.07.011>
- Austin, P.C., Grootendorst, P., Normand, S.-L.T., Anderson, G.M., 2007. Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a Monte Carlo study. *Stat. Med.* 26, 754–768. <https://doi.org/10.1002/sim.2618>
- Austin, P.C., Stuart, E.A., 2015. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* 34, 3661–3679.  
<https://doi.org/10.1002/sim.6607>
- Basu, R., 2009. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environ. Health Glob. Access Sci. Source* 8, 40.  
<https://doi.org/10.1186/1476-069X-8-40>

- Bell, M.L., Samet, J.M., Dominici, F., 2004. Time-Series Studies of Particulate Matter. *Annu. Rev. Public Health* 25, 247–280.  
<https://doi.org/10.1146/annurev.publhealth.25.102802.124329>
- Benmarhnia, T., Deguen, S., Kaufman, J.S., Smargiassi, A., 2015. Review Article: Vulnerability to Heat-related Mortality: A Systematic Review, Meta-analysis, and Meta-regression Analysis. *Epidemiology* 26, 781–793.  
<https://doi.org/10.1097/EDE.0000000000000375>
- Benmarhnia, T., Sottile, M.-F., Plante, C., Brand, A., Casati, B., Fournier, M., Smargiassi, A., 2014. Variability in Temperature-Related Mortality Projections under Climate Change. *Environ. Health Perspect.* 122, 1293–1298. <https://doi.org/10.1289/ehp.1306954>
- Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L., Armstrong, B., 2013. Time series regression studies in environmental epidemiology. *Int. J. Epidemiol.* 42, 1187–1195.  
<https://doi.org/10.1093/ije/dyt092>
- Braitman, L.E., Rosenbaum, P.R., 2002. Rare Outcomes, Common Treatments: Analytic Strategies Using Propensity Scores. *Ann. Intern. Med.* 137, 693–695.  
<https://doi.org/10.7326/0003-4819-137-8-200210150-00015>
- Carpenter, J., Bithell, J., 2000. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat. Med.* 19, 1141–1164.  
[https://doi.org/10.1002/\(SICI\)1097-0258\(20000515\)19:9<1141::AID-SIM479>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F)
- Chen, K., Bi, J., Chen, J., Chen, X., Huang, L., Zhou, L., 2015. Influence of heat wave definitions to the added effect of heat waves on daily mortality in Nanjing, China. *Sci. Total Environ.* 506–507, 18–25. <https://doi.org/10.1016/j.scitotenv.2014.10.092>
- Corso, M., Pascal, M., Wagner, V., 2017. Impacts de la chaleur et du froid sur la mortalité totale en France entre 2000 et 2010. *Bull. Epidemiol. Hebd.* 31, 634–40.
- Croissant, Y., 2020. *pglm: Panel Generalized Linear Models*.
- D'Ippoliti, D., Michelozzi, P., Marino, C., de'Donato, F., Menne, B., Katsouyanni, K., Kirchmayer, U., Analitis, A., Medina-Ramón, M., Paldy, A., Atkinson, R., Kovats, S., Bisanti, L., Schneider, A., Lefranc, A., Iñiguez, C., Perucci, C.A., 2010. The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project. *Environ. Health* 9, 37. <https://doi.org/10.1186/1476-069X-9-37>
- Dominici, F., Samet, J.M., Zeger, S.L., 2000. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *J. R. Stat. Soc. Ser. A Stat. Soc.* 163, 263–302. <https://doi.org/10.1111/1467-985X.00170>
- Gasparrini, A., Armstrong, B., Kenward, M.G., 2010. Distributed lag non-linear models. *Stat. Med.* 29, 2224–2234. <https://doi.org/10.1002/sim.3940>
- Gasparrini, A., Guo, Y., Hashizume, M., Kinney, P.L., Petkova, E.P., Lavigne, E., Zanobetti, A., Schwartz, J.D., Tobias, A., Leone, M., Tong, S., Honda, Y., Kim, H., Armstrong, B.G., 2015a. Temporal Variation in Heat–Mortality Associations: A Multicountry Study. *Environ. Health Perspect.* 123, 1200–1207. <https://doi.org/10.1289/ehp.1409070>

- Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B., Leone, M., Sario, M.D., Bell, M.L., Guo, Y.-L.L., Wu, C., Kan, H., Yi, S.-M., Coelho, M. de S.Z.S., Saldiva, P.H.N., Honda, Y., Kim, H., Armstrong, B., 2015b. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet* 386, 369–375. [https://doi.org/10.1016/S0140-6736\(14\)62114-0](https://doi.org/10.1016/S0140-6736(14)62114-0)
- Gayat, E., Pirracchio, R., Resche-Rigon, M., Mebazaa, A., Mary, J.-Y., Porcher, R., 2010. Propensity scores in intensive care and anaesthesiology literature: a systematic review. *Intensive Care Med.* 36, 1993–2003. <https://doi.org/10.1007/s00134-010-1991-5>
- Greifer, N., 2021. cobalt: Covariate Balance Tables and Plots.
- Gronlund, C.J., 2014. Racial and Socioeconomic Disparities in Heat-Related Health Effects and Their Mechanisms: a Review. *Curr. Epidemiol. Rep.* 1, 165–173. <https://doi.org/10.1007/s40471-014-0014-4>
- Guo, Y., Gasparrini, A., Armstrong, B., Li, S., Tawatsupa, B., Tobias, A., Lavigne, E., de Sousa Zanotti Stagliorio Coelho, M., Leone, M., Pan, X., Tong, S., Tian, L., Kim, H., Hashizume, M., Honda, Y., Guo, Y.-L.L., Wu, C.-F., Punnasiri, K., Yi, S.-M., Michelozzi, P., Saldiva, P.H.N., Williams, G., 2014. Global Variation in the Effects of Ambient Temperature on Mortality: A Systematic Evaluation. *Epidemiology* 25, 781–789. <https://doi.org/10.1097/EDE.0000000000000165>
- Guo, Y., Gasparrini, A., Armstrong, B.G., Tawatsupa, B., Tobias, A., Lavigne, E., Coelho, M. de S.Z.S., Pan, X., Kim, H., Hashizume, M., Honda, Y., Guo, Y.-L.L., Wu, C.-F., Zanobetti, A., Schwartz, J.D., Bell, M.L., Scortichini, M., Michelozzi, P., Punnasiri, K., Li, S., Tian, L., Garcia, S.D.O., Seposo, X., Overcenco, A., Zeka, A., Goodman, P., Dang, T.N., Dung, D.V., Mayvaneh, F., Saldiva, P.H.N., Williams, G., Tong, S., 2017. Heat Wave and Mortality: A Multicountry, Multicommunity Study. *Environ. Health Perspect.* 125, 087006. <https://doi.org/10.1289/EHP1026>
- Hansen Alana, Bi Peng, Nitschke Monika, Ryan Philip, Pisaniello Dino, Tucker Graeme, 2008. The Effect of Heat Waves on Mental Health in a Temperate Australian City. *Environ. Health Perspect.* 116, 1369–1375. <https://doi.org/10.1289/ehp.11339>
- Imbens, G.W., 2004. Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Rev. Econ. Stat.* 86, 4–29. <https://doi.org/10.1162/003465304323023651>
- Kent, S.T., McClure, L.A., Zaitchik, B.F., Smith, T.T., Gohlke, J.M., 2014. Heat Waves and Health Outcomes in Alabama (USA): The Importance of Heat Wave Definition. *Environ. Health Perspect.* 122, 151–158. <https://doi.org/10.1289/ehp.1307262>
- Kim, H., Lee, J.-T., Fong, K.C., Bell, M.L., 2021. Alternative adjustment for seasonality and long-term time-trend in time-series analysis for long-term environmental exposures and disease counts. *BMC Med. Res. Methodol.* 21, 2. <https://doi.org/10.1186/s12874-020-01199-1>



- Kovats, S., Ebi, K.L., Menne, B., 2003. Methods of assessing human health vulnerability and public health adaptation to climate change, Health and Global Environmental Change. WHO Europe.
- Laaidi, K., Pascal, M., 2004. Système d'alerte canicule et santé 2004 (Sacs 2004) : Rapport opérationnel. Institut de veille sanitaire, Saint-Maurice, France.
- Laaidi, K., Pascal, M., Baffert, E., Strauss, B., Ledrans, M., Empereur-Bissonnet, P., 2005. Système d'alerte canicule et santé 2005 (Sacs 2005) : Rapport opérationnel. Institut de veille sanitaire, Saint-Maurice, France.
- Le Tertre, A., Lefranc, A., Eilstein, D., Declercq, C., Medina, S., Blanchard, M., Chardon, B., Fabre, P., Filleul, L., Jusot, J.-F., Pascal, L., Prouvost, H., Cassadou, S., Ledrans, M., 2006. Impact of the 2003 Heatwave on All-Cause Mortality in 9 French Cities. *Epidemiology* 17, 75–79. <https://doi.org/10.1097/01.ede.0000187650.36636.1f>
- Littnerova, S., Jarkovsky, J., Parenica, J., Pavlik, T., Spinar, J., Dusek, L., 2013. Why to use propensity score in observational studies? Case study based on data from the Czech clinical database AHEAD 2006–09. *Cor Vasa, Heart Failure* 55, e383–e390. <https://doi.org/10.1016/j.crvasa.2013.04.001>
- McElroy, S., Schwarz, L., Green, H., Corcos, I., Guirguis, K., Gershunov, A., Benmarhnia, T., 2020. Defining heat waves and extreme heat events using sub-regional meteorological data to maximize benefits of early warning systems to population health. *Sci. Total Environ.* 721, 137678. <https://doi.org/10.1016/j.scitotenv.2020.137678>
- McGregor, G.R., Menne, B., 2015. Vagues de chaleur et santé: guide pour l'élaboration de système d'alerte, OMM. Organisation météorologique mondiale ; Organisation Mondiale de la Santé, Genève, Suisse.
- Muggeo, V.M., Hajat, S., 2009. Modelling the non-linear multiple-lag effects of ambient temperature on mortality in Santiago and Palermo: a constrained segmented distributed lag approach. *Occup. Environ. Med.* 66, 584–591. <https://doi.org/10.1136/oem.2007.038653>
- Nairn, J.R., Fawcett, R.J.B., 2015. The Excess Heat Factor: A Metric for Heatwave Intensity and Its Use in Classifying Heatwave Severity. *Int. J. Environ. Res. Public. Health* 12, 227–253. <https://doi.org/10.3390/ijerph120100227>
- Nitschke, M., Tucker, G.R., Bi, P., 2007. Morbidity and mortality during heatwaves in metropolitan Adelaide. *Med. J. Aust.* 187, 662–665. <https://doi.org/10.5694/j.1326-5377.2007.tb01466.x>
- Nori-Sarma, A., Benmarhnia, T., Rajiva, A., Azhar, G.S., Gupta, P., Pednekar, M.S., Bell, M.L., 2019. Advancing our Understanding of Heat Wave Criteria and Associated Health Impacts to Improve Heat Wave Alerts in Developing Country Settings. *Int. J. Environ. Res. Public. Health* 16. <https://doi.org/10.3390/ijerph16122089>
- Pascal, M., 2021. Quels indicateurs pour faciliter la prise en compte de la santé publique dans les politiques d'adaptation au changement climatique ?, Méthode. Santé publique France, Saint-Maurice, France.

- Pascal, M., Wagner, V., Corso, M., Laaidi, K., Le Tertre, A., 2019. Évolution de l'exposition aux canicules et de la mortalité associée en France métropolitaine entre 1970 et 2013. Santé publique France, Saint-Maurice, France.
- Pascal, M., Wagner, V., Corso, M., Laaidi, K., Ung, A., Beaudreau, P., 2018. Heat and cold related-mortality in 18 French cities. *Environ. Int.* 121, 189–198. <https://doi.org/10.1016/j.envint.2018.08.049>
- Pascal, M., Wagner, V., Le Tertre, A., Laaidi, K., Honoré, C., Bénichou, F., Beaudreau, P., 2013. Definition of temperature thresholds: the example of the French heat wave warning system. *Int. J. Biometeorol.* 57, 21–29. <https://doi.org/10.1007/s00484-012-0530-1>
- Peduzzi, P., Concato, J., Feinstein, A.R., Holford, T.R., 1995. Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* 48, 1503–1510. [https://doi.org/10.1016/0895-4356\(95\)00048-8](https://doi.org/10.1016/0895-4356(95)00048-8)
- Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R., 1996. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* 49, 1373–1379. [https://doi.org/10.1016/S0895-4356\(96\)00236-3](https://doi.org/10.1016/S0895-4356(96)00236-3)
- Peng, R.D., Dominici, F., Louis, T.A., 2006. Model choice in time series studies of air pollution and mortality. *J. R. Stat. Soc. Ser. A Stat. Soc.* 169, 179–203.
- Perkins, S.E., 2015. A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes at the global scale. *Atmospheric Res.* 164–165, 242–267. <https://doi.org/10.1016/j.atmosres.2015.05.014>
- Perkins, S.E., Alexander, L.V., 2013. On the Measurement of Heat Waves. *J. Clim.* 26, 4500–4517. <https://doi.org/10.1175/JCLI-D-12-00383.1>
- Perrakis, K., Gryparis, A., Schwartz, J., Tertre, A.L., Katsouyanni, K., Forastiere, F., Stafoggia, M., Samoli, E., 2014. Controlling for seasonal patterns and time varying confounders in time-series epidemiological models: a simulation study. *Stat. Med.* 33, 4904–4918. <https://doi.org/10.1002/sim.6271>
- R Core Team, 2022. R: A language and environment for statistical computing.
- Robinson, P.J., 2001. On the Definition of a Heat Wave. *J. Appl. Meteorol. Climatol.* 40, 762–775. [https://doi.org/10.1175/1520-0450\(2001\)040<0762:OTDOAH>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0762:OTDOAH>2.0.CO;2)
- Rocklöv, J., Forsberg, B., 2008. The effect of temperature on mortality in Stockholm 1998–2003: A study of lag structures and heatwave effects. *Scand. J. Public Health* 36, 516–523. <https://doi.org/10.1177/1403494807088458>
- Rosenbaum, P.R., Rubin, D.B., 1985. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *Am. Stat.* 39, 33–38. <https://doi.org/10.1080/00031305.1985.10479383>
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55. <https://doi.org/10.1093/biomet/70.1.41>

- Rubin, D.B., 2007. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat. Med.* 26, 20–36. <https://doi.org/10.1002/sim.2739>
- Schwartz, J., Zanobetti, A., Bateson, T., 2003. Morbidity and mortality among elderly residents in cities with daily PM measurements. *Revis. Anal. Time-Ser. Stud. Air Pollut. Health* 25–58.
- Schwarz, L., Hansen, K., Alari, A., Ilango, S.D., Bernal, N., Basu, R., Gershunov, A., Benmarhnia, T., 2021. Spatial variation in the joint effect of extreme heat events and ozone on respiratory hospitalizations in California. *Proc. Natl. Acad. Sci.* 118. <https://doi.org/10.1073/pnas.2023078118>
- Sekhon, J.S., 2011. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching package for R. *J. Stat. Softw.* 42, 1–52. <https://doi.org/10.18637/jss.v042.i07>
- Soubeyroux, J.-M., Bernus, S., Corre, L., Drouin, A., Dubuisson, B., Etchevers, P., Gouget, V., Josse, P., Kerdoncuff, M., Samacoits, R., Tocquer, F., 2020. Les nouvelles projections climatiques de référence DRIAS-2020 pour la métropole. Météo-France, Cerfacs, CNRM, IPSL, Paris, France.
- Soubeyroux, J.-M., Ouzeau, G., Schneider, M., Cabanes, O., Kounkou-Arnaud, R., 2016. Les vagues de chaleur en France : analyse de l'été 2015 et évolutions attendues en climat futur. *La Météorologie* 8, 45. <https://doi.org/10.4267/2042/60704>
- Tong, S., Wang, X.Y., Barnett, A.G., 2010. Assessment of heat-related health impacts in Brisbane, Australia: comparison of different heatwave definitions. *PloS One* 5, e12155. <https://doi.org/10.1371/journal.pone.0012155>
- Ung, A., Corso, M., Pascal, M., Laaidi, K., Wagner, V., Beaudeau, P., 2019. Évaluation de la surmortalité pendant les canicules des étés 2006 et 2015 en France métropolitaine. *Santé publique France, Saint-Maurice, France.*
- Vaidyanathan, A., Saha, S., Vicedo-Cabrera, A.M., Gasparrini, A., Abdurehman, N., Jordan, R., Hawkins, M., Hess, J., Elixhauser, A., 2019. Assessment of extreme heat and hospitalizations to inform early warning systems. *Proc. Natl. Acad. Sci.* 116, 5420–5427. <https://doi.org/10.1073/pnas.1806393116>
- Vicedo-Cabrera, A.M., Guo, Y., Sera, F., Huber, V., Schleussner, C.-F., Mitchell, D., Tong, S., Coelho, M. de S.Z.S., Saldiva, P.H.N., Lavigne, E., Correa, P.M., Ortega, N.V., Kan, H., Osorio, S., Kyselý, J., Urban, A., Jaakkola, J.J.K., Rytí, N.R.I., Pascal, M., Goodman, P.G., Zeka, A., Michelozzi, P., Scortichini, M., Hashizume, M., Honda, Y., Hurtado-Díaz, M., Cruz, J., Seposo, X., Kim, H., Tobias, A., Íñiguez, C., Forsberg, B., Åström, D.O., Ragettli, M.S., Röösli, M., Guo, Y.L., Wu, C., Zanobetti, A., Schwartz, J., Bell, M.L., Dang, T.N., Do Van, D., Heaviside, C., Vardoulakis, S., Hajat, S., Haines, A., Armstrong, B., Ebi, K.L., Gasparrini, A., 2018. Temperature-related mortality impacts under and beyond Paris Agreement climate change scenarios. *Clim. Change* 150, 391–402. <https://doi.org/10.1007/s10584-018-2274-3>

- Viechtbauer, W., 2010. Conducting Meta-Analyses in R with the metafor Package. *J. Stat. Softw.* 36, 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Belesova, K., Berry, H., Bouley, T., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Chambers, J., Daly, M., Dasandi, N., Davies, M., Depoux, A., Dominguez-Salas, P., Drummond, P., Ebi, K.L., Ekins, P., Montoya, L.F., Fischer, H., Georgeson, L., Grace, D., Graham, H., Hamilton, I., Hartinger, S., Hess, J., Kelman, I., Kieseewetter, G., Kjellstrom, T., Kniveton, D., Lemke, B., Liang, L., Lott, M., Lowe, R., Sewe, M.O., Martinez-Urtaza, J., Maslin, M., McAllister, L., Mikhaylov, S.J., Milner, J., Moradi-Lakeh, M., Morrissey, K., Murray, K., Nilsson, M., Neville, T., Oreszczyn, T., Owfi, F., Pearman, O., Pencheon, D., Pye, S., Rabbaniha, M., Robinson, E., Rocklöv, J., Saxer, O., Schütte, S., Semenza, J.C., Shumake-Guillemot, J., Steinbach, R., Tabatabaei, M., Tomei, J., Trinanes, J., Wheeler, N., Wilkinson, P., Gong, P., Montgomery, H., Costello, A., 2018. The 2018 report of the Lancet Countdown on health and climate change: shaping the health of nations for centuries to come. *The Lancet* 392, 2479–2514. [https://doi.org/10.1016/S0140-6736\(18\)32594-7](https://doi.org/10.1016/S0140-6736(18)32594-7)
- Weitzen, S., Lapane, K.L., Toledano, A.Y., Hume, A.L., Mor, V., 2004. Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiol. Drug Saf.* 13, 841–853. <https://doi.org/10.1002/pds.969>
- Wood, S.N., 2017. *Generalized Additive Models: An Introduction with R*, 2nd ed. Chapman and Hall/CRC, New York. <https://doi.org/10.1201/9781315370279>
- Xu, Z., Etzel, R.A., Su, H., Huang, C., Guo, Y., Tong, S., 2012. Impact of ambient temperature on children's health: A systematic review. *Environ. Res.* 117, 120–131. <https://doi.org/10.1016/j.envres.2012.07.002>
- Xu, Z., FitzGerald, G., Guo, Y., Jalaludin, B., Tong, S., 2016. Impact of heatwave on mortality under different heatwave definitions: A systematic review and meta-analysis. *Environ. Int.* 89–90, 193–203. <https://doi.org/10.1016/j.envint.2016.02.007>
- Yu, W., Mengersen, K., Wang, X., Ye, X., Guo, Y., Pan, X., Tong, S., 2012. Daily average temperature and mortality among the elderly: a meta-analysis and systematic review of epidemiological evidence. *Int. J. Biometeorol.* 56, 569–581. <https://doi.org/10.1007/s00484-011-0497-3>