Advancing understanding of climate's impacts on individual wellbeing and social justice

Abstract

Climate change interacts with other environmental stressors to create and potentially exacerbate diverse hazards across geographies and socio-economic conditions – rendering some people and places far more vulnerable than others. Current assessments on the associations between the environment and wellbeing rely on coarse, aggregated data that do not reflect these distinctions, leaving a gap in understanding about whether proposed climate actions can achieve fair and equitable outcomes for different groups. We empirically demonstrate, with our new, publicly accessible data, that using more granular longitudinal data can advance understanding about climate impacts and shed new light on the effectiveness and fairness of climate actions and policies.

Main text

The Glasgow Climate Pact adopted at the 26th United Nations Conference of Parties (COP26) climate conference calls for an improved understanding of the geography of climate change impacts, of the related adaptation needs, and of response options. Climate and environmental risk affect people in different ways, depending on the context in which they live and on their own, individual characteristics. (Hsiang, Oliva, and Walker 2019; Vona 2021).

Impact and vulnerability analyses conducted at the territorial level provide important insights on the regional dimensions of climate and environmental vulnerability. Yet, aggregate studies fail to inform the environmental justice debate (Banzhaf, Ma, and Timmins 2019) because they do not address how environmental risk affects the wellbeing of different groups within these wider geographies over time and across generations (Mitchell and Norman 2012, Ara Begum et al. 2022). Moreover, the studies cannot create the quasi-experimental settings needed to evaluate the effectiveness of adaptive behaviors, risk-avoidance efforts, or environmental regulations. Wellbeing is a complex and contested concept (Lamb and Steinberger 2017). Yet, health-related dimensions that incorporate physical health, mental health, perceived general health, and one's working environment are, unambiguously, some of its defining dimensions. Vulnerability links to numerous individual characteristics, among them age, gender, education, and socio-economic status, and to many health-related dimensions, such as pre-existing health conditions, lifestyles, and awareness of risk (**Figure 1**). Actions to reduce both exposure and vulnerability often involve a complex array of actions and situations, such as access to safe housing, access to and use of appropriate healthcare, and the ability to devote resources to medical expenditures in times of need. How these forces interact ultimately affects individuals' wellbeing (Jafino, Hallegatte, Rozenberg 2021).

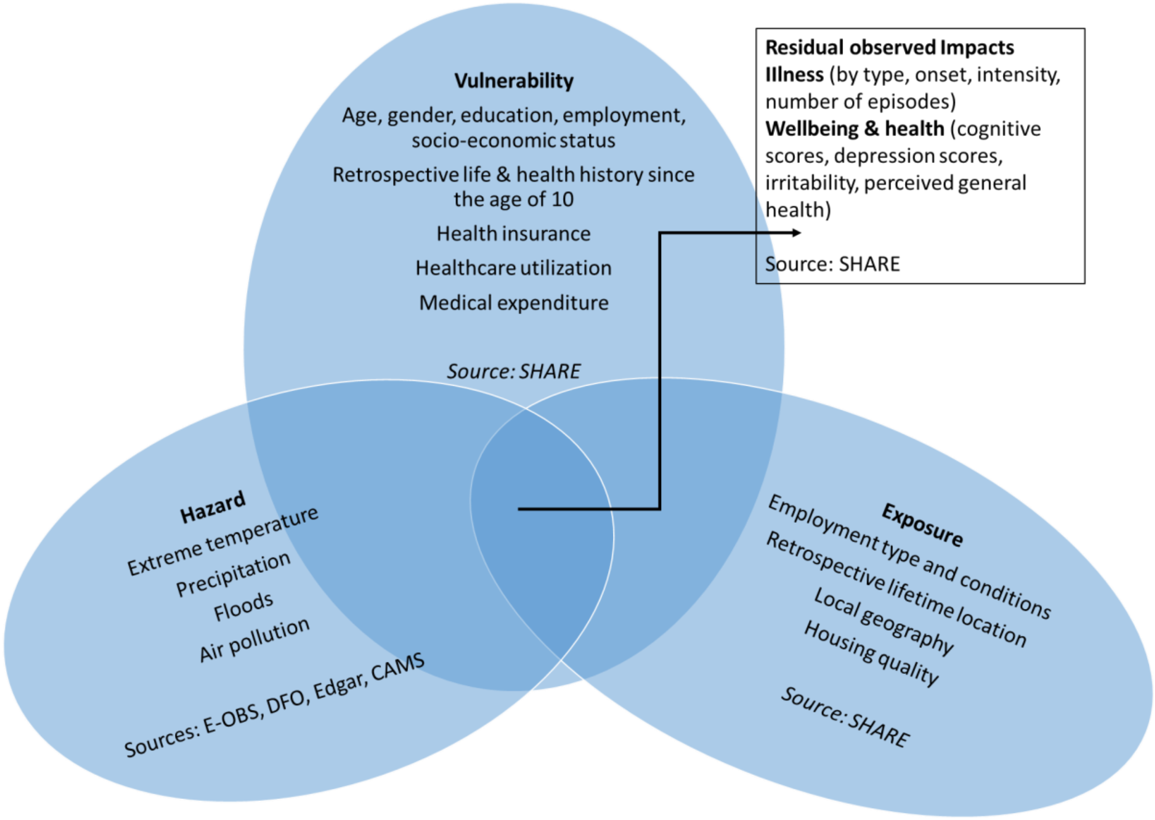


Figure 1 (TO BE EDITED FOR THE FINAL SUBMISSION). Interactions between environmental hazards, vulnerability, exposure to risk, and their impacts on health outcomes and wellbeing. The figure relies on data from the Survey of Health, Ageing and Retirement in Europe (SHARE). These data have been augmented with data from the Global Land Data Assimilation System (GLDAS), the Dartmouth Flood Observatory (DFO), the Emission Database for Global Atmospheric Research (EDGAR), and the Copernicus Atmosphere Monitoring Service (CAMS).

Here, we argue that more granular, longitudinal data would advance the causal assessment of the impacts of environmental risk and of the effectiveness of adaptation interventions, delivering the much-needed information and methods for evaluation of climate actions and the pursuit of climate justice (Breil et al. 2021, Ara Begum et al. 2022). Longitudinal studies following individuals over long periods of time can uncover causal relationships between exposure, vulnerability, and policy interventions or actions. The existing evidence in literature is still piecemeal and confined to a few regions, mostly in the United States (Zivin and Neidell 2016; Voorheis 2017; Zivin, Hsiang, and Neidell 2018; Sun et al. 2020; Park et al. 2020; Vona 2021).

We bring forward a new source of data, the augmented SHARE dataset, which expands on data from the European Union (EU)-funded initiative, the longitudinal Survey on Health, Ageing and Retirement in Europe (SHARE). We demonstrate that these data can be used to study relationships between environment and wellbeing, and therefore to advance the climate adaptation and climate policy literatures by expanding understanding of the links between climate change and human health. Since 2004, SHARE has regularly surveyed about 120,000 randomly selected individuals over the age of 50 in 28 EU countries. The regular panel waves (2004-2019) have followed these individuals and their spouses over time. Two specific interviews conducted in the third and seventh waves reconstruct the retrospective life history of those individuals, providing information on their early life conditions, health, healthcare access, and working lives. These histories include key focal points, such as the age at which a person left school, dates when the person started and ended a given job, dates of the onset of any illness, details about any changes in housing circumstances, and dates of the birth of any children. SHARE contains a rich inventory of types of morbidity and illnesses. It provides information on illness onset, frequency, and severity. Likewise, it provides numerous variables that can be used to construct subjective and objective indicators of wellbeing at home and at work (see Supplementary Information, SI). We provide open access to this extended SHARE database that has been expanded to include cumulative and yearly exposure to extreme temperatures, floods, solar radiation, and air pollution.

We use this augmented SHARE dataset as proof of concept of the potential of more detailed geographical and longitudinal health data to shed light on the differing ramifications of climate change for different populations. We use this data to analyse the following three issues: i) the effects of climate an illustrative type of illness, the *prevalence of breathlessness;* ii) the effects of climate on one's *perceived health status through life*; and iii) the effects of climate on the *perceived comfort of one's job*. We are cautious about statements of causality precisely because we know details of where people live only to the level of one's territorial region (i.e., the Nomenclature of Territorial Units for Statistics - NUTS1 and 2 depending on the country, see SI); thus, these data imply that all individuals who lived in the same regions throughout their lives face the same, average environmental conditions. Results are summarized in **Table 1** and presented in more detail in the Supplementary Table S4.

Prevalence of breathlessness

Having ever experienced breathlessness in one's lifetime is related to average exposure to pollution (concentrations and emissions of fine particulate matter, PM2.5), but the relative impact of actual exposure grows once one accounts for other, relevant, individual-level variables: average income, age, smoking history, body max index (BMI), frequency with which the individual practices sports (1=more than once a week, 2=once a week, 3=one to three times a month, 4=hardly ever, or never), whether the individual's job is uncomfortable (1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree), and whether the individual had any illness at birth. This illustrates how considering individual characteristics can help uncover the expected associations between environmental hazards and poor health.

Perceived health status

Cumulative exposure to extreme temperatures affects one's *perceived health status* differently, depending on when in one's lifetime the question is posed. This measure provides a subjective assessment by individuals of their health status, which is ranked between excellent (=5) and poor (=1). Exposure to both extremely high and extremely low temperatures is associated with worse perceived physical health in old age. When we consider only the information provided in the first wave of individual interviews, only *extremely low* temperatures are significantly associated with worse health. By contrast, when we consider the most recent wave, in which individuals are considerably older (69 years old on average, 6 years older than the average age in their first wave), only *extremely high* temperatures are significantly associated with worse health status (see Supplementary Table S5). This example shows how such geographically granular, longitudinal data (such as those we leverage through the augmented SHARE database) can distinguish among the different associations that can emerge between climate and human health throughout one's life.

As another example, in the context of *perceived health status,* we interact the two extreme-temperature-exposure variables with adaptation or risk-avoiding behavior: having *air-conditioning (AC)* and *central heating*. We find that extremely low temperatures are associated with worse health outcomes across all specifications. At extremely high temperatures there is an apparent, protective effect of having air-conditioning (AC) – that is, the effects of extremely high temperatures are less pronounced for individuals who have AC. However, once we consider education levels, this protective effect markedly wanes (becoming nonsignificant at the 10% level at average temperatures above 27.5°C). Higher education might be associated with improved information on the possible negative effects of extreme temperatures; thus, individuals might acquire air-conditioning as one measure among other protective steps they take. This example illustrates the potential for more granular data to help provide a source of information for evaluating the effectiveness of individual adaptation actions. Indeed, in this regard the augmented SHARE dataset provides underpinning material that is almost completely missing in the literature, with only a few exceptions in the United States, (Barreca et al. 2016).

Another illustration concerns *perceived health in young age* (until the age of 15), which is positively related to exposure to more frequent extreme high temperatures. Such relationship remains equally strong once we consider the significant positive effect of average solar radiation, (positively correlated with high temperature extremes). A possible channel through which frequent high temperatures might have a positive impact on young age health is by allowing children to engage in more activities outdoors, a behaviour we do not observe. Our augmented dataset makes it also possible to put the role of environmental hazards into perspective and offer a relative assessment in comparison to other important variables. The magnitude of the association between higher temperatures/higher average solar radiation and improved health is two orders of magnitude smaller than the association between one's health status and any of the following: ever being poor in childhood; ever having experienced physical harm; or ever living in a childhood home that lacked a fixed bath, running water supply, inside toilet, or central heating (Supplementary Table 4).

Perceptions about whether one's job is uncomfortable

As a final illustration, we investigate how average temperatures and average solar radiation affect the perception of *having an uncomfortable job*. This analysis can be of particular interest as a first-stage regression in the context of how exposure to environmental hazards at work can lead to loss of productivity, the onset of illness, and/or early retirement, a less-considered aspect of labor-output loss. We show that higher summer temperatures and higher summer radiation averages are associated with a higher probability of stating that one's job is uncomfortable. In winter, temperatures that are milder/less cold are associated with a lower probability of having a job perceived as uncomfortable. The effect of exposure to these environmental hazards is, as expected, more pronounced in jobs which are physically demanding.

Potential for new insights

Geographically localized, longitudinal data would open new opportunities to advance the climate adaptation literature and to respond to the mission of climate justice. As outlined and illustrated in this brief, the use of the extended SHARE database for analyses of climate impacts underscores the potential of more granular data and longitudinal panels to characterize the potential vulnerabilities of different populations; in turn, this provides an opportunity to better understand whether climate actions and policies are likely to be fair and equitable in terms of their impacts. Existing and regularly updated initiatives – not limited to SHARE – could easily move in this direction. Providing the location of individuals with a buffer zone of a few kilometres would ensure compliance with privacy regulations and open promising, new research avenues

Table 1. Exploring the association between environmental hazards, health outcomes, and risk-avoiding behaviors

	1. Ever experienced breathlessness (0=no, 1=yes)		2. Young age (<15) perceived reported health (1=poor; 5=excellent)		3. Old age (>49) perceived reported health (1=poor; 5=excellent)			4. Uncomfortable job (0=no, 1=yes)	
Exposure	Avg. pm2.5 conc. median (μg/m3)	0.003*** (0.0007)	Avg. first 15 years exposure to negative temperature (# days)	+0.0001 (0.0004)	Avg. lifetime exposure to temperature > 30°C (# days)	-0.002** (0.0008)	-0.001 (0.001)	Average winter temperature	0.007 (0.001)
	Avg. cum. lifetime exposure to negative temperature (# days)	2.48e-05 (9.09e-05)	Avg. first 15 years exposure to temperature > 30°C (# days)	+0.003* (0.002)	Avg. cum. lifetime exposure to negative temperature (# days)	-0.001*** (0.0004)	-0.001*** (0.0004)	Average summer temperature	-0.003*** (0.001)
	Avg. lifetime exposure to temperature > 30°C (# days)	-0.0004* (0.0002)	Avg. first 15 years solar radiation (W/m^2)	+0.002* (0.001)				Average radiation	0.0001 (0.0004)
Exposure x Individual characteristics or risk-avoiding behaviors					AC x avg. Lifetime exposure to temperature > 30°C (# days)	+0.002*** (0.0006)	+0.002** (0.0007)	Job is physical x average winter temperature	-0.011*** (0.002)
					Central heating x avg. cum. lifetime exposure to negative temperature (# days)	+0.0004 (0.0005)	+0.0007 (0.0005)	Job is physical x average summer temperature	0.008*** (0.001)
					AC, central heating	not significant	not significant	Job is physical x average radiation	0.001*** (0.0003)
Individual characteristics	Current age; ever smoked; BMI; sports hardly ever;	+	Mother + father ISCED-1997 education; N. rooms/people living in the house; year of birth	-	Education (ISCED level 1 to 6)		-	Job is physical	-
	Average income; agree job is uncomfortable; born with any illness;	-	Ever experienced physical harm, ever lonely, ever poor, born with any illness, no feature, % of time in a suburb of big city	+	Income, wealth, ever smoked, BMI, sports frequency, depression scores, year of birth, female, job comfort	Y	Y	Average income (euros PPP)	-

Notes: Linear probability models. Models 1 and 2 include country fixed effects. Model 2 includes fixed effects of the International Standard Classification of Occupations (ISCO) (at the one-digit level). Model 3 includes year and country fixed effects. Model 4 includes fixed effects for the ISCO (at the one-digit level), the International Standard Classification of Education (ISCED), and the country. The corresponding questions to outcome variables 1 to 4 are the answers to the following questions or statement. *Ever experienced breathlessness*: "For the past six months at least, have you been bothered by any of the health conditions of breathlessness or difficult breathing?" Responses indicate whether they selected this symptom in any survey wave. *Young age perceived health:* "Would you say that your health during your childhood was in general excellent, very good, good, fair, or poor?" Responses were coded as follows: excellent 5, very good 4, good 3, fair 2, poor 1. *Old age perceived health:* "Would you say your health is Excellent, Very Good, Good, Fair, (or) Poor?" Responses were coded as follows: excellent 5, very good 4, good 3, fair 2, poor 1. *Uncomfortable job:* "My immediate work environment was uncomfortable (for example, because of noise, heat, crowding)." Answers were coded as follows: one ((has an uncomfortable job) for "Strongly Agree" or "Agree;" zero (does not have an uncomfortable job) for "Disagree" and "Strongly Disagree."

Acknowledgments

This research received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant agreement No. 756194, "ENERGYA," and from the European Union's Horizon 2020 under the Marie Skłodowska-Curie grant agreement No 956107, "Economic Policy in Complex Environments (EPOC)." This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7,(DOIs: 10.6103/SHARE.w1.800, 10.6103/SHARE.w2.800, 10.6103/SHARE.w3.800, 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800. The SHARE data collection was funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Contributions

C.M. and E.D.C designed and performed the analysis. C.M., S.P. processed climate data under the supervision of M.N.M. and E.D.C.; E.D.C wrote the first draft of the paper. All contributed to editing the paper.

Methods

Environmental data

We compiled a database that combines a set of environmental hazards (extreme temperatures, solar radiation exposure, heavy precipitation, and air pollution hazards related to ozone, nitrogen dioxide, and particulate matter (PM2.5 and PM10)) with a comprehensive set of variables about individual-level health (characterizing individuals' health status, behavioral risks, and risk-averting behaviors at different points in life in Europe from the SHARE database, a longitudinal random sample of Europeans aged 50+ for 28 countries (Börsch-Supan et al. 2013). The Supplementary Information (SI) lists examples of the morbidity variables that can be used to study the relationship between health outcomes and environmental exposure.

The environmental-hazard variables included in our database have been assembled using the high-resolution, daily, near-surface temperature and precipitation gridded-observational data (E-OBS, version 19.0e) made available by the European Climate Assessment & Dataset (ECA&D) at 0.1° X0.1° resolution (~10km x 10km at the equator) (Haylock, M. R. 2008; Cornes et al. 2018). We use bins of daily mean temperature (TG), daily minimum temperature (TN) and daily maximum temperature (TX) (see SI). Because a substantial number of individuals in SHARE were born before 1950, we construct the variables for environmental exposure over one's full lifetime and early life by extending the time frame for variables back to the 1920s. Additional variables that in our dataset include average seasonal temperature, heating and cooling degree days, and solar radiation. We also include floods events from the Dartmouth Flood Observatory (DFO) database (Brakenridge 2021, see SI for further details on the individual variables and data sources). Air pollution data are from: (i) the Copernicus Atmosphere Monitoring Service (CAMS) global reanalysis (EAC4) monthly averaged fields (Inness et al. 2019) made available by the Copernicus Atmosphere Data Store (https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4?tab=overview), and (ii) the Emissions Database for Global Atmospheric Research (EDGAR) ver 5.0 (Crippa et al 2019) covering the period from 1970 to 2015 , made available by the European Commission Joint Research Centre (JRC) (https://data.jrc.ec.europa.eu/dataset/377801af-b094-4943-8fdc-f79a7c0c2d19). We also construct cumulative variables of exposure to environmental hazards, reflecting not the exposure to, for instance, extreme temperatures in the year of a wave, but instead exposure that had occurred from the time an individual was born until the wave in question.

The SHARE database

SHARE regions are those that respondents indicate as the location of their accommodation in the retrospective accommodation wave (SHARELIFE waves, or the NUTS in which the household was located at the moment of sampling). The regions from the retrospective waves most often have a direct correspondence to NUTS regions (with certain exceptions; for instance, for Luxembourg, respondents indicate the canton). Whenever regions were the combination of two NUTS regions, we reported the environmental-hazard variables at the level of those combined regions. We then detect in which SHARE region the grid cells are located by overlaying them with a shapefile of the SHARE regions, constructed resorting to EUROSTAT NUTS shapefiles (downloadable from EUROSTAT) and to a shapefile of Luxembourg cantons (downloadable from data.public.lu, see SI for more details on the NUTS classifications used.) Depending on the country, SHARE regions are either NUTS1 (Belgium, France, Germany and one region of Hungary, Central Hungary), NUTS2 (Austria, Bulgaria, Croatia, Czechia, Denmark, Finland, Latvia, Lithuania, Greece, Hungary (except for Budapest and Pest, which are reported together as the NUTS1 region of Central Hungary), Italy, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden).

The resulting database consists of seven modules:

1. yearly_module: yearly exposure in year of wave (and one and two years before the wave) – unique identifier {individual,wave}
2. individual_year_panel: yearly exposure in years since birth up to the most recent participation in SHARE – unique identifier {individual,year}
3. life_module: rolling exposure throughout life – unique identifier {individual,wave}
4. young_age_module: cumulative exposure over the first five, ten and fifteen years of life – unique identifier {individual}
5. job_module: cumulative exposure during the years at one's most recent job – unique identifier {individual}
6. illness_before_module: cumulative exposure during one-, three-, and five-year periods before the onset of illness – unique identifier {individual}
7. illness_during_module: rolling exposure during periods of illness – unique identifier {individual,wave}. Variables differ between waves only for individuals for whom the illness period intersects with the SHARE interview period.

Three of these (the yearly_module, individual_year_panel, and illness_during_module) are longitudinal.

The first module, “yearly_module”, refers to yearly variables (i.e., environmental-hazard exposure in a specific year, as opposed to cumulative exposure or averages over longer time periods). For each individual-wave observation, we report environmental-hazard exposure in the year of that wave, in the year before, and in the year two years before, signaled by suffixes “t0”, “t_1bf,” and “t_2bf,” respectively. Such module only provides information on the waves in which respondents participated (alongside the information from one year and two years immediately prior to those waves).

The second module, individual_year_panel, has the same yearly variables. The difference, as the name indicates, is that it is not merged with current-wave information and, instead, provides a full individual-year panel for the period from birth until most recent participation in SHARE. This dataset can be of particular interest to be merged with other retrospective modules of SHARE, such as the jobs-episode module. A long-term longitudinal analysis is then feasible.

Examples in this article use three of the seven different modules: the young_age module, the job_module and the life_module, respectively. These examples should not be considered as encompassing analysis of research questions. For collecting sound empirical evidence, we recommend exploring the panel component of the longitudinal augmented SHARE modules, through modules one, two, three, and seven.

References

Ara Begum, Rawshan et al. (2022). Chapter 1: Point of Departure and Key Concepts. IPCC WGII Sixth Assessment Report

Banzhaf, Spencer, Lala Ma, and Christopher Timmins. 2019. “Environmental Justice: The Economics of Race, Place, and Pollution.” *Journal of Economic Perspectives* 33 (1): 185–208. https://doi.org/10.1257/jep.33.1.185.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro. 2016. “Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century.” *Journal of Political Economy* 124 (1): 105–59. https://doi.org/10.1086/684582.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., ... & Zuber, S. (2013). Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). International journal of epidemiology, 42(4), 992-1001.

Breil, Margaretha, Marianne Zandersen, Petya Pishmesheva, Anders Branth Pedersen, Linda Romanovska, Ingrid Coninx, Magdalena Rogger, and Katie Johnson. 2021. “‘Leaving No One Behind’ in Climate Resilience Policy and Practice in Europe Overview of Knowledge and Practice for Just Resilience,” 108. https://doi.org/10.25424/cmcc/justtrans.

Cornes, Richard C., Gerard van der Schrier, Else J.M. van den Besselaar, and Philip D. Jones. 2018. “An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets.” *Journal of Geophysical Research: Atmospheres* 123 (17): 9391–9409. https://doi.org/10.1029/2017JD028200.

Haylock, M. R., et al. 2008. “A European Daily High-resolution Gridded Data Set of.Pdf.” *Journal of Geophysical Research Atmospheres*.

Hsiang, Solomon, Paulina Oliva, and Reed Walker. 2019. “The Distribution of Environmental Damages.” *Review of Environmental Economics and Policy* 13 (1): 83–103. https://doi.org/10.1093/reep/rey024.

Inness, Antje, Melanie Ades, Anna Agustí-Panareda, Jérôme Barr, Anna Benedictow, Anne Marlene Blechschmidt, Juan Jose Dominguez, et al. 2019. “The CAMS Reanalysis of Atmospheric Composition.” *Atmospheric Chemistry and Physics* 19 (6): 3515–56. https://doi.org/10.5194/acp-19-3515-2019.

Kahn, Matthew E. 2016. “The Climate Change Adaptation Literature.” *Review of Environmental Economics and Policy* 10 (1): 166–78. https://doi.org/10.1093/reep/rev023.

Lamb, William F., and Julia K. Steinberger. 2017. “Human Well-Being and Climate Change Mitigation.” *Wiley Interdisciplinary Reviews: Climate Change* 8 (6): 1–16. https://doi.org/10.1002/wcc.485.

Mitchell, Gordon, and Paul Norman. 2012. “Longitudinal Environmental Justice Analysis: Co-Evolution of Environmental Quality and Deprivation in England, 1960-2007.” *Geoforum* 43 (1): 44–57. https://doi.org/10.1016/j.geoforum.2011.08.005.

Park, R. Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. “Heat and Learning.” *American Economic Journal: Economic Policy* 12 (2). https://doi.org/10.1257/POL.20180612.

Sun, Yi, Sindana D. Ilango, Lara Schwarz, Qiong Wang, Jiu Chiuan Chen, Jean M. Lawrence, Jun Wu, and Tarik Benmarhnia. 2020. “Examining the Joint Effects of Heatwaves, Air Pollution, and Green Space on the Risk of Preterm Birth in California.” *Environmental Research Letters* 15 (10). https://doi.org/10.1088/1748-9326/abb8a3.

Vona, Francesco. 2021. “Managing the Distributional Effects of Environmental and Climate Policies : The Narrow Path for a Triple Dividend.” Environment Working Paper No. 188.

Voorheis, John. 2017. “Longitudinal Environmental Inequality and Environmental Gentrification: Who Gains From Cleaner Air?” *CARRA Working Paper Series* 04. https://www.census.gov/content/dam/Census/library/working-papers/2017/adrm/carra-wp-2017-04.pdf.

Zivin, Joshua Graff, Solomon M. Hsiang, and Matthew Neidell. 2018. “Temperature and Human Capital in the Short and Long Run.” *Journal of the Association of Environmental and Resource Economists* 5 (1): 77–105. https://doi.org/10.1086/694177.

Zivin, Joshua Graff, and Matthew Neidell. 2016. “Environment , Health , and Human Capital” 51 (3): 689–730.

Supplementary Information

Environmental Variables

a. Climate data

Temperature Bins

For yearly measures of the full temperature distribution, we focus on bins of temperature, i.e., the number of days in a year where the minimum (TN), mean (TG) and maximum (TX) temperature fall in one of the sixteen 2.5°C temperature intervals: <-5, -5 to -2.5, -2.5 to 0, 0 to 2.5, 2.5 to 5, 5 to 7.5, 7.5 to 10, 10 to 12.5, 12.5 to 15, 15 to 17.5, 17.5 to 20, 20 to 22.5, 22.5 to 25, 25 to 27.5, 27.5 to 30 and > 30, computed at the grid cell level. The use of temperature bins allows flexibility in considering the non-linear impacts of temperature on health and other variables of interest. We then assign the grid cells to the SHARE regions by employing a shapefile of the SHARE regions and geospatial routines from R packages sf and raster, constructed resorting to EUROSTAT NUTS shapefiles (downloadable from EUROSTAT) and to a shapefile of Luxembourg cantons (downloadable from data.public.lu). Depending on the country, SHARE regions are either NUTS1 (Belgium, France, Germany and one region of Hungary, Central Hungary), NUTS2 (Austria, Bulgaria, Croatia, Czechia, Denmark, Finland, Greece, Hungary except for Budapest and Pest, which are reported together as the NUTS1 region of Central Hungary, Italy, Latvia, Lithuania, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden).[1] Once the bins are computed at grid cell level and georeferenced to a SHARE region, we aggregate them into two regional measures: median and mean. We also calculate the standard deviation between the cells of a SHARE region, given that, especially for larger regions, spatial variability might be substantial. Accordingly, the variable names end with '_median', '_mean' or '_std'.

Average (seasonal) temperature

We calculate the average annual temperature and the average seasonal temperatures – spring, summer, fall and winter in the SHARE region where the respondent lived in a certain year. These are calculated for each grid cell as the average of the mean temperature in all days of the year, or in the days pertaining to each season (December, January and February were allocated to winter; March, April and May to spring; June, July and August to summer; and September, October and November to fall). These grid cells values are aggregated to the SHARE region through both the median and the mean.

Heating and Cooling Degree days

[1] Wave 3 was conducted in almost all countries in 2008/2009 while Wave 7 was conducted in 2017. This would, at first, lead us to use NUTS2006 and NUTS2016 respectively. In practice, the SHARE regions indicated by respondents are consistent from Wave 3 to Wave 7, i.e., they do not change even if there were changes in the NUTS structure. France and Poland are the two examples – there are region changes in the NUTS, but not in the SHARE regions, which remain with a direct correspondence of names to NUTS2006. Therefore, for the two countries, we resort to NUTS2006 (shapefile NUTS2013 since there was no change to the NUTS boundaries of the two countries from NUTS2006 to NUTS2013). For the remaining countries, we resort to the NUTS2016 shapefile.

Following the EUROSTAT definitions (https://ec.europa.eu/eurostat/cache/metadata/en/nrg_chdd_esms.htm), at each grid cell we calculate the number of heating degree days (HDD) and cooling degree days (CDD) using the E-OBS dataset. Thus, for HDD, we sum over a year, for each gridcell, the differences between 18ºC and the recorded mean daily temperatures, for every day when the temperature in that grid cell was equal or below 15ºC (average temperature coming from TG variable of E-OBS). For CDD, the process is analogous, except we sum the differences between the recorded mean daily temperature and 21ºC, only for those days where the mean temperature was above 24ºC.

Each grid cell thus has, for each year, an HDD and a CDD index. These are aggregated to the SHARE regions through both the median and the mean, as with the remaining variables.

- Radiation

The 0.1° gridded E-OBS dataset provides data on daily radiation starting in 1950 through variable QQ. For each grid cell, we calculate for any given year, the average of the radiation over all the days in that year, or in the days pertaining to each season. These grid cells values are aggregated to the SHARE region through both the median and the mean.

- Precipitation

For precipitation we likewise provide yearly variables and cumulative variables calculated from them, starting from the E-OBS dataset, resorting to daily near-surface precipitation (RR). At each grid cell, we calculate, the number of days in each year where the sum of precipitation exceeds 10 mm and 20 mm - heavy and very heavy precipitation days-, as defined in the Agroclimatic indicators datasets part of the C3S Global Agriculture Sectoral Information Systems (SIS). As with temperature variables, these are georeferenced with SHARE regions, and aggregated using the median and mean, alongside the standard deviation to analyse intra-region variation.

b. Flood events

For floods, we resort to the DFO dataset (Brakenridge 2021), which provides information on flood events from 1985 until the present. We report 6 variables: the number of flood events, the number of casualties, the number of displaced individuals, a weighted number of flood events (weighted by an indicator 1, 1.5 or 2, representing the severity of the flood event), the total days during which there were floods events, and the weighted total days (weighted by an indicator 1, 1.5 or 2 representing the severity of the flood event).

The variables correspond to whether the individual was living in a region considered in the dataset to be affected by the flood event (more specifically, if the region where the individual was living overlaps with the region provided as 'affected' in the DFO dataset). Since depending on the country, individuals might report a NUTS2 or NUTS1 region, other 12 variables are created. The first 6 refer to whether the NUTS1 region where the individual resided was affected by flood events and the latter 6 to whether the NUTS2 region where the individual resided was affected by flood events.

c. Pollution data

The variables considered for pollution relate to the four most explored pollutants in the context of health: particulate matter 2.5 microns (in diameter) (PM2.5), particulate matter 10 microns (PM10), ozone (O_3) and nitrogen dioxide (NO_2) (as put forward in the WHO Review of evidence on health aspects of air pollution (WHO 2014).

Concentration

For PM2.5, PM10 and NO_2, there is limited evidence for the existence of a threshold below which health effects are negligible. Negative health outcomes have been found at very low concentrations (WHO 2014). We therefore resort to yearly average exposures, starting from the dataset CAMS global reanalysis (EAC4) on monthly averaged fields whose first year is 2003.

The original CAMS EAC4 monthly dataset resolution is 0.75° X 0.75°. We disaggregate the dataset into 0.1° X 0.1° through bilinear interpolation, and, at the grid-cell level, take the average of the 12 months of each year. As done with the temperature dataset, each grid-cell is associated with the SHARE region when its centroid falls within the region boundary, and the three variables: mean, median and standard deviation, are then constructed.

For O_3, the literature documents mixed evidence on the existence of thresholds. Several papers find an association between health outcomes and summer ozone concentration, but not winter season concentration; a finding attributed to the existence of a threshold by some studies or due to confounding effects or seasonal behavioural differences (Gryparis et al. (2004). Other studies that specifically analyse the threshold question arrive to different conclusions (e.g., evidence of thresholds is found in Kim et al. (2004) but not in Bell et al. (2006)). We follow the recent literature on long-term effects of ozone exposure and operate with yearly averages of daily maxima and warm-season averages of daily maxima (Kazemiparkouhi et al. (2020), Lim et al. (2019), Malley et al. (2017)). The dataset used is CAMS EAC4[2], whose first year is 2003, from which we use the average O_3 concentration at 3-hour intervals of each day, at the surface level. For each day, we keep the maximum of the 6 observations reported, at the grid-cell level (after disaggregating the spatial resolution from the gridded 0.75° to 0.1° as mentioned above). We then take either the yearly average or the warm months average (April to September) of the daily maxima, for each grid cell. The grid cells are overlapped with the SHARE regions, as with the temperature datasets, and we calculate the mean, median, and standard deviation at the SHARE region level.

Emissions

The datasets on pollution concentration mentioned begin in 2003 (or in 2004 for O_3), thus, enabling coverage for the regular SHARE waves (which start in 2004), but not for the cumulative exposure. To allow us to go further back in time we use a dataset not on pollution concentration, but on pollutant emissions, the EDGAR v5.0 Global Air Pollutant Emissions[3] dataset, which covers the period 1970-2015 (Crippa et al 2019). The relevant variable for direct health effects is concentration, thus, the health impacts of emissions will be different across regions, depending, namely, on meteorological conditions and topography. Even so, especially given that emissions are the variable which can be affected policy-wise, considering their (indirect) effects on other variables can be of interest. The variables obtained from EDGAR are estimates of yearly emissions of PM2.5 and PM10 at the grid cell level which are overlapped with SHARE regions to obtain yearly mean, median and standard deviation at the region level. Information on concentration could also be derived from the EDGAR dataset if

[2] Inness et al. (2019), http://www.atmos-chem-phys.net/19/3515/2019/
[3] https://edgar.jrc.ec.europa.eu/gallery?release=v50_AP&substance=PM10&sector=TOTALS

combined with advanced chemical transport models (CTMs). The resolution of the original dataset is available a 0.1° X0.1° resolution which we then aggregated through mean, median and standard deviation to the SHARE regions.

d. Cumulative variables

The SHARE dataset is a panel dataset, though only covering individuals aged 50 and above. Environmental hazards might have a cumulative impact on health. Situations which took place at a young age might also only later on transpire into health consequences.

We therefore construct cumulative variables of exposure to environmental hazards, reflecting not the exposure to for instance extreme temperatures in the year of a wave, but instead exposure since an individual was born until the wave in question, amongst other cumulative indicators.

If a variable has no prefix, it refers to the exposure to the environmental hazard in the year of the wave. Prefixes starting with 's' correspond to a rolling sum of exposure, with the simple 's_' corresponding to the rolling sum of exposure from birth (or from the oldest year available) up until the year of the wave in question.

The prefixes starting with 'y' are simple sums instead of rolling sums; they correspond to total exposure during certain, relevant, years. For early age exposure, 'y5_', 'y10_' and 'y15_' correspond to total exposure during the first 5, 10 and 15 years of age. 'yjob_' corresponds to exposure during the years at current job or at the most recent job. We also generate variables for exposure to environmental in the years preceding periods of ill health during adulthood. Respondents indicate up to 3 periods where they experienced ill health, specifying the start and end (more details in section Morbidity). For individuals indicating illness periods, we construct variables with prefix 'yill1_', 'yill2_' and 'yill3_' denoting exposure during the years of illness periods 1,2 and 3 respectively. We construct variables with prefix 'y1bf_', 'y3bf_' and 'y5bf_' to represent exposure to hazards during the 1 year, the 3 years and the 5 years preceding the start of each illness period.

We generate cumulative variables since birth for 6 of the 16 temperature bins, on the low extremes and on the high extremes, i.e., for temperatures below 5ºC, between -5ºC and -2.5 ºC and between -2.5 ºC and 0 ºC; and for temperatures between 25 ºC and 27.5 ºC, between 27.5 ºC and 30 ºC, and above 30 ºC. Others can be made readily available on request. On the temperature variables, we report cumulative exposure since birth for CDD and HDD. Cumulative variables since birth are also available for precipitation. We report cumulative variables for flood variables as well.

As auxiliary variables, we report the rolling sum of the number of years for which cumulative measures were computed. We choose to provide both cumulative exposures and years for which cumulative exposure is available, instead of only averages, since even for the same variable, there is not available information for the same number of years for all individuals. This is for two reasons: i) individuals who were born before the years where the environmental variables start and ii) periods in which individuals were outside their country of interview. By providing both cumulative and years available, averages can be readily computed through their ratio, if averages are the variables of interest, and simultaneously, subsets of the sample based on the number of years available (e.g., necessarily all years since birth) can be analysed separately.

We report as well average spring, summer, fall, winter, and yearly temperatures and average radiation, since birth and during the first 5, 10 and 15 years of life. For these, we provide directly these averages alongside the rolling sum of the number of years, instead of cumulative exposure as we do for the remaining (count) variables.

1.1.1. Environmental Variables list

Table S1

Variable name	Variable description
Temperature Variables	
Bins	
[tn/tg/tx]_neg5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. below -5ºC ([median/mean] gridcells) at [year of wave/year before/2years before]
[tn/tg/tx]_neg5_neg2p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt -5 and -2.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_neg2p5_0_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt -2.5 and 0ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_0_2p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 0 and 2.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_2p5_5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 2.5 and 5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_5_7p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 5 and 7.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_7p5_10_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 7.5 and 10ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_10_12p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 10 and 12.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_12p5_15_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 12.5 and 15ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_15_17p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 15 and 17.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_17p5_20_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 17.5 and 20ºC ([median/mean] gridcells) at [year wave/year before/2years before]

[tn/tg/tx]_20_22p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 20 and 22.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_22p5_25_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 22.5 and 25ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_25_27p5_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 25 and 27.5ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_27p5_30_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. bt 27.5 and 30ºC ([median/mean] gridcells) at [year wave/year before/2years before]
[tn/tg/tx]_g30_[median/mean][none/_t1bf/_t2bf] (3*2*3 vars)	No. days [min/avg/max] temp. above 30ºC ([median/mean] gridcells) at [year wave/year before/2years before]
<u>Average temperatures</u>	
temperature_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg daily mean temperature ([median/mean] gridcells) at [year wave/year before/2years before]
summer_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg summer daily mean temperature ([median/mean] gridcells) at [year wave/year before/2years before]
spring_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg spring daily mean temperature ([median/mean] gridcells) at [year wave/year before/2years before]
fall_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg fall daily mean temperature ([median/mean] gridcells) at [year wave/year before/2years before]
winter_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg winter daily mean temperature ([median/mean] gridcells) at [year wave/year before/2years before]
<u>CDD/HDD</u>	
CDD_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	EUROSTAT Cooling degree days index ([median/mean] gridcells) at [year wave/year before/2years before]
HDD_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	EUROSTAT Heating degree days index ([median/mean] gridcells) at [year wave/year before/2years before]
Radiation Variables	
radiation_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Average daily radiation ([median/mean] gridcells) at [year wave/year before/2years before]
radiation_spring_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Average daily radiation in spring months of year ([median/mean] gridcells) at [year wave/year before/2years before]
radiation_summer_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Average daily radiation in summer months of year ([median/mean] gridcells) at [year wave/year before/2years before]

radiation_fall_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Average daily radiation in fall months of year ([median/mean] gridcells) at [year wave/year before/2years before]
radiation_winter_[median/mean]_[t0/t_1bf/t_2bf] (2*3 vars)	Average daily radiation in winter months of year ([median/mean] gridcells) at [year wave/year before/2years before]
Precipitation Variables	
prec10_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	No. days total precipitation above 10mm ([median/mean] gridcells) at [year wave/year before/2years before]
prec20_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	No. days total precipitation above 20mm ([median/mean] gridcells) at [year wave/year before/2years before]
Flood Variables	
fl_no_floods_[SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	No. flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
fl_tot_dead_[SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	No. casualities flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
fl_tot_displaced_[SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	No. displaced by flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
fl_weighted_floods_[SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	Weighted No. flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
fl_tot_days_[SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	No. days flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
fl_weighted_[days_SHARE/NUTS1/NUTS2] [none/_t1bf/_t2bf] (3*3 vars)	Weighted No. days flood events in SHARE region / NUTS1 region / NUTS2 region at [year wave/year before/2years before]
Pollution vars	
<u>Concentration</u>	
conc_pm2p5_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg monthly concentration PM2.5 ([median/mean] gridcells) at [year wave/year before/2years before]
conc_pm10_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg monthly concentration PM10 ([median/mean] gridcells) at [year wave/year before/2years before]
conc_no2_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg monthly concentration NO2 ([median/mean] gridcells) at [year wave/year before/2years before]
conc_yearly_o3_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg daily max O3 concentration ([median/mean] gridcells) at [year wave/year before/2years before]
conc_warm_o3_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Avg daily max O3 concentration in warm months (Apr-Sep) ([median/mean] gridcells) at [year wave/year before/2years before]
<u>Emissions</u>	
emissions_PM25_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Yearly emissions of PM2.5 ([median/mean] gridcells) at [year wave/year before/2years before]
emissions_PM10_[median/mean][none/_t1bf/_t2bf] (2*3 vars)	Yearly emissions of PM10 ([median/mean] gridcells) at [year wave/year before/2years before]

The cumulative variables are created using the yearly variables, therefore, their names are the same, but with added prefixes which indicate over what period are the cumulative measures taken.

Table S2 Cumulative Variables

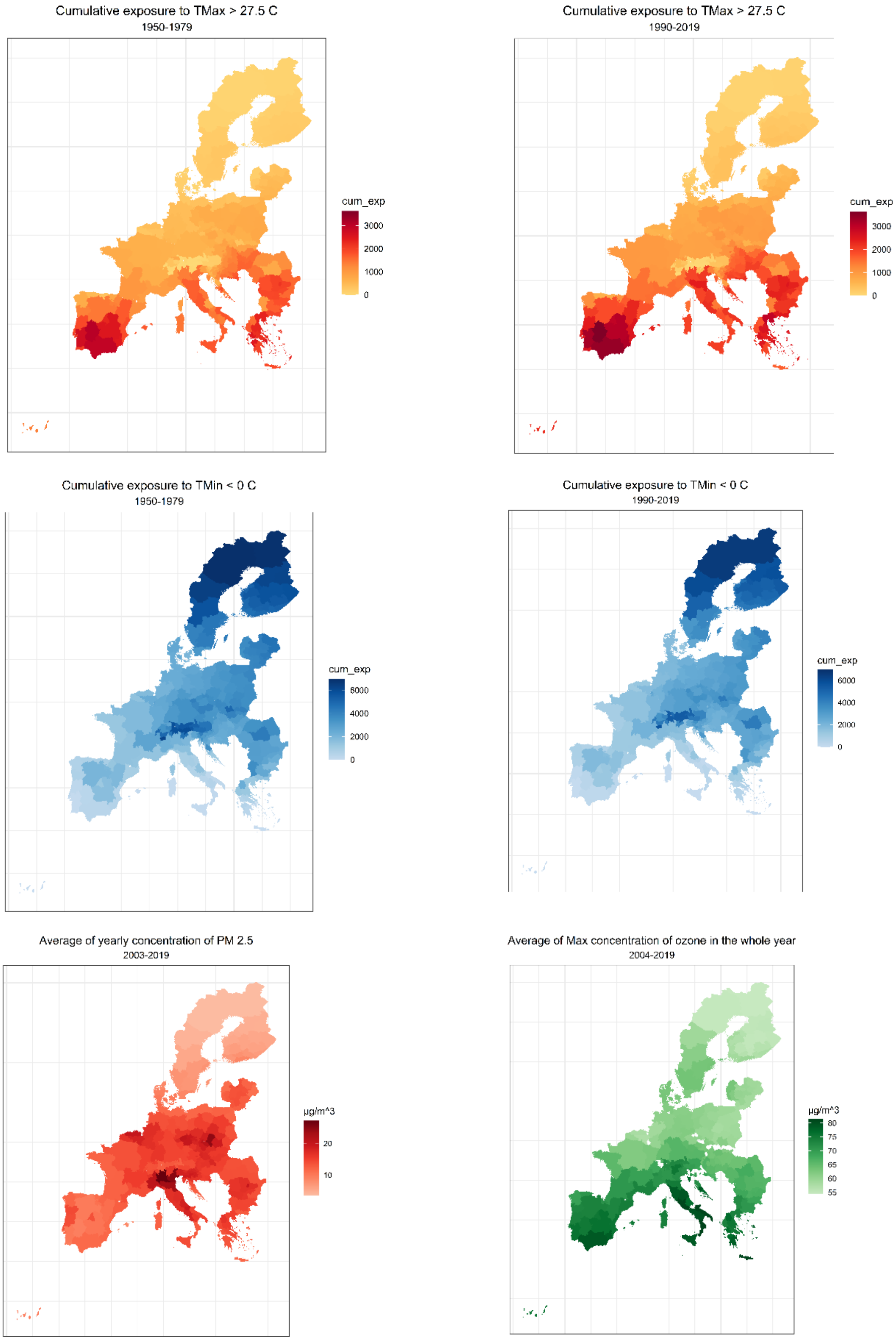
Main exposure variables			
Cumulative variable prefix	**Prefix meaning**	**Yearly Variables for which the cumulative measure is calculated**	**Module**
s_	Rolling sum since birth (or earliest available year) until present wave	Bin variables (1920), HDD (1920), CDD (1920), Precipitation Variables (1920), Flood Variables (1985)	life_module
avg_	Rolling average since birth (or earliest available year) until present wave	Average temperature variables (1920), Radiation Variables (1950), Concentration Variables (2003/2004), Emissions Variables (1970),	life_module
y5_ / y10_ / y_15	Cumulative exposure during the first 5/10/15 years of life	Bin variables, HDD, CDD, Precipitation Variables (no flood variables since no individual is born after 1970, thus, not 15 after 1985)	young_age_module
avg5_ / avg10_ / avg15_	Average during the first 5/10/15 years of life	Average temperatures variables, Radiation Variables, (no concentration variables), Emission Variables	young_age_module
yjob_	Cumulative exposure during the most recent job	Bin variables, HDD, CDD, Precipitation Variables, (no flood variables)	job_module
avgjob_	Average exposure during the most recent job	Average temperature variables (1920), Radiation Variables (1950), Concentration Variables (2003/2004),	job_module

		Emissions Variables (1970),	
yill[1/2/3]_	Cumulative exposure during illness period 1/2/3	Bin variables, HDD, CDD, Precipitation Variables, Flood variables	illness_during _module
avgill[1/2/3]_	Average exposure during illness period [1/2/3]	Average temperature variables (1920), Radiation Variables (1950), Concentration Variables (2003/2004), Emissions Variables (1970),	illness_during _module
y[1/3/5]bf_[1/2/3]_	Cumulative exposure during the [1/3/5] year(s) preceding illness period [1/2/3]	Bin variables, HDD, CDD, Precipitation Variables, Flood variables	illness_before _module
avg[1/3/5]bf_[1/2/3]_	Average exposure during the [1/3/5] year(s) preceding illness period [1/2/3]	Average temperature variables (1920), Radiation Variables (1950), Concentration Variables (2003/2004), Emissions Variables (1970),	illness _before_mod ule
Auxiliary variables (denominator for averages)			
rol_years_exposure_ [temp / prec/rad/ fl[SHARE/NUTS1/NUTS2] /conc_noto3/ conc_o3/emissions]	Rolling sum of non-empty years of [temperature/ precipitation/ radiation / flood pertaining to [SHARE region /NUTS1 region/NUTS2 region] / non-O3 concentration/ O3 concentration/emission s] variables	All variables	life_module

tot_years_exposure_ [temp/prec/ rad/ fl[SHARE/NUTS1/NUTS2]/ conc_noto3 /conc_o3/emissions]	Maximum of rol_years_exposure_ [temp/prec/ rad/fl[SHARE/NUTS1/NUTS2] /conc_noto3 /conc_o3/emissions]	All variables	life_module
years_present_ [temp/prec/rad/emissions] _ outof[5/10/15]	Years for which there is information on [temperature /precipitation /radiation/emissions] variables out of the first [5/10/15] years of life	All variables except flood and concentration variables (do not go back in time sufficiently to catch the first 15 years of life of respondents)	young_age_module
job_years_exposure_ [temp/prec/rad/emissions /con_noto3/conc_o3]	Years in which individual was at most recent job for which there is information on [temperature /precipitation /radiation/emissions /non-ozone concentration/ozone concentration] variables	All variables except flood variables (flood events during years at most recent job not considered a variable of interest)	job_module
Ill_years_exp_dur[1/2/3]_ [temp/prec /rad/emissions/ fl[SHARE/NUTS1/NUTS2]/ conc_noto3/conc_o3/emissions]_	Years for which during period of illness [1/2/3] there is info on [temperature/precipitation/radiation/ emissions/floods at [SHARE level/NUTS1 level/NUTS2 level]/non-ozone concentration/ozone concentration/emissions] variables	All variables	illness_during_module
ill_y_exp_[1/3/5]_bf_[1/2/3]_ [temp/prec /rad/emissions/ fl[SHARE/NUTS1/NUTS2]/ con_noto3/con_o3/emissions]	Years for which there is information out of the [1/3/5] years before illness period [1/2/3] on [temperature/precipitation/ radiation/	All variables	illness_before_module

	Emissions/floods at [SHARE level/NUTS1 level/NUTS2 level]/not-ozone concentration/ozone concentration/emissions] variables		

Figure S1. Selected environmental variables



SHARE data

SHARE is a longitudinal random sample of Europeans aged 50+ for 28 countries. SHARE contains approximately 120,000 individuals and 300,000 interviews. The regular panel waves (2004-2019) of SHARE follow individuals (respondents and their spouses) over time. Household members are being interviewed every two years. In addition, the specific module on SHARELIFE (waves 3 and 7) reconstructs the retrospective life history of the same people, which includes information on early life conditions, health, health care and working life. The data has been collected by making use of a "timeline" methodology, which carefully reconstructs the life of individuals by gradually including focal points such as school-leaving age, job started/finished, onset of illness, housing and birth of children.

1.1.2. Regional information used for construction of environmental exposure variables

It is the regional information present in SHARE that allows us to merge in our environmental hazard indicators, described in the preceding section.

SHARE regions are those which respondents indicate as the location of their accommodation in the retrospective accommodation waves in waves 3 and 7 (SHARELIFE waves), modules AC and RA respectively, or the NUTS in which the household was located at the moment of sampling, provided in modules gv_housing. The regions from the retrospective waves most often have a direct correspondence to NUTS regions (with certain exceptions, for instance, for Luxembourg, respondents indicate the canton). Whenever regions were the combination of two NUTS regions, we reported the environmental hazard variables at the level of those combined regions.

The gv_housing modules allow us to go beyond the retrospective waves data in two ways: i) extending forward cumulative exposure information for individuals who participated in waves beyond wave 3, since the regional information on their accommodation only extends up to wave 3 using the retrospective waves; ii) adding year-on-year information and cumulative exposure since first wave of participation until last for individuals who did not reply to either retrospective accommodation module (either because they died, were not yet part of the survey, or simply did not participate in wave 3 and 7 or did not answer to the retrospective accommodation questions in either of these waves).To keep exposure information always at the same disaggregation level, where regional information from the gv_housing modules differed from that from the retrospective accommodation modules, we transformed it to the latter.

Whenever individuals lived in a country different to that in which they were now sampled, we do not know in which region they lived, but only their country. Country-level information is considered too aggregate to provide useful environmental exposure measures. Thus, for periods where respondents were outside the country, we do not have any environmental information. Cumulative exposure variables, therefore, do not consider such years. A correction can be done by dividing cumulative exposures by the number of years for which there is information (which excludes the years when individuals were abroad).

1.1.3. Outcome variables of particular interest for environmental hazards

The SHARE database contains variables that can be used to characterize the impacts on well-being of environmental risk. All individuals living within a SHARE region in a certain year are treated as facing the same average environmental conditions. Here we highlight some of the SHARE variables which can be outcome variables of interest in the context of morbidity. In the same way as for exposure/vulnerability variables we have built some variables based on underlying SHARE information which can facilitate empirical analyses.

Morbidity

For morbidity we have perceived health status at the individual level (1 to 5, excellent to poor, based on question ''Would you say your health is.. 'Excellent (1), Very Good (2), Good (3), Fair (4), Poor (5)). We also know if an individual has being ever diagnosed or bothered by a disease, if he or she is taking drugs for certain illnesses, and the age of the onset for a range of illnesses, such as: heart attack, stroke, high blood pressure, asthma, lung disease, cancer, diabetes, arthritis and other, Alzheimer, Parkinson, Mental disorders, depression. Respondents also provide information on up to three periods of ill health throughout life, with a start and an end year, and what health conditions were responsible for such periods. Questions on severity of the illness include whether they brought on negative consequences at work, limited social life and leisure activities or impacted the family negatively.

There are questions specific to childhood health. Beyond perceived childhood health status (as with adult health status, variable takes values from 1 to 5, excellent to poor, based on question 'Would you say that your health during your childhood was in general: excellent (1), very good (2), good (3), fair (4), or poor (5)'), other questions measure possible severity of health conditions, namely, if the respondents ever missed school for at least one month, were ever confined to their beds for at least one month, committed to a hospital for one month or longer, or in a single year, hospitalized at least three times.

Respondents answer as well whether they had any out of a list of illnesses during childhood, of note, infectious diseases, asthma, respiratory problems other than asthma, allergies, severe diarrhoea, severe headaches, emotional problems, childhood diabetes and heart trouble. Regarding such conditions, respondents do not provide exact start and end dates for the illness, but state whether the condition lasted for at least one year, and whether it took place from 0 to 5 years old, from 6 to 10, or from 11 to 15 years old.

We give particular attention to illnesses, which in the literature are found to be associated with environmental factors, particularly extreme temperatures. From the list of illnesses SHARE provides information on, we focus on the following, which we designated 'environmentally-related' – angina or heart attack, stroke, asthma, (other) respiratory problems, migraines, emotional distress, fatigue, infectious diseases and allergies. We generate several variables which can facilitate analysis of environmental factors, as described in Table *S3* Morbidity generated variablesTable . These variables are provided as part of the 'illness_before_module' and 'illness_during_module'.

Table S3 Morbidity generated variables, provided in example SHARE dataset

Variable	Variable Description
ill_length_[1/2/3]	Length of illness period [1/2/3]
ill_age_onset_[1/2/3	Age of onset of illness period 1/2/3

ill_start_[1/2/3]	Year when illness period 1/2/3 started
ill_end_[1/2/3]	Year when illness period 1/2/3 ended
Ill_any_issue[1/2/3]	Any issue in period 1/2/3
Ill_any_env_related_issue[1/2/3]	Any environmentally-related issue in period 1/2/3
[Environmentally-related illness name]_[1/2/3]	Whether it was [angina or heart attack/stroke/asthma/other respiratory problems/migrains/emotional distress/fatigue/infectious diseases/allergies] (one of) the issue(s) responsible for illness period [1/2/3]

<u>Note:</u> Environmentally-related issues are angina or heart attack, stroke, asthma, (other) respiratory problems, migraines, emotional distress, fatigue, infectious diseases and allergies.

There are other measures of health outcomes which are clinically measured, some of which are especially targeted to older age individuals. These are:

- Depression scores;
- Cognitive scores for different cognitive functions;
- Physical Health measures, including: Difficulties with Activities of Daily Living (ADL), difficulties with Instrumental Activities of Daily Living (IADL), Lung functioning, Walking speed, Grip Strength and Dried Blood Spots.

In addition to the morbidity outcome, a wide range of other individual and household-level characteristics are available. These include, for example, quality of housing, location of dwelling (big city, the suburbs or outskirts of a big city, a large town, a small town, a rural area or village), type of housing situation (e.g. owner versus renter), occupation including ISCO coding, education including ISCED codes, job conditions. Information about income/wealth/material wellbeing, migration, behavioural risks (e.g. Smoking, drinking; stress levels; parental behavioural risks) are also available.

Association between environmental hazards and subjective and objective health outcomes

Table S4: Extensive regression results of Main Text Table 4

Variables	1. Ever experienced breathlessness		2. Young age (<15) perceived reported health (1=poor; 5= excellent)		3. Old age (>49) perceived reported health (1=poor; 5= excellent)			4. Uncomfortable job	
Exposure Variables	**Avg. pm2.5 conc. median (µg/m3)**	0.00351***	**Avg. first 15 years exposure to negative temperature (# days)**	0.000122	**Avg. Lifetime exposure to temperature > 27.5°C (# days)**	-0.00164**	-0.00115	**Average winter temperature**	0.00710***
		(0.000728)		(0.000383)		(0.000826)	(0.000884)		(0.00130)
	Avg. Cum. lifetime exposure to negative temperature (# days)	2.48e-05	**Avg. first 15 years exposure to temperature > 30°C (# days)**	0.00288*	**Avg. Cum. lifetime exposure to negative temperature (# days)**	-0.00130***	-0.00134***	**Average summer temperature**	-0.00263***
		(9.09e-05)		(0.00151)		(0.000426)	(0.000427)		(0.000981)
	Avg. Lifetime exposure to temperature > 30°C (# days)	-0.000410*	**Avg. first 15 years solar radiation (W/m^2)**	0.00221*				**Average radiation**	0.000115
		(0.000243)		(0.00133)					(0.000366)
Exposure x Individual characteristics					**AC x Avg. Lifetime exposure to temperature > 27.5°C (# days)**	0.00233***	0.00161**	**Job is physical x average winter temperature**	-0.0112***
						(0.000649)	(0.000678)		(0.00155)
					Central heating x Avg. Cum. lifetime exposure to negative temperature (# days)	0.000428	0.000691	**Job is physical x average summer temperature**	0.00850***
						(0.000476)	(0.000472)		(0.00135)

					Job is physical x average radiation	0.00133***
						(0.000261)
AC, Central heating, whether job is physical						
AC			-0.0208	-0.0227		
			(0.0376)	(0.0380)		
Central Heating			0.0242	-0.0469		
			(0.0461)	(0.0460)		
Job is physical						-0.00280
						(0.0300)
Occupation and edu fixed effects						
Occupation ISCO codes (1digit)						Y
Parental ISCO codes (1digit)	N/A	Y	N/A	N/A		
Mother ISCED edu level		0.00347				
		(0.00464)				
Father ISCED edu level		0.00294				
		(0.00229)				
ISCED edu level 2				0.120***		-0.0156*
				(0.0251)		(0.00882)
ISCED edu level 3				0.183***		-0.0442***
				(0.0220)		(0.00826)
ISCED edu level 4				0.276***		-0.0623***
				(0.0447)		(0.0123)
ISCED edu level 5				0.362***		-0.0806***
				(0.0250)		(0.00908)
ISCED edu level 6				0.377***		-0.123***
				(0.124)		(0.0205)
Socio-Ecconomic conditions controls						
Household Networth			2.33e-07***	1.78e-07***		

				(2.83e-08)	(2.67e-08)		
Household Income (current/average)	average	-5.20e-07***		6.99e-07**	5.04e-07***	average	-3.38e-07***
		(8.09e-08)		(2.80e-07)	(1.65e-07)		(1.07e-07)
House at 15 years had no basic ammenities			-0.125***				
			(0.0276)				
Rooms / people when 10 years old			0.0388*				
			(0.0199)				
Ever poor in childhood			-0.172***				
			(0.0209)				
% of time in urban area			-0.000159				
			(0.0214)				
Physical harm in childhood			-0.102***				
			(0.0163)				
Loneliness in childhood			-0.280***				
			(0.0204)				
Behavioural variables							
BMI (Body Mass Index)	average	0.00985***		-0.0269***	-0.0242***		
		(0.000483)		(0.00186)	(0.00188)		
Ever smoked		0.0393***		-0.0387**	-0.0433***		
		(0.00386)		(0.0158)	(0.0159)		
Sports more than once a week		0		0	0		
		(0)		(0)	(0)		
Sports once a week		0.00136		-0.136***	-0.158***		
		(0.00570)		(0.0227)	(0.0229)		
Sports once to three times a month		0.0111*		-0.152***	-0.162***		
		(0.00620)		(0.0252)	(0.0253)		
Sports hardly ever		0.0813***		-0.492***	-0.493***		
		(0.00481)		(0.0192)	(0.0193)		
Health variables							
Depression				-0.412***	-0.404***		
				(0.0168)	(0.0169)		

Born with an illness	0.130***	-0.674***			
	(0.0276)	(0.0906)			
Job comfort level					
Job is uncomfortable [strongly disagree]	0		0	0	
	(0)		(0)	(0)	
Job is uncomfortable [disagree]	0.0183***		-0.106***	-0.0867***	
	(0.00465)		(0.0204)	(0.0203)	
Job is uncomfortable [agree]	0.0388***		-0.233***	-0.171***	
	(0.00565)		(0.0235)	(0.0238)	
Job is uncomfortable [strongly agree]	0.0613***		-0.298***	-0.216***	
	(0.00717)		(0.0268)	(0.0276)	
Fixed at birth controls					
Female			-0.0407**	-0.0172	
			(0.0163)	(0.0165)	
Current Age	0.00489***				
	(0.000219)				
Year of Birth		0.0106***	0.0208***	0.0178***	
		(0.00113)	(0.000910)	(0.000951)	
Country Fixed Effects	Y	Y	Y	Y	Y
Year Fixed Effects	N/A	N/A	Y	Y	N/A
Observations	33,511	16,086	14,714	14,296	44,308
R-squared	0.081	0.081	0.294	0.302	0.205

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table S5: Effect of temperature exposure on old age health at first wave of participation (younger age) and at last wave of participation (older age)

	Old age (>49) perceived reported health (1=poor; 5= excellent)	
	Fiirst wave of participation	Last wave of participation
Avg. Lifetime exposure to temperature > 27.5°C (# days)	-0.000500	-0.00169***
	(0.000491)	(0.000492)
Avg. Cum. lifetime exposure to negative temperature (# days)	-0.000826***	-5.47e-05
	(0.000259)	(0.000242)
Occupation and educational effects		
ISCED edu level 2	0.0827***	0.0559***
	(0.0160)	(0.0152)
ISCED edu level 3	0.208***	0.136***
	(0.0149)	(0.0140)
ISCED edu level 4	0.330***	0.163***
	(0.0295)	(0.0252)
ISCED edu level 5	0.398***	0.249***
	(0.0165)	(0.0158)
ISCED edu level 6	0.395***	0.294***
	(0.0514)	(0.0463)
Household Networth	8.02e-08***	1.25e-07***
	(1.92e-08)	(2.38e-08)
Household Income (current/average)	8.88e-07***	2.15e-06***
	(1.70e-07)	(2.67e-07)
Behavioural variables		
BMI (Body Mass Index)	-0.0302***	-0.0286***
	(0.00109)	(0.00104)
Ever smoked	-0.0798***	-0.0652***
	(0.00942)	(0.00888)
Sports more than once a week	0	0
	(0)	(0)

Sports once a week	-0.136***	-0.0699***
	(0.0135)	(0.0154)
Sports hardly ever	-0.489***	-0.377***
	(0.0113)	(0.0125)
Depression	-0.395***	-0.448***
	(0.00981)	(0.0117)
Job is uncomfortable [strongly disagree]	0	0
	(0)	(0)
Job is uncomfortable [disagree]	-0.106***	-0.0874***
	(0.0116)	(0.0108)
Job is uncomfortable [agree]	-0.191***	-0.170***
	(0.0136)	(0.0128)
Job is uncomfortable [strongly agree]	-0.231***	-0.226***
	(0.0170)	(0.0161)
Female	-0.00685	0.0539***
	(0.00975)	(0.00931)
Year of Birth	0.0153***	0.0205***
	(0.000550)	(0.000519)
Country Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
Observations	41,296	46,110
R-squared	0.270	0.224

*** p<0.01, ** p<0.05, * p<0.1

References

Brakenridge, G. R., DFO Flood Observatory, University of Colorado, 07/09/2021. Global Active Archive of Large Flood Events, 1985-Present. Available from Robert.Brakenridge@Colorado.edu.

Bell, M. L., Peng, R. D., & Dominici, F. (2006). The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environmental Health Perspectives*, *114*(4), 532–536.

Gryparis, A., Forsberg, B., Katsouyanni, K., Analitis, A., Touloumi, G., Schwartz, J., Samoli, E., Medina, S., Anderson, H. R., Niciu, E. M., Wichmann, H. E., Kriz, B., Kosnik, M., Skorkovsky, J., Vonk, J. M., & Dörtbudak, Z. (2004). Acute effects of ozone on mortality from the "Air Pollution and Health: A European Approach" project. *American Journal of Respiratory and Critical Care Medicine*, *170*(10), 1080–1087

Kazemiparkouhi, F., Eum, K.-D., Wang, B., Manjourides, J., & Suh, H. H. (2020). Long-term ozone exposures and cause-specific mortality in a US Medicare cohort. *Journal of Exposure Science & Environmental Epidemiology*, *30*, 650–658.

Kim, S. Y., Lee, J. T., Hong, Y. C., Ahn, K. J., & Kim, H. (2004). Determining the threshold effect of ozone on daily mortality: An analysis of ozone and mortality in Seoul, Korea, 1995-1999. *Environmental Research*, *94*(2), 113–119.

Lim, C. C., Hayes, R. B., Ahn, J., Shao, Y., Silverman, D. T., Jones, R. R., Garcia, C., Bell, M. L., & Thurston, G. D. (2019). Long-term exposure to ozone and cause-specific mortality risk in the United States. American Journal of Respiratory and Critical Care Medicine, 200(8), 1022–1031.

Malley, C. S., Henze, D. K., Kuylenstierna, J. C. I., Vallack, H. W., Davila, Y., Anenberg, S. C., Turner, M. C., & Ashmore, M. R. (2017). Updated Global Estimates of Respiratory Mortality in Adults ≥30Years of Age Attributable to Long-Term Ozone Exposure. Environmental Health Perspectives, 125(8), 087021

World Health Organization (2014.) Review of evidence on health aspects of air pollution – REVIHAAP" project Technical Report, available at https://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-report.pdf