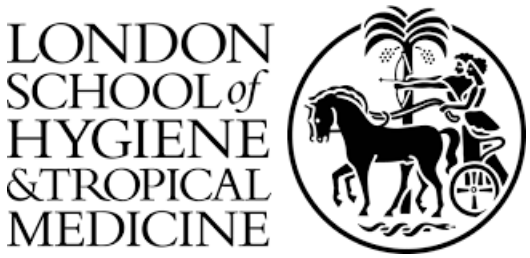


An extended random-effects framework for complex meta-analysis, with applications in environmental epidemiology

Francesco Sera

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy
February, 2023

Department of Health Services Research and Policy
Faculty of Public Health and Policy
London School of Hygiene and Tropical Medicine

No funding was received

Declaration of Own Work

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

Abstract

Standard applications of meta-analysis consider a single effect-size estimated from independent studies. However, extensions to deal with more complex meta-analytical problems have been presented, such as multivariate, network, multilevel, dose-response, longitudinal meta-analysis and meta-regression. These extensions are characterised by non-independence among effect-sizes with a complex correlation structures that need to be modelled or accounted for. In my PhD, I reviewed and brought together these different extensions, developing a coherent extended mixed-effects framework for meta-analysis. The framework is built on the link between meta-analysis and linear mixed-effects models, where patterns of effect sizes are modelled through a flexible structure of fixed and random terms. The extended mixed-effects framework for meta-analysis has been implemented in the R package mixmeta. Meta-analytic models are often applied to environmental epidemiology using two-stage designs. In this setting, location-specific exposure-response associations are estimated in the first stage, and then the estimates are pooled using meta-analytic methods in the second stage. In my PhD, I illustrated multiple design extensions of the classical two-stage method, all implemented using the extended mixed-effects framework described above. In addition, I applied the framework and related software to show the advantages of using the extended two-stages design in environmental epidemiology studies, allowing a clearer characterisation of the short-term health effects of environmental stressors. In these applications, I first explored the role of urban characteristics in modifying the effects of temperature on health. Then I used a multi-country, multi-city, longitudinal design to quantify the independent role of air conditioning in the attenuation of heath related health risk. Finally, I developed a two-stage ecological modelling approach to examine the impact of meteorological variables on SARS-CoV-2 transmission. The extended mixed-effects framework for meta-analysis and related software has proved to be a valid and useful analytical tool to address research questions on environmental health risks and beyond.

Contents

List of Tables

List of Figures

Acknowledgements

I am grateful to my supervisor, Prof. Antonio Gasparrini, for his supervision over the last five years. His support, patience and guidance allowed me overcoming the challanges related to this PhD project. I would also like to thank the members of my advisory board: Prof. Ben Armstrong, for his wise guidance on substantive issues related to my research topics; Prof. Harvey Goldstein, for his stimulus on addressing the problems on a different perspective; and Prof. Marta Blangiardo for her competent comments and inputs in the different phases of my PhD. This PhD project benefitted from the intense research activity and network that characterise my supervisor's team. I am grateful to Ana Vicedo-Cabrera, Rochelle Schneider, Pierre Masselot, and Malcom Mistry for all the discussion made on the research topics related to this PhD and beyond. I would also like to remember and thank all the researchers I have encountered in this long journey in the community of epidemiologists and biostatisticians over the last 25 years, in particular Mariarosa Corona, Domenico Palli, Annibale Biggeri, Carol Dezateux, and Harvey Goldstein. My gratitude also goes to my family and friends, especially during the nice time spent in London. The final thought is for Gianluca and Deborah, for the marvellous life we are having togheter and for making everything possible.

Preface

This PhD thesis consists of a collection of research papers and software documentation. These publications are related to the same research topic, but they have been published as independent research contributions. As a result, concepts and definitions could be repeated in different papers, and, more importantly, their content is not uniformly linked and standardized. The thesis is therefore divided into three main parts, where the selected publications are preceded by an introduction and followed by a final discussion. The aim is to present my research activity during the PhD project as a coherent body of work.

The introduction in Part I contains two main chapters. The context of complex meta-analysis, as a general analytical tool as well as applications in environmental epidemiology, is illustrated in Chapter 1. Chapter 2 offers a summary of the publications, also introducing the main statistical developments and the related software implementation. In Part III, Chapter 8, I provide a final discussion and describe potential directions for future research.

Part II includes the selected five publications in different chapters. The order has been chosen to reflect the research steps of the PhD project. The first publication in Chapter 3 present the extended framework for complex meta-analysis. The article in Chapter 4 illustrates multiple design extensions of the classical two-stage method in environmental epidemiology, all implemented using the extended random-effects framework for complex meta-analysis. The publication in Chapter 5 is an example of a standard two-stage design applied in environmental epidemiology. The last Chapters 6 and 7 include two publications that describe the extensions of the classical two-stage method in environmental epidemiology.

Part I

Introduction

Chapter 1

Background

1.1 Historical facts

The term "Meta-Analysis" was introduced in 1976 by Gene V. Glass in his seminal paper "Primary, Secondary, and Meta-Analysis of Research" (Glass, 1976). He defined "Meta-Analysis" as "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings". We can recognise two critical aspects in his definition: firstly, the number of results from individual studies is "large", with difficulties in extracting meaningful information; secondly, the stochastic nature of the results is assumed, and statistical methods are claimed to integrate the findings. According to Glass, the meta-analysis should be a "rigorous alternative to the casual, narrative discussions of research studies". In the same period, in medical statistics, there were the first attempts to combine information from randomised controlled trials (RCT) (Altman, 2015; Peto *et al.*, 1977). These techniques were appealing as the RCTs were of a small sample size with low power to detect effects and a high probability of false-positive findings. Meta-analytic works continued to increase in medical sciences in the 1980s, and 1990s (Shadish and Lecy, 2015) along with the "Evidence-Based Medicine" paradigm.

Several authors pointed out the misleading use of meta-analytic methods (Bailar III, 1995; Egger and Smith, 1995; Eysenck, 1994); concerns were expressed about the biases that could affect the individual studies, the lack of similarity across studies on crucial aspects of the study design (individual, setting, treatment, and outcome), and the presence of selective publication processes. These concerns were amplified by the tendency to apply the meta-analytic method to summarise results from observational studies (Egger *et al.*, 1998; Shapiro, 1997). Observational studies do not rely on randomisation to ensure homogeneity of exposure groups, but rather the researchers use selection procedures, information on personal characteristics, and analytic methods to minimize possible bias due to confounding. The heterogeneity of the population investigated, and the different techniques used to minimise information and selection biases required statistical methods (e.g., "meta-regression") that allow studying possible factors explaining the heterogeneity of the study results.

Individual Patient Data (IPD) meta-analysis could partially address the problems of applying meta-analytic techniques to summarise RCTs and observational studies by obtaining and synthesising raw participant data from each study (Riley *et al.*, 2010). However, IPD meta-analysis is resource-intensive as it requires a considerable effort in building the research network and creating a team with skills and competence to collect the data, perform the quality checks, and the data analysis. As advantages, I can mention that the researcher can adopt common and predefined selection criteria, standardise the analysis plan across the studies, and check the assumptions of statistical models used (Burke *et al.*, 2017; Riley *et al.*, 2010). Another advantage of IPD meta-analysis is the possibility to evaluate modifiers explaining the heterogeneity of the study results with higher power and avoiding ecological bias when compared to "meta-regression" techniques (Thompson and Higgins, 2002). There are two statistical approaches for IPD meta-analysis: the one-stage or the two-stage approach (Burke *et al.*, 2017; Higgins *et al.*, 2001). In the two-stage approach, the row participant data are analysed in each study, summary effect estimates are calculated along with a measure of their precision (e.g., standard error or confidence intervals), then the estimates are pooled using standard meta-analytic models. The one-stage meta-analysis analyses all the row data in a single model, for example, using a fixed or random-effects (hierarchical, multilevel) model to account for the study's clustering of participants. Overall no clear evidence can be draw on the choice between one-stage or two-stage approach (Burke

et al., 2017; Crowther *et al.*, 2014; Goldstein *et al.*, 2000; Higgins *et al.*, 2001; Thompson *et al.*, 2001; Turner *et al.*, 2000; Whitehead *et al.*, 2001).

The last two decades have seen an expansion of the scope of meta-analyses. Within medical research two major directions can be identified. First, from the initial emphasis on RCTs and observational studies the attention has moved also to diagnostic (Ma *et al.*, 2016), prognostic (Arends *et al.*, 2008; Jackson *et al.*, 2014), and genetic (Bagos and Liakopoulos, 2010; Hong and Breitling, 2008) studies. Second, meta-analysis characterised by a complex study design has been performed (Ishak *et al.*, 2007; Riley *et al.*, 2017). Other development were systematic reviews of systematic reviews (panoramic meta-analysis) (Hemming *et al.*, 2012), cross-design synthesis (Larose and Dey, 1997; Prevost *et al.*, 2000), and split/analysis/meta-analysis methods applied to Big Data (Cheung and Jak, 2016).

1.2 Applications in environmental epidemiology

In environmental epidemiological studies, it is common practice to investigate short-term associations between environmental exposures and health outcomes by analysing time-series data collected from multiple locations. The pooling of data collected in multiple locations increases the statistical power, thus facilitating the detection of small risks usually associated with environmental stressors (Armstrong *et al.*, 2020). Moreover, the analysis of large datasets collected across multiple populations increases the representativeness of the findings.

In this setting, the hierarchical structure of the time-series data should be considered when performing the analysis, e.g., using fixed or mixed-effects (multilevel) models. Moreover, to remove unmeasured time-varying confounders, the model usually requires an aggressive control for trends with the definition of highly-parametrised models, and the assumption that these trends vary across locations. As a results, the number of parameters to be estimated could be very high. It could be computationally unfeasible to perform the analysis with a single model (one-stage approach), and alternative approaches are needed.

An analytical approach applied in this setting is based on the two-stage design, which has become the standard method for the analysis of multi-location data (Baccini *et al.*, 2008; Basagana *et al.*, 2018; Berhane and Thomas, 2002; Bobb *et al.*, 2014; Chen *et al.*, 2012; Dominici *et al.*, 2000; Gasparrini *et al.*, 2015; Liu *et al.*, 2019; Romieu *et al.*, 2012; Samoli *et al.*, 2005; Schwartz, 2000; Wong *et al.*,

2008). The design is based on the separation of the analysis into two stages: In the first stage, location-specific exposure-response associations are estimated while adjusting for various (often time-varying) confounders; then, in the second stage, the estimates are pooled using meta-analytic methods, which can potentially incorporate location-specific meta-predictors.

The separation of the analysis into two stages provides a flexible and computationally efficient analytical framework compared to one-stage approaches (Berhane and Thomas, 2002; Dominici *et al.*, 2000; Gasparrini *et al.*, 2012; Schwartz, 2000). Moreover, another important advantage of the two-stage design is the enhanced ability to examine heterogeneity in risk across populations, which can be linked to contextual characteristics.

1.3 Statistical models for meta-analysis

The main objective of a meta-analysis is to obtain a summary (pooled) estimate from those presented in several studies. The basic meta-analysis method is to obtain a weighted average of estimates from each study with weights based on the precision of the estimates (Sutton and Higgins, 2008).

The pooled estimate is usually achieved using fixed-effects or random-effects assumptions (Higgins *et al.*, 2009; Riley *et al.*, 2011; Sutton and Higgins, 2008). A fixed-effects meta-analysis relies on the assumption of a common underlying effect across all studies. The pooled estimate is calculated as a weighted average of the study summary estimates, with weights equal to the inverse of the estimated variance. In contrast, random-effects meta-analysis assumes that each study has an underlying "true" estimate, and these estimates are "similar" or "exchangeable" in Bayesian terminology. As a consequence, random-effects models incorporate the underlying between-study variation of the estimates into the weights (DerSimonian and Laird, 1986). An important component of the meta-analytic process is the assessment of "heterogeneity" of the underlying effect and factors that could explain such variation. To assess the presence of heterogeneity, often a chi-square test based on the Q statistics is undertaken against the null hypothesis of no heterogeneity (Borenstein *et al.*, 2021), but this test has shown low power (Jackson, 2006). For this reason, measures that quantify the extent of the heterogeneity are preferable; among those, the I^2 statistics measure the proportion of total variation in study estimates attributable to heterogeneity (Higgins and Thompson, 2002; Higgins *et al.*, 2003).

The second step in the "heterogeneity" assessment process is the analysis of factors that could explain between-study differences. A key tool is meta-regression, which combines meta-analytic and regression-based approaches (Thompson and Higgins, 2002).

Estimation Methods

In the previous section, we briefly described the standard meta-analytic models. In this setting, the parameters (using frequentist terminology) to be estimated are the pooled estimates, the between-studies variance, and the coefficients of the meta-regression model that describe how the pooled estimate varies according to the characteristics of the single studies. Several estimation methods have been proposed. These estimation methods will be briefly reviewed in this paragraph, specifically focusing on random-effects models.

Method of Moments. The traditional approach for random-effects meta-analytical models is represented by the estimator proposed by DerSimonian and Laird based on the method of moments (DerSimonian and Laird, 1986). The method of moments is non-iterative and relies on equating the sample statistic of Q to its expected value; then, the method of moments estimates of the between-study variability is plugged-in into the pooled effect formula. It is important to note that this method does not consider uncertainty in the variance estimates when making inferences on the pooled effect.

Likelihood Methods. Likelihood-based methods derive the (restricted) likelihood of the model and use optimization procedures to maximize them in terms of the model parameters (Brockwell and Gordon, 2001; Hardy and Thompson, 1996); moreover, some applications deal with the problem of uncertainty associated with estimating between-study variance (Hardy and Thompson, 1996). The Maximum Likelihood (ML) estimate of the between-study variance may be downward biased. To overcome this problem, some authors (Goldstein *et al.*, 2000; Thompson *et al.*, 2001; Turner *et al.*, 2000) have proposed Restricted (or Residual) Maximum Likelihood (REML) inferential procedures for the estimation of the pooled effect and between-study variance parameter. The REML procedures consider the degree of freedom spent for inference on the fixed effect (mean effect size) while estimating between-study variance parameters.

Non parametric approaches. Non parametric approaches have been proposed to avoid the assumption of normal distribution of the random effects; among those are the non-parametric maximum likelihood (Aitkin, 1999) and Estimating Equations procedures (Ma and Mazumdar, 2011; Ritz *et al.*, 2008).

Bayesian approach. Several authors have proposed Bayesian inferential procedure for meta-analysis (Higgins *et al.*, 2009; Sutton and Abrams, 2001; Thompson *et al.*, 2001). A main advantage of the Bayesian approach is a clearer interpretation of the hypothesis of the random effect. Within the frequentist (or classic) framework, the effect sizes are supposed to be drawn from an "infinite" population of effect sizes represented by a random variable, while in the Bayesian approach to random-effects meta-analysis is based on the concept of *exchangeability*. Exchangeability is the judgment that the effect sizes, even if they were not identical, cannot be differentiated a priori. Bayesian and likelihood approaches can be viewed as particular cases of hierarchical, mixed effects, or multilevel models (Goldstein *et al.*, 2000).

Computational techniques

The simple calculations used in the DerSimonian and Laird method have made their approach very popular, as the author's derived non-iterative closed formulae for calculating the underlying mean effect (and its variance) and between-study variability. In contrast, in the simple univariate random-effects model, the maximum likelihood procedure is based on the score functions of two parameters (underlying pooled effect and between-study variance), but there is no closed solution because their respective estimators depend on each other. To solve this problem, iterative solutions of the two score equations were proposed (Berkey *et al.*, 1998, 1995; Gumedze and Dunne, 2011). Another possible approach to maximize the likelihood is to use numerical iterative procedures. These procedures were used to calculate ML estimates of the underlying mean effect and between-study variance components (Goldstein *et al.*, 2000; Kalaian and Raudenbush, 1996; Konstantopoulos, 2011; Stram, 1996; Thompson *et al.*, 2001; Turner *et al.*, 2000). In particular, Konstantopolous (Konstantopoulos, 2011) and Stram (Stram, 1996) used a Fisher scoring algorithm, while Kalaian and Raudenbusch (Kalaian and Raudenbush, 1996) used an Expectation-Maximization algorithm, while other authors (Goldstein

et al., 2000; Thompson *et al.*, 2001; Turner *et al.*, 2000) have used Iterative Generalised Least Square (IGLS) algorithm (Goldstein, 1986, 2011). Bagos (Bagos, 2015) proposed a ML procedure based on numerical integration by adaptive quadrature as implemented in the Stata package *gllamm* (Rabe-Hesketh *et al.*, 2001).

The maximum likelihood problem can also be addressed through the maximisation of the profile likelihood function over the between-study variance parameters after fixing the mean effect estimated through the generalised least square formula (equivalent to ML estimators). This approach has computational advantages as the iterative algorithm needs only to find the optimum on the between-study variance parameters space. Gasparrini and colleagues (Gasparrini *et al.*, 2012) proposed a numeric version of the Newton-Raphson algorithm to calculate ML estimates of the between-study variance components.

Numerical iterative algorithms have also been used within REML inferential procedures where the estimates of the between-study variance parameters are plugged into the generalised least square formula for the calculation of the mean summary effect (Gasparrini *et al.*, 2012; Goldstein *et al.*, 2000; Van den Noortgate *et al.*, 2013, 2015; Riley *et al.*, 2007a; Thompson *et al.*, 2001; Turner *et al.*, 2000). Alternative Bayesian models (Higgins *et al.*, 2009; Sutton and Abrams, 2001; Thompson *et al.*, 2001) consider non-informative priors with different distributional choices for the variance-components.

Inferential issues

Inference in the DerSimonian and Laird approach is based on the central limit theorem, which ensures that the standardised mean effect follows a standard normal distribution asymptotically on n, where n is the number of the studies (Borenstein *et al.*, 2021). Several authors have proposed alternatives to the DerSimonian and Laird approach, which considers uncertainty in the between-study variance in the inferential procedure on the mean effect (Guolo and Varin, 2017; Sidik and Jonkman, 2007; Veroniki *et al.*, 2016). Among these authors, Hartung and Knapp (Hartung, 1999; Hartung and Knapp, 2001a,b) proposed a simple adjustment of the standard error of the mean effect and the use of a T distribution for statistical tests and confidence intervals; Sidik and Janckman proposed a similar method (Sidik and Jonkman, 2002, 2005). Similarly, the likelihood approach to inference is based on the Wald test that

asymptotically follows a standard normal distribution (Brockwell and Gordon, 2001). This approach relies on the between-study variance estimated through maximum likelihood and does not consider uncertainty in the estimation. To consider the uncertainty in estimating the between-study variance, Hardy and Thompson describe the procedure to derive profile likelihood-based confidence intervals (Hardy and Thompson, 1996); this approach was also used by Turner and colleagues (Turner *et al.*, 2000). Turner and colleagues have also considered parametric bootstrap (Turner *et al.*, 2000) to take into account uncertainty in the estimation of the between-study variances, while other authors have considered a t-distribution for the Wald statistics (Kalaian and Raudenbush, 1996), suggesting use of t_{k-4} (Berkey *et al.*, 1995), t_{k-2} (Higgins *et al.*, 2009; Riley *et al.*, 2007a) or t_{k-1} (Follmann and Proschan, 1999) test statistics. For finite-sample inference with REML estimation, Kenward and Roger proposed an approximate correction of the covariance matrix and degree of freedom (Kenward and Roger, 1997). This correction has been recently applied to standard univariate meta-analysis (Morris *et al.*, 2018).

Most interest in the inferential procedures is in the underlying mean effect. Inferential procedures on the between-study variance parameters have received less attention, and includes methods based on likelihood-based confidence intervals (Hardy and Thompson, 1996), Wald type (Konstantopoulos, 2011; Stram, 1996) and parametric bootstrap (Turner *et al.*, 2000). These methods have been reviewed by Viechtbauer (Viechtbauer, 2007) and Veronicki and colleagues (Veroniki *et al.*, 2016). Kalaian and Raudenbush (Kalaian and Raudenbush, 1996) used Likelihood Ratio tests to compare models with different structures of the between studies (co)variance matrix, but these tests have problems due to boundary conditions of the parameter space of the null hypothesis (Pinheiro and Bates, 2006).

Non-parametric inferential approaches, e.g. (permutation, and resampling procedures (Guolo and Varin, 2017)), and parametric bootstrap (Turner *et al.*, 2000)) could gain robustness and closer nominal level of type I error concerning standard meta-analytic methods, but are computationally intensive with a possible loss of power. Problems related to the uncertainty of the between-study variance components and their effects on the precision of the mean effect sizes are naturally solved in the Bayesian framework, modelling these quantities as random variables. Posterior distributions of between-study variance and mean effect sizes can be calculated using point estimates and credible intervals derivation.

1.4 More complex settings for meta-analysis

In the last decades, there have been extensions in the scope of the meta-analysis. Meta-analytic techniques have been applied in complex study settings; these include, among others, studies with multiple treatments (network meta-analysis) or multiple outcomes (Riley *et al.*, 2017), and longitudinal studies (repeated-effects sizes) (Ishak *et al.*, 2007) The standard meta-analytic models, introduced in the previous chapter, assume independence across estimates obtained from different studies. In contrast, recent applications include more complex settings where the estimates show a complex pattern of dependencies. Ignoring this correlation pattern between estimates could lead to inefficient and or biased estimates (Jackson *et al.*, 2011, 2017; Riley *et al.*, 2017). The following sections will show different examples that involve complex patterns of dependencies among effect measures and require more advanced study designs.

Multiple outcomes and multiple treatments

In some new applications of meta-analysis, multiple estimates are reported on each study; for example, studies could report more than one clinical outcome (e.g. disease-free and overall survival times in patients with cancer), multiple measures or sub-scales of the same psychometric test, various biomarkers, or multiple measurements of accuracy or performance (Riley *et al.*, 2017). These estimates could be correlated at the study level because, as shown by Gleser and Olkin (Olkin and Gleser, 2009), the correlation at the individual level implies a correlation between effect sizes at the population (study) level. Correlation among effect estimates could also arise when multiple treatments are compared against a common control group. As shown by Gleser and Olkin (Olkin and Gleser, 2009), in this case, the correlation is related to the between-study variation of the baseline risk (or value).

Multiple parameters also arise in two-stage designs applied to environmental epidemiological studies. In this context, in each study area, multiple parameters could be used to represent complex exposure-response dependencies, such as non-linear and lagged temperature-health associations of temperature (Gasparrini, 2014; Gasparrini *et al.*, 2012), or alternatively correlated effects of multiple exposures, such as different pollutants included in the same first-stage model (Dominici *et al.*, 2004).

Multivariate meta-analysis deals with all these issues by jointly modelling the correlated outcome,

assuming a multivariate normal distribution with a known within-study covariance matrix. (Berkey *et al.*, 1998; Bujkiewicz *et al.*, 2013; Kalaian and Raudenbush, 1996; Nam *et al.*, 2003; Olkin and Gleser, 2009; Raudenbush *et al.*, 1988; Ritz *et al.*, 2008; Stram, 1996; Van Houwelingen *et al.*, 2002; Wei and Higgins, 2013a). It is important to note here that the different effect estimates are not pooled all together, but the multivariate meta-analysis produces summary estimates for each outcome or treatment (Riley *et al.*, 2017). An important application related to multivariate meta-analysis with multiple treatments (White *et al.*, 2012) is Network meta-analysis or Multiple Treatment Comparisons (MTC). In this application, the set of treatments under comparison can be different across studies, and gain in efficiency from correlation among effect sizes is reached through the assumption of consistency (Riley *et al.*, 2017).

Methods that handle multiple effect sizes generally assume that the within-study covariance matrix is known. For some complex applications, it is possible to reconstruct the within-study covariance matrix (Olkin and Gleser, 2009). In other applications, often the within-study correlation matrix is not known. Riley and colleagues (Riley, 2009) showed that ignoring the within-study correlation could result in estimates with inferior inferential properties. For example, estimates could have higher mean square error and standard error, or could be even biased if the missing mechanism is not ignorable. These inefficiencies and biases depend on the scale of the within-study correlation respect to the between-study correlation and the variability of the within-study correlation among the studies. Several authors have proposed inferential procedures to take into account the missing within-covariance matrix (Hedges *et al.*, 2010; Riley *et al.*, 2007b; Tipton, 2015; Wei and Higgins, 2013b).

Multiple levels

In the same applications, summary effect sizes have natural levels of hierarchy (Stevens and Taylor, 2009); for example, two or more study reports can come from the same research group or laboratory. Multiple effect sizes could also have been calculated for different higher-level units such as firms, hospitals, school districts, neighbourhoods or cities. In these situations, the issue is that the inner-level effect size estimates within the outer level could be correlated. These correlations could produce unexplained heterogeneity that must be modelled to account for the hierarchical structure (Konstantopoulos, 2011;

Stevens and Taylor, 2009). As pointed out by Stevens and Taylor (Stevens and Taylor, 2009), it is essential to consider these sources of correlations among studies, as correlated estimates tend to over-influence the pooled estimates, and the down-weighting implied by random effects models could reduce the influence of these studies. Some attempts at multilevel meta-analytic models have been proposed to combine the results of several systematic reviews (panoramic meta-analysis), for example, when the same treatment has been associated with different surgery or cancer type (Hemming *et al.*, 2013). Multilevel meta-analysis models also allow modelling the covariance among multiple outcomes without knowing the within-study covariance matrix (Van den Noortgate *et al.*, 2013, 2015).

Longitudinal studies

Another example of recent applications of meta-analysis characterised by a complex study design is when meta-analysis is applied to studies with a longitudinal design, and these studies could report effect sizes measured at several time points. As the estimates are calculated on the same subjects, these estimates could be correlated. Meta-analysis of longitudinal studies reporting multiple effect sizes could use the information on the correlation among repeated-effect size measures to produce unbiased and more efficient estimates. Some authors have proposed multivariate or multilevel meta-analytic models to pool repeated summary estimates at common observations time-points (Ahn and French, 2010; Ishak *et al.*, 2007; Lopes *et al.*, 2003; Musekiwa *et al.*, 2016; Peters and Mengersen, 2008; Trikalinos and Olkin, 2012). Multivariate meta-analysis allow the user to model the autocorrelation among effect estimates using different parametrisation of the within-study and between study covariance matrix; for example, within or between study covariance terms could be parametrised to model an autoregressive process (AR(1)) of the effect estimates, but these models require estimates to be calculated at common, evenly pre-defined time-points across all the studies.

No previous study has investigated the problem of pooling multiple results from different studies obtained at different time points. In this case, the between-study covariance matrix could be modelled to give the covariance between random terms as a function of the "temporal distance" among measures, as for Gaussian Markov Fields Models (Rue and Held, 2005).

Non linear or complex associations

Dose-response meta-analytic models have been used to summarise a linear association between the exposure-response across epidemiological studies (Berlin *et al.*, 1993). The standard approach consists of a two-stage procedure where, in the first stage, for each study, the slope of the association and its precision are estimated, taking into account the within-study correlation among effect estimates at different levels of the exposure. In the second stage, the slopes are pooled using standard fixed or random effects methods. Within a study, the association between the exposure and the outcome could be non-linear. To describe the non-linear association, a set of parameters would be needed (e.g. using polynomials, fractional polynomials or splines), and these parameters are correlated as estimates in the same sample. In this case, a two-stage procedure using a multivariate meta-analytic model has been proposed (Gasparrini *et al.*, 2012; Liu *et al.*, 2009; Orsini *et al.*, 2011; Rota *et al.*, 2010): in the first stage, a set of parameters characterising the non-linear associations and their covariance matrix are estimated on each study taking into account the within-study correlation among effect estimates at different levels of the exposure, and at the second stage, the set of parameters are pooled using multivariate meta-analysis.

Other non-standard applications

Other non-standard meta-analysis are the estimation of the relationship between the outcome of interest and surrogate markers (Buyse *et al.*, 2000), studies of accuracy of diagnostic/screening tests (Ma *et al.*, 2016), cross-design synthesis (Larose and Dey, 1997; Prevost *et al.*, 2000), and split/analysis/meta-analysis methods applied to Big Data (Cheung and Jak, 2016).

A higher grade of complexities arises when some of the proposed extensions overlap; Riley and colleagues (Riley *et al.*, 2017) presented examples of meta-analysis with multiple treatments and multiple outcomes, or meta-analysis with multiple treatments, with each of the multiple treatments characterised via a dose-response effect on the outcome.

Software for more complex meta-analysis

Software for multivariate meta-analysis

Multivariate meta-analysis models can be fitted using specific software: for instance the packages metafor, mvmeta, mmeta, mtvmeta, metaSEM, and robumeta in the R environment, the user-written package *mvmeta* in Stata, as well as using statistical software specifically designed to fit multilevel models, such as *MlWin* (Goldstein *et al.*, 2000; Rasbash *et al.*, 2000; Thompson *et al.*, 2001; Turner *et al.*, 2000).

The mvmeta package in R was developed by my supervisor, Prof. Antonio Gasparrini, and it performs fixed and random-effects (single level) multivariate meta-analysis and meta-regression. The mvmeta package computes maximum likelihood and restricted maximum likelihood estimates through a Quasi-Newton algorithm. In the metafor package, the ML and REML estimates are calculated through iterative numerical procedures based on the Fisher scoring algorithm, but the implementation is based on the full marginal model and could face computational issues with complex random-effects structure and a high number of studies and or repetitions. The metaSEM implements the Structural Equation Model (SEM) approach for meta-analysis proposed by Cheung (Cheung, 2014a,b, 2015). This framework use iterative numerical methods to estimate fixed effects and random-variances parameters, through ML and REML procedures. Although the approach linking meta-analysis and SEM is interesting, it requires additional knowledge of SEM, and it may not be easy for the user to specify a complex model using SEM parametrisation. The rubumeta package estimates the between-study variances using moment methods. It calculates standard errors of the mean summary effects using a robust method that does not need information on the within-study covariance matrix (Hedges *et al.*, 2010; Tipton, 2015). Within Stata, the *mvmeta* programme was developed by Ian White (White *et al.*, 2011). In this package, between-study variances are estimated using ML and REML through the Newton-Raphson method and an extension of the moment methods described by Jackson and colleagues (Jackson *et al.*, 2010). The inference is based on the asymptotic standard distribution for the Wald statistics. The standard error provided for an REML analysis allows uncertainty in estimating the between-study covariance matrix by inverting the second derivative matrix of the restricted likelihood.

Software for multilevel meta-analysis

The two meta-analysis packages metafor and metaSEM, both in the R environment, also allow the user to specify and estimate multilevel meta-analytic models, but both packages only allow up to two levels. These packages were described in the previous paragraph.

Software for longitudinal meta-analysis

Longitudinal meta-analysis with repeated measures at common observation points can be fitted using the packages that allow multivariate meta-analysis models. Some of these programs, such as metafor, and mvmeta, enable the user to specify the between-study covariance matrix in terms of the autoregressive process. This specification is useful when the different points are evenly spaced. To date, only the statistical package metafor can fit longitudinal meta-analysis with a continuous-type autoregressive structure in which the estimates refer to different time points that are not evenly spaced.

Software for dose-response meta-analysis

The R package dosresmeta, developed by Crippa and colleagues, estimates the parameters of non-linear dose-response meta-analysis using a two-stage approach. Interestingly, the package has been recently updated with the possibility to model non-linear meta-analysis with a one-stage model (Crippa *et al.*, 2019).

Software for specific application meta-analysis

Several softwares has been developed specifically for extended applications of meta-analysis. For example, within the R environment we could cite the packages MAMA for meta-analysis of gene-expression studies, mada for meta-analysis of diagnostic studies, and netmeta to perform network meta-analysis (Polanin *et al.*, 2017; Schwarzer *et al.*, 2015).

Meta-analysis with general statistical software

Given the link with mixed-effects (multilevel or hierarchical) models, standard and more complex meta-analytic models can be estimated using general statistical software. For example, several authors

use the *SAS PROC MIXED* procedure (Konstantopoulos, 2011; Van den Noortgate *et al.*, 2013, 2015; Riley *et al.*, 2007a; Singer, 1998; Van Houwelingen *et al.*, 2002), while Bagos (Bagos, 2015) shows how to fit complex meta-analytic models using Stata *gllamm*. Notably, several scholars used software specifically designed to fit multilevel models *MlWin* (Goldstein *et al.*, 2000; Rasbash *et al.*, 2000; Thompson *et al.*, 2001; Turner *et al.*, 2000). Within the R environment, the package nlme allows to fit linear and non-linear mixed effects models and could be used to fit (basic or complex) meta-analytic models. However, there are differences in the estimates of the fixed-effects coefficients and their standard errors with respect to other packages (e.g. metafor) both using maximum likelihood and restricted maximum likelihood procedures.

These are important contributions, but they require advanced knowledge of statistical and computational aspects and could be difficult for the user to customize complex meta-analytic models. Moreover, using general statistical software for multilevel modelling needs to fix the within-study error as known, and this could be feasible for standard univariate meta-analysis on which there is a single known variance for each estimate, but it can be more problematic for the extensions (e.g., correlated known errors) on which the between study covariance matrix needs to be supplied by the user.

1.5 Aims and objectives

The literature review presented in the previous sections highlights that meta-analysis of complex studies is a hot topic with contributions from several scholars. Despite the interest in the topic, there are still difficulties in applying complex meta-analytic methods in applied meta-research problems, e.g. for modelling complex risks associated with environmental factors (Riley *et al.*, 2017; Sutton and Higgins, 2008). These difficulties are partly due of a lack of a unified framework on which all the complex models can be formulated. In section 1.4, I showed several models for the different specific settings for complex meta-analysis, but these were presented as separate developments rather than as various extension within the same modelling structure. The main aim of this PhD is to develop a general (united) framework for complex meta-analysis and to derive multiple design extensions of the classical two-stage method, all implemented using the extended random framework for complex meta-analysis. The definition of a framework and related optimal inferential procedures will be implemented in a user-

friendly software to further increase the use of theoretically-sound and efficient statistical methods in complex meta-analytic research with a particular focus on environmental epidemiology.
In particular, the specific objectives of the PhD are:

1. Objective 1: to develop an extended random-effects framework for complex meta-analysis. The extended framework will allow the specification of several non-standard applications of meta-analysis (e.g., multivariate, multilevel, longitudinal, dose-response, and their combination) within a common model using consistent estimation and inferential procedures.

2. Objective 2: to derive multiple design extensions of the classical two-stage method, all implemented using the extended random framework for complex meta-analysis. This framework will offer a flexible and generally applicable tool to implement extensions of the classical two-stage study design used in environmental epidemiology.

3. Objective 3: to implement the analytic and inferential extended framework in a new R Package mixmeta. The availability of user-friendly software designed to be coherent with the formulation of an extended random-effects framework for complex meta-analysis will facilitate the implementation, use, and dissemination of complex meta-analytic procedures in traditional and non-standard applications.

4. Objective 4: to illustrate the application of the extended random-effect framework for complex meta-analysis in environmental epidemiology studies. These studies are characterised by complex settings with multiple levels of hierarchies, non-linearity, spatial and temporal structure and represent an ideal context to demonstrate the advantages of the extended framework.

Chapter 2

Contribution of selected publications

In this chapter, I provide a summary of the contribution of my research to the field of statistical methods for complex meta-analysis and its application in environmental epidemiology, coherent with the PhD objectives illustrated in Section 1.5. In the next Section 2.1, I will introduce the main methodological developments of this thesis (PhD's Objectives 1 and 2), while the implementation of these methods on the statistical environment R (PhD's Objective 3) is illustrated in Section 2.2. Finally, Section 2.3 will provide an overview of the publications. I will illustrate each publication in this section, emphasising this PhD's substantive contribution to methodological developments (PhD's Objectives 1 and 2) and the environmental epidemiology field (PhD's Objective 4). Moreover, I will highlight my role in the various steps from study planning to article publication.

2.1 Methodological developments

In Section 1.4, I presented extensions that deal with more complex meta-analytical problems. These include, potentially among others, multivariate models for pooling multiple outcomes or multi-parameter associations (Gasparrini *et al.*, 2012; Jackson *et al.*, 2011), network meta-analysis for indirect mixed-treatment comparison (Riley *et al.*, 2017), multilevel versions for hierarchically-structured studies

(Stevens and Taylor, 2009), dose-response meta-analysis (Crippa *et al.*, 2019; Orsini *et al.*, 2011), and longitudinal meta-analysis for studies reporting multiple estimates at different times (Ishak *et al.*, 2007). Although these extensions were presented separately, all of them can be described as examples on which multiple estimates are obtained within each study; this data structure creates a dependence among estimates. Consequently, the dependence within and or between studies creates more complex correlation structures that need to be modelled or accounted for in obtaining pooled estimates.

In my PhD, coherently with Objective 1, I reviewed and brought together these different developments into a coherent extended mixed-effects framework for meta-analysis (Sera *et al.*, 2019a). The extended framework for meta-analysis is built on the known link between meta-analysis and linear mixed-effects (LME) models, where patterns of effect sizes are modelled through a flexible structure of fixed and random terms (Bagos, 2015; Berkey *et al.*, 1998; Goldstein *et al.*, 2000; Konstantopoulos, 2011; Van den Noortgate *et al.*, 2013; Stram, 1996; Thompson *et al.*, 2001; Turner *et al.*, 2000; Van Houwelingen *et al.*, 2002).

As discussed in section 1.2, the two-stage design has become a standard tool in environmental epidemiology to model multi-location data. However, its standard form is rather inflexible and poses important limitations for modelling complex risks associated with environmental factors. In my contribution, in accordance with Objective 2 of this PhD, I illustrate multiple design extensions of the classical two-stage method (Sera and Gasparrini, 2022), all implemented using the extended random framework for complex meta-analysis . This framework offers a flexible and generally applicable tool to implement extensions of the classical two-stage study design used in environmental epidemiology.

2.2 Software implementation

The extended mixed-effects framework for meta-analysis has been implemented within R, a free programming language and software environment for statistical computing and graphics.

The choice to create the two R packages has been motivated by several considerations. First, the approach is based on relatively complex statistical methods and routines requiring non-trivial computing skills to provide stable results. The availability of fully-documented packages in a freely-available software can promote the application of the techniques by other research teams. Second, the production of

the packages involves generalising the methodologies beyond the specific data and models I have used in my research. The packages are therefore expected to be applicable in a broader range of analyses and potentially easier to improve and extend.

Following Objective 3 of this PhD, the extended random-effects framework for complex meta-analysis have been implemented in the R package mixmeta (`https://cran.r-project.org/web/packages/mixmeta/index.html`). The package was first released on CRAN on July 2019. At the time of writing, the current version is 1.2.0, after ten updates.

The main function in the package is `mixmeta()`, which performs fixed and random-effects meta-analysis and meta-regression. This regression-type function contains arguments to define formulae which specify the outcomes and the fixed effect linear predictors and random terms that allow setting the random structure of the model. The `mixmeta()` function calls internal functions to compute maximum likelihood and restricted maximum likelihood estimates through a hybrid Quasi-Newton and iterative reweighted least square (IGLS) algorithm. Additional functions are used, among other purposes, to obtain predictions and best linear unbiased predictions, to run heterogeneity tests and to compute fit statistics.

The package includes several datasets used for applications of the extended meta-analytical framework. Documentation of the package is provided through the reference manual (Gasparrini *et al.*, 2021), reported in Appendix A. The manual also provides examples of specific models and fully demonstrates the flexibility of the extended meta-analytical framework.

2.3 Overview of the publications

The six publications summarise my research activity within the PhD project. They include five research papers, as chapters in Part II, and a package reference material in Appendix A. I am the first and corresponding author of the five research papers and the co-author of the package reference material. The five research papers have already been published in high-impact international scientific journals.

The order of the publications has been chosen to describe a coherent research project. However, the manuscripts have been published or submitted as independent contributions, and the text included in the different chapters is not consistently linked. This section aims to provide the reader with a

summary of each publication, progressively illustrating my research's conceptual, methodological and more applied steps. In particular, the first two papers present the main methodological developments, in accordance of Objectives 1 and 2 of this PhD, while the last three research papers are related to Objective 4 of this PhD and illustrate the substantive contribution of this PhD in environmental epidemiology.

Research paper I

The first research paper, originally published as Sera *et al.* (2019a) and included in Chapter 3, represents the first methodological output coherent with Objective 1 of this PhD. In this contribution, I illustrate a general framework for meta-analysis based on linear mixed-effects models, where potentially complex patterns of effect sizes are modelled through an extended and flexible structure of fixed and random terms. This definition includes, as special cases, a variety of meta-analytical models that have been separately proposed in the literature, such as multivariate, network, multilevel, dose-response, longitudinal meta-analysis and meta-regression. As the paper's first author, I structured the methodological description of the general framework for meta-analysis based on linear mixed-effects models in agreement with the co-authors. I developed the algebra for this model family. I independently chose the example included in the paper and performed the analysis. I was the lead author of the manuscript and acted as the corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Research paper II

The second research paper, originally published as Sera and Gasparrini (2022) and included in Chapter 4, along with the documentation added in the online Appendix, represents the second methodological output and follows the Objective 2 of this PhD. In this contribution, I illustrate extensions of the classical two-stage study design used in environmental epidemiology. I introduce the extended two-stage design and its features in the article, including the design structure and related modelling framework. Then, after presenting the specific example and the related dataset, I demonstrate applications of the various design extensions in multiple case studies using multi-location analyses of

temperature and air pollution health risks. Methodological notes, data, and R scripts for reproducing the examples are added as supplementary material.
As the first author of the paper, I structured the paper. I independently chose the example included in the paper and performed the analysis. I was the lead author of the manuscript and acted as the corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Research paper III

The third research paper, originally published as Sera *et al.* (2019b), is included in Chapter 5. In this contribution, related to Objective 4 of the PhD, I explored the role of urban characteristics in modifying the direct effects of temperature on health. In particular, I used a multi-country dataset to study the effect modification of temperature-mortality relationships by a range of city-specific indicators. This study is an example of a standard two-design applied in environmental epidemiology. I used distributed lag non-linear time-series models in the first stage and multivariate meta-regression models in the second stage to estimate fractions of mortality attributable to heat and cold (AF%) in each city. In the last third stage, I used meta-regression models to evaluate each indicator's effect modification across cities. The analysis, even if it presents complex aspects as the multivariate meta-regression models used in the second stage, represents an example of a "standard" two-stage design from which extensions, e.g. multilevel and or longitudinal can be derived.
As the first paper's first author, I coordinate the work with all the co-authors discussing the study design, research question, and the relevant epidemiological and public health issues. I independently conducted the analysis, discussing the analytical approaches, interpretation and conclusions with the research team. I took the lead in writing the manuscript and acted as the corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Research paper IV

The fourth research paper, originally published as Sera *et al.* (2020) is included in Chapter 6. In this contribution, I used a multi-country, multi-city, longitudinal design to quantify the independent role

of air conditioning in attenuating heath-related health risk. This study is an example of a two-stage design with a longitudinal structure that can be analysed using the extended framework for complex meta-analysis (Objective 4 of the PhD). I used collected the daily time series of mortality, mean temperature, and yearly air conditioning prevalence for 311 locations in Canada, Japan, Spain, and the USA between 1972 and 2009. In the first stage, I fitted a quasi-Poisson regression combined with distributed lag non-linear models for each city and sub-period to estimate summer-only temperature–mortality associations. In the second stage, I used the extended random-effects framework to implement a multilevel, multivariate spatio-temporal meta-regression model to evaluate the effect modification of air conditioning on heat–mortality associations. The modelling strategy allows to compute relative risks and fractions of heat-attributable excess deaths under observed and fixed air conditioning prevalences. As the paper's first author, I coordinate the work with all the co-authors discussing the study design, research question, and the relevant epidemiological and public health issues. I independently conducted the analysis, discussing the analytical approaches, interpretation and conclusions with the research team. I took the lead in writing the manuscript and acted as the corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Research paper V

The fifth research paper, originally published as Sera *et al.* (2021), is included in Chapter 7. In this contribution, I used a two-stage ecological modelling approach to examine the impact of meteorological variables on SARS-CoV-2 transmission between cities across the globe while accounting for confounding of non-pharmaceutical interventions and city-level covariates. It represents an excellent example of a two-stage design with a multilevel structure with cities nested within a country (Objective 4 of the PhD). In the first stage, I used estimated the effective reproduction number (Re), in each city, over a city-specific time window early in the epidemic. In the second 'cross-sectional' stage, I estimated the association between city-level Re with each meteorological variable, controlling for confounding by total population, population density, GDP per capita, percentage of population >65 years, pollution levels (i.e. particulate matter, PM2.5), and the lagged OxCGRT Government Response Index at the end-point of the selected time-window (lagged by ten days). In the model I considered the two-level

(cities and countries) structure of the data using a multilevel meta-regression model implemented trought the extended random-effects framework for complex meta-analysis.

As the paper's first author, I coordinate the work with all the co-authors discussing the study design, research question, and the relevant epidemiological and public health issues. I independently conducted the analysis, discussing the analytical approaches, interpretation and conclusions with the research team. I took the lead in writing the manuscript and acted as the corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Part II

Selected publications

Chapter 3

Research paper I

Title: An extended mixed-effects framework for meta-analysis.

Author(s): Francesco Sera, Benedict Armstrong, Marta Blangiardo, Antonio Gasparrini.

Journal/Publisher: Statistics in Medicine.

Type of publication: Research paper.

Stage of publication: Published online on October 24, 2019 as doi.org/10.1002/sim.8362.

URL: `https://onlinelibrary.wiley.com/doi/10.1002/sim.8362`.

Academic peer-reviewed: Yes.

Candidate's role: See Section 2.3.

Senior author: (Prof. Antonio Gasparrini)

Received: 14 January 2019 | Revised: 31 July 2019 | Accepted: 12 August 2019

DOI: 10.1002/sim.8362

RESEARCH ARTICLE

WILEY Statistics in Medicine

An extended mixed-effects framework for meta-analysis

Francesco Sera[1,2] | **Benedict Armstrong**[1,2] | **Marta Blangiardo**[3] | **Antonio Gasparrini**[1,2]

[1]Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London, UK

[2]Centre for Statistical Methodology, London School of Hygiene & Tropical Medicine, London, UK

[3]Department of Epidemiology and Biostatistics, Imperial College London, London, UK

Correspondence
Francesco Sera, Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.
Email: francesco.sera@lshtm.ac.uk

Funding information
Medical Research Council UK, Grant/Award Number: MR/M022625/1 and MR/R013349/1

Standard methods for meta-analysis are limited to pooling tasks in which a single effect size is estimated from a set of independent studies. However, this setting can be too restrictive for modern meta-analytical applications. In this contribution, we illustrate a general framework for meta-analysis based on linear mixed-effects models, where potentially complex patterns of effect sizes are modeled through an extended and flexible structure of fixed and random terms. This definition includes, as special cases, a variety of meta-analytical models that have been separately proposed in the literature, such as multivariate, network, multilevel, dose-response, and longitudinal meta-analysis and meta-regression. The availability of a unified framework for meta-analysis, complemented with the implementation in a freely available and fully documented software, will provide researchers with a flexible tool for addressing nonstandard pooling problems.

KEYWORDS

dose-response, longitudinal, meta-analysis, mixed-effects models

1 | INTRODUCTION

Meta-analysis has become a standard method to summarize evidence in various scientific fields.[1] Traditional applications require a set of single effect size estimates that are collected from multiple independent studies. However, extensions to deal with more complex meta-analytical problems have been presented. These include, potentially among others, multivariate models for pooling multiple outcomes or multiparameter associations,[2,3] network meta-analysis for indirect mixed-treatment comparison,[4] multilevel versions for hierarchically structured studies,[5] dose-response meta-analysis,[6,7] and longitudinal meta-analysis for studies reporting multiple estimates at different times.[8] Although these extensions were presented separately, all of them can be described as cases where multiple observations are collected within each study, and their dependence within and/or between studies creates more complex correlation structures that need to be modeled or accounted for.

In this contribution, we review and bring together these different developments into a coherent unified framework, built on the known link between meta-analysis and linear mixed-effects (LME) models, where patterns of effect sizes are modeled through a flexible structure of fixed and random terms.[9-17] The manuscript is organized as follows: the analytic formulation of the unified framework is introduced in Section 2, followed by estimation and inferential procedures in Section 3. Specific applications are presented in Section 4, including analytic definitions linked to the general framework,

Abbreviations: AIC, Akaike information criteria; BCG, Bacillus Calmette-Guerin; BIC, Bayesian information criteria; BLUP, best linear unbiased prediction; GLS, generalized least squares; LME, linear mixed-effects; ML, maximum likelihood; OR, odds ratio; REML, restricted maximum likelihood; (R)IGLS, (restricted) iterated generalized least squares; RMSE, root mean square error; RR, incidence relative rate; TB, tuberculosis.

and illustrations through real-data examples. Section 5 describes the software implementation of the modeling framework in the new R package mixmeta, while Section 6 presents the results of a simulation study. Section 7 draws some conclusions. R code and data for replicating examples and simulation results are added as supplementary material, with an updated version available at the personal website and GitHub page of the last author.

2 | A MIXED-EFFECTS FRAMEWORK FOR META-ANALYSIS

A unified modeling framework can be defined by casting the meta-analytical problem as a LME model. In general terms, we assume that there is a set of n total measures effect sizes (observations) of k different outcomes, representing *units* of analysis aggregated in $i = 1, \dots, m$ *groups* that are considered independent. Additional $L-1$ inner levels of grouping could exist within each of the m outer groups, for a total of L grouping levels. Grouping levels can be represented by studies themselves, as in standard meta-analysis, or be defined either between or within studies. An extended mixed-effects metaregression model for the $\mathbf{y}_i$ effect sizes (outcomes) in group i can be generally written as

$$\begin{aligned} &\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, i = 1, \dots, m, \\ &\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}_i), \boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{S}_i). \end{aligned} \tag{1}$$

Here, $\mathbf{X}_i\boldsymbol{\beta}$ defines the fixed effects that represent the population-averaged outcomes in terms of p unit-level meta-predictors in the design matrix $\mathbf{X}_i$, with fixed-effects coefficients $\boldsymbol{\beta}$. The random part of the model, $\mathbf{Z}_i\mathbf{b}_i$, describes the deviation from the population averages in terms of q predictors defined at different grouping levels and composing the random-effects design matrix $\mathbf{Z}_i$, with coefficients $\mathbf{b}_i$. The vector $\boldsymbol{\epsilon}_i$ defines the unit-level sampling errors. The model has marginal distribution $\mathbf{y}_i \sim \mathrm{N}(\mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}_i)$, where the marginal (co)variance matrix $\boldsymbol{\Sigma}_i = \mathbf{S}_i + \mathbf{Z}_i\boldsymbol{\Psi}_i\mathbf{Z}_i^{\mathrm{T}}$ is given by the sum of within-group errors (assumed known) and between-group random effects, defined by (co)variance matrices $\mathbf{S}_i$ and $\boldsymbol{\Psi}_i$, respectively. The latter is composed of a block-diagonal form of level-specific matrices $\boldsymbol{\Psi}^{(1)}, \dots, \boldsymbol{\Psi}^{(L)}$ (from outer to inner levels), defined by a set of parameters $\boldsymbol{\xi}$ dependent on their specific form (eg, unstructured, (heterogeneous) compound symmetry, and (heterogeneous) autoregressive of first order)[18] and on constraints for ensuring positive definiteness. These matrices are expanded consistently with the inner structure of each group, similarly to $\mathbf{Z}_i$ (see Section 4.3 for algebraic details).

3 | ESTIMATION

Likelihood functions

The unknown parameters of the model in Equation (1) are the vector $\boldsymbol{\beta}$ of fixed effects and the vector $\boldsymbol{\xi}$ that characterizes the set of level-specific (co)variance matrices of random effects composing $\boldsymbol{\Psi}_i$. These can be estimated through (restricted) maximum likelihood (ML and REML) estimators, with the marginal (restricted) log-likelihood functions derived from the LME framework[19,20] as

$$\begin{aligned} l(\boldsymbol{\beta}, \boldsymbol{\xi}|\mathbf{y}) &= -\frac{1}{2}n\log(2\pi) - \frac{1}{2}\sum_{i=1}^{m}\log|\boldsymbol{\Sigma}_i| - \frac{1}{2}\sum_{i=1}^{m}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}), \\ l_R(\boldsymbol{\xi}|\mathbf{y}) &= -\frac{1}{2}(n-p)\log(2\pi) + \frac{1}{2}\log\left|\sum_{i=1}^{m}\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i\right| - \frac{1}{2}\log\left|\sum_{i=1}^{m}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right| \\ &\quad - \frac{1}{2}\sum_{i=1}^{m}\log|\boldsymbol{\Sigma}_i| - \frac{1}{2}\sum_{i=1}^{m}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}). \end{aligned} \tag{2}$$

Note that the REML version l_R only depends on $\boldsymbol{\xi}$, as it is obtained by reprojecting $\mathbf{y}$ using a set of $n-p$ orthogonal error contrasts, a transformation which is reflected algebraically in the inclusion of the two additional determinant terms in Equations (2). The estimators derived from REML are generally considered superior, particularly in regard to the estimation of random components, as they account for the loss of degrees of freedom in the estimation of $\boldsymbol{\beta}$ that induces a downward bias in the ML counterpart. However, they pose limitations for hypothesis testing, as discussed in the following.

For known $\boldsymbol{\xi}$, ML estimates of the fixed-effects coefficients and their associated (co)variance matrix can be easily obtained by generalized least squares (GLS) estimators, expressed in closed form as

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \left(\sum_{i=1}^{m}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right)^{-1}\sum_{i=1}^{m}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{y}_i,\\ V(\hat{\boldsymbol{\beta}}) &= \left(\sum_{i=1}^{m}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right)^{-1}.\end{aligned} \tag{3}$$

Fixed-effects meta-analytical models can be simply estimated using Equation (3) by setting $\boldsymbol{\Sigma}_i = \mathbf{S}_i$. For random-effects models, when the random part is unknown, the joint estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ requires iterative methods for maximizing the likelihood functions in Equations (2). For computational convenience, a *profiled* approach is preferable, where iterative algorithms are defined in terms of parameters $\boldsymbol{\xi}$ only, and values of $\hat{\boldsymbol{\beta}}$ are obtained by Equation (3) and plugged in at each iteration, until convergence. Alternative algorithms have been proposed, such as Newton-Raphson, expectation-maximization, and (restricted) iterative generalized least squares (IGLS and RIGLS), each of them with different properties.[18,21,22] See Section 5 for additional details.

Hypothesis testing and model comparison

Inferential procedures follow standard LME theory and concern the fixed-effects parameter vector $\boldsymbol{\beta}$ and the set of random-effects (co)variance matrices $\boldsymbol{\Psi}^{(\ell)}$, with $\ell = 1, \ldots, L$. Regarding the fixed effects, under the marginal model and replacing $\boldsymbol{\Sigma}_i$ in Equation (3) with its ML or REML estimate through $\hat{\boldsymbol{\xi}}$, the vector $\hat{\boldsymbol{\beta}}$ follows a multivariate normal distribution with (co)variance matrix $V(\hat{\boldsymbol{\beta}})$. These results can be used to derive approximated confidence intervals and (multivariate) Wald tests for specific coefficients or their linear combinations. If the Wald test gives significant results, a common question is which particular linear combinations of the coefficients are significantly different from zero. The common example is where we find a difference on the k effect sizes, and we wish to perform all possible comparisons. A simultaneous comparison procedure that maintains the overall type I error was proposed by Goldstein.[23] Comparison between nested models can be performed through likelihood ratio (LR) tests, or more generally using fit statistics such as the Akaike or Bayesian information criteria (AIC and BIC), each of which is easily computed using the (restricted) ML values from Equation (2).

LR tests and AIC/BIC can also be used for hypothesis testing and model selection, for instance by comparing alternative structures for random-effects (co)variance matrices $\boldsymbol{\Psi}^{(\ell)}$ or by assessing the presence of heterogeneity at each grouping level ℓ. However, it must be noted that the chi-square distribution is a poor approximation to the actual distribution of the LRT statistic when applied to a large number of parameters, and when testing heterogeneity, with the null hypothesis $\boldsymbol{\Psi}^{(\ell)} = \mathbf{0}$ lying on the boundary of the parameters space. More importantly, the REML log-likelihood function is not invariant to one-to-one reparametrization of the fixed effects, as this changes the specification of the error contrasts; therefore, LR tests and AIC/BIC can only be used to compare REML models with the same fixed-effects specification.

In addition to inferential tools borrowed directly from LME models, other statistics traditionally used in meta-analysis to assess the presence and amount of heterogeneity can be easily extended in this more general mixed-effects framework, for instance the Cochran Q and I^2.[24,25] These can be defined as

$$\begin{aligned}Q &= \sum_{i=1}^{m}(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}})^{\mathrm{T}}\mathbf{S}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}),\\ I^2 &= \max\left\{\frac{Q-n+p}{Q}, 0\right\},\end{aligned} \tag{4}$$

where $\hat{\boldsymbol{\beta}}$ are estimated by the correspondent fixed-effects model with no random term. The Cochran Q statistic follows a χ^2_{n-p} distribution under the hypothesis of no heterogeneity, and can be used to define the related test, while the I^2 statistic quantifies the amount of heterogeneity as the proportion of total variation above that related to sampling error.

Prediction

In this complex meta-analytic setting, inferential procedures can be complemented with prediction tools that inform about potentially complex relationships that are pooled across studies, including for example multivariate and non-linear

associations.[3] In this context, predictions offer a method to link specific values of metaregressors defined at any grouping level with effect size expectations. Given a set of unit-level metapredictors $\mathbf{x}_0$ that form the design matrix $\mathbf{X}_0$ depending on the specific model (see Section 4), the (marginal) predicted mean $\hat{\mathbf{y}}_0$ with (co)variance matrix $V(\hat{\mathbf{y}}_0)$ are obtained by

$$\begin{aligned} \hat{\mathbf{y}}_0 &= \mathbf{X}_0\hat{\boldsymbol{\beta}}, \\ V(\hat{\mathbf{y}}_0) &= \mathbf{X}_0 V(\hat{\boldsymbol{\beta}})\mathbf{X}_0^{\mathrm{T}}. \end{aligned} \tag{5}$$

In addition to the marginal level, improved study-specific estimates can be obtain as *best linear unbiased predictions* (BLUPs). These are interpreted as trade-off between $\mathbf{y}_i$ and $\hat{\mathbf{y}}_i$, with estimates of effect sizes borrowing information within and/or between studies. BLUPs can be defined as conditional expectations given the random effects, and its empirical version $\hat{\mathbf{y}}_{b_i}$ and (co)variance matrix $V(\hat{\mathbf{y}}_{b_i})$ are provided as

$$\begin{aligned} \hat{\mathbf{y}}_{b_i} &= \mathbf{X}_0\hat{\boldsymbol{\beta}} + \mathbf{Z}_i\hat{\boldsymbol{\Psi}}_i\mathbf{Z}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_i^{-1}(\hat{\mathbf{y}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}), \\ V(\hat{\mathbf{y}}_{b_i}) &= \mathbf{X}_0 V(\hat{\boldsymbol{\beta}})\mathbf{X}_0^{\mathrm{T}} + \mathbf{Z}_i\hat{\boldsymbol{\Psi}}_i\mathbf{Z}_i^{\mathrm{T}} - \mathbf{Z}_i\hat{\boldsymbol{\Psi}}_i\mathbf{Z}_i^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_i^{-1}\mathbf{Z}_i\hat{\boldsymbol{\Psi}}_i\mathbf{Z}_i^{\mathrm{T}}. \end{aligned} \tag{6}$$

It is interesting to note that, in a multilevel context, BLUPs can be defined also as predictions at higher levels of grouping. For instance, BLUPs at level $\ell \leq L$ are derived by including in $\hat{\boldsymbol{\Psi}}_i$ and $\mathbf{Z}_i$ only the random-effects components corresponding to the grouping levels in ℓ and above (see Section 4.3 for an algebraic definition of levels).

4 | SPECIFIC APPLICATIONS

Different models for meta-analysis can be expressed as special versions of the general framework in Equation (1). These includes the standard methods, extensions mentioned above, and their combinations, among potentially other models. In this section, we describe the most common cases, graphically represented in Figure 1, highlighting their distinctive aspects and their link with the general framework.

Figure 1 shows how extensions of the standard model are generally characterized by repeated measures and groupings that induce patterns of correlation across effect sizes. As in LME models, these potentially complex structures can be flexibly modeled by a combination of fixed and random terms, optionally including meta-predictors with alternative parameterizations, for instance indicators and continuous smooth functions. In the following, we illustrate each case by

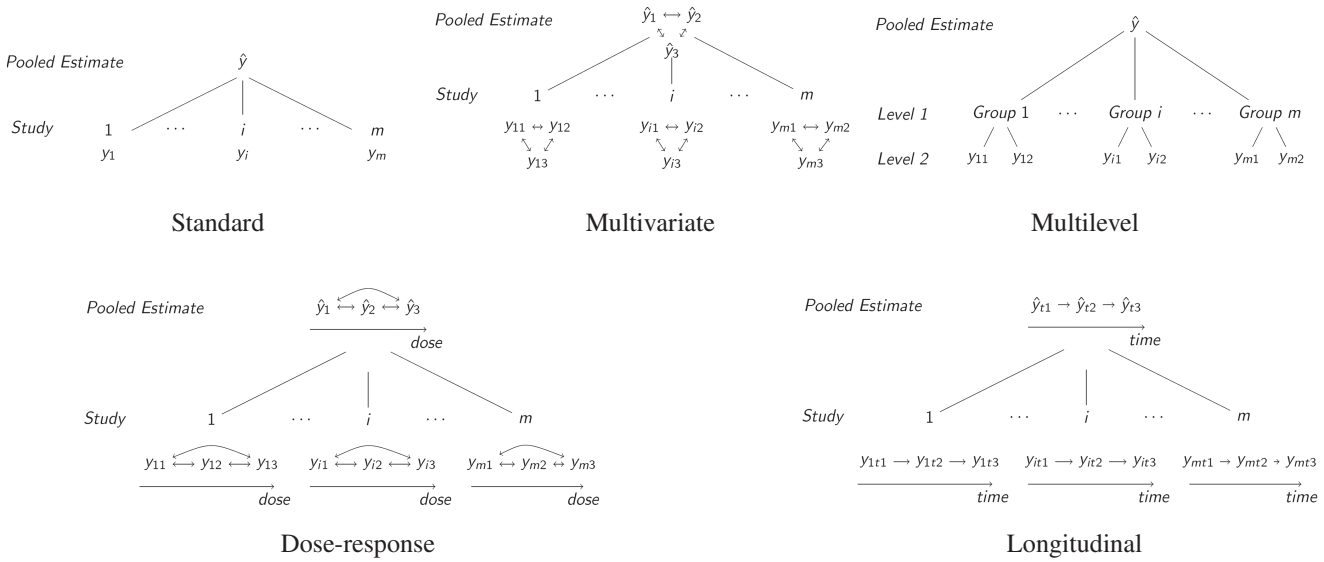


FIGURE 1 Graphical illustration of data structures in specific applications of the extended framework for meta-analysis

replicating and extending real-data meta-analyses from published studies, reproduced in the R scripts provided in the supplementary material. Details on the data and substantive context can be found in the references or in the help pages of the R package mixmeta.

4.1 | Standard meta-analysis

The objective of a *standard meta-analysis* is to obtain a summary (pooled) estimate from single effect sizes estimated separately in independent studies. The basic estimation procedures are based on the computation of weighted averages across studies, with weights proportional to the precision of the estimates.[26] The pooled estimate can be derived under *fixed-effects* or *random-effects* assumptions,[27,28] with random-effects models incorporating the underlying between-study variation into the weights.[24]

Analytic formulation

Several authors have already pointed out that the random-effects meta-analysis can be expressed as a LME model in a regression context.[14,29] Specifically, the standard model for a set of effect sizes y_i can be defined as

$$\begin{aligned} &y_i = \beta_0 + b_i + \epsilon_i, i = 1, \ldots, m, \\ &b_i \sim \mathrm{N}(0, \tau^2), \epsilon_i \sim \mathrm{N}\left(0, s_i^2\right), \end{aligned} \tag{7}$$

where β_0 is the pooled effect, b_i are the study-specific random effects distributed with between-study variance τ^2, and ϵ_i is the error term distributed with known within-study variance s_i^2. This standard model represents the simplest case of the general extended framework in Equation (1), with $n = m$ (a single estimate from separate studies), and scalar quantities $\mathbf{X}_i = \mathbf{Z}_i = 1$, $\boldsymbol{\beta} = \beta_0$, $\mathbf{b}_i = b_i$, $\boldsymbol{\epsilon}_i = \epsilon_i$, $\boldsymbol{\Psi} = \tau^2$, and $\mathbf{S}_i = s_i^2$. In fixed-effects models, the term b_i does not exist. The model in Equation (7) can be extended to *meta-regression* by defining a set of study-level predictors $\mathbf{x}_i = [x_{i1}, \ldots, x_{ip}]^{\mathrm{T}}$ and by setting $\mathbf{X}_i = \mathbf{x}_i^T$, where usually $x_{i1} = 1$ specifies the intercept term.

Illustrative example

In this first example we consider a meta-analysis and meta-regression performed by Colditz and colleagues that evaluate the efficacy of the Bacillus Calmette-Guerin (BCG) vaccine for preventing tuberculosis (TB).[30] The dataset was used by several authors to illustrate their random-effects regression models.[15,17,29] The data refers to 13 prospective clinical trials that estimated the odds ratio (OR) of TB between groups vaccinated with the (BCG) vaccine and non-vaccinated control populations.

We apply the general framework to estimate the parameters for the log-OR β_0 and between-study variance τ^2 in Equation (7) using an ML estimator (see Section 3), replicating the results reported by Van Houwelingen and colleagues.[17] Consistently, the estimated OR is 0.476 (95%CI: 0.336 to 0.675), with a clear indication of a protective effect of BCG vaccine, and the estimated τ^2 is 0.302, with suggestions of a large heterogeneity ($I^2 = 92.6\%$). Similarly to the original analysis, we can investigate the influence on vaccine efficacy of various meta-predictors such as study location, year of publication and method of treatment allocation. For instance, adding latitude in a meta-regression model reduces the between-study variance and residual heterogeneity ($\tau^2 = 0.004$, and $I^2 = 56.2\%$). The coefficient for latitude is -0.033 (95%CI: -0.039 to -0.026), indicating an improved efficacy of the vaccine at higher latitudes.

4.2 | Multivariate meta-analysis

An important extension of the standard univariate model in Equation (7) is *multivariate meta-analysis*, in which each study still reports single estimates, but for multiple effect sizes referring to different outcomes, such as disease free and overall survival risks in cancer patients.[2,31] The same model has been extended to other contexts, for instance to pool results from multiparameter functions defining nonlinear relationships,[3] or meta-analysis of diagnostic accuracy tests.[32] A common application of multivariate models is for network meta-analysis applied in mixed-treatment comparisons, where efficiency can be gained by exploiting the correlation among effect sizes that measure relative effects across different treatments.[4,33]

TABLE 1 Example of multivariate (network) meta-analysis of 24 trials comparing alternative treatments of smoking cessation, using a consistency model and a structured between-study (co)variance matrix. Previously reported by White[35]

Comparison	Estimated log-OR	Standard error	*p*-value
Self-help versus no contact (B versus A)	0.398	0.330	0.227
Individual counselling versus no contact (C versus A)	0.702	0.196	<0.0001
Group counseling versus no contact (D versus A)	0.866	0.373	0.020

Analytic formulation

In all these applications, the k-dimensional vector $\mathbf{y}_i$ contains estimates for multiple (potentially correlated) effect sizes in each study. A model for random-effect multivariate meta-analysis can be represented as follows:

$$\begin{aligned} &\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{b}_i + \boldsymbol{\epsilon}_i, i = 1, \ldots, m, \\ &\mathbf{b}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}), \boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \mathbf{S}_i). \end{aligned} \tag{8}$$

This can be written in terms of the general framework in Equation (1) by setting $\mathbf{X}_i = \mathbf{Z}_i = \mathbf{I}_k$, $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_k]^{\mathrm{T}}$, $\mathbf{b}_i = [b_{i1}, \ldots, b_{ik}]^{\mathrm{T}}$, and $k \times k$ between-study $\boldsymbol{\Psi}$ and within-study $\mathbf{S}_i$ (co)variance matrices. Here, each of the m studies represents a group with multiple estimates, with a total of $n = k \cdot m$ units in the balanced (full-outcome) case. Missing outcomes (unbalanced case) can be accounted for by excluding related rows in the matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$. Similarly to the univariate case in Section 7, the model can extended to *multivariate meta-regression* by setting $\mathbf{X}_i = \mathbf{x}_i^T \otimes I_k$, with $\otimes$ as the Kronecker product operator and $\boldsymbol{\beta}$ being the kp-dimensional coefficient vector that defines the association of the k outcomes with the p predictors. This notation defines the same set of metapredictors for all outcomes, but it allows different associations for each of them. Alternative parameterizations can be used to define outcome-specific sets of meta-predictors or to impose the same effects across outcomes.

Illustrative example

As an example of multivariate models, we consider an application of network meta-analysis on 24 trials that compare four alternative interventions to promote smoking cessation, labelled here A–D (see Table 1).[34] Each trial compares only two or three interventions, and the joint meta-analysis allows to gain information through indirect comparisons. Treatment A is used here as the reference, and trials without an arm A were augmented with 0.01 individuals and 0.001 successes. Here, the $\mathbf{y}_i$ and $\mathbf{S}_i$ represent the log-OR of cessation and associated (co)variance error matrix of treatments B, C, and D versus A estimated in each trial, including missing values. In the following, we replicate results previously presented in the article by White.[35]

The first model is formulated under the assumption of consistency, that is allowing heterogeneity between studies but with no systematic variation across trial designs (defined by groups of trials reporting the same comparisons).[36] This model can be fitted using the general framework in Equation (1) expressed as Equation (8), with $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]^{\mathrm{T}}$ representing the three comparisons of treatments B, C, and D versus A. Following White,[35] we impose a parsimonious structure to the random-effects (co)variance matrix $\boldsymbol{\Psi}$, assuming the same variance τ^2 for all the comparisons and fixing their correlation to 0.5. The results of the consistency models are reported in Table 1: Treatments C and D are effective with respect treatment A, with substantial heterogeneity among studies ($\tau^2 = 0.454$ and $I^2 = 86.3\%$). The consistency assumption that direct and indirect evidence agree with each other can be relaxed by defining and testing design-by-treatment interactions.[33,36] This *inconsistency* model has ten fixed-effects coefficients compared with the three of the simpler version, and a global Wald test with a p-value of 0.646 fails to reject the consistency assumption.

4.3 | Multilevel meta-analysis

The previous models work under the assumption that studies independently provide single estimates of one or multiple outcomes. This setting can be too simplistic for some applications of meta-analysis. For example, some studies can report multiple estimates of the same effect size, either at different stages or for separate groups. Similarly, studies can exhibit nested levels of hierarchy, with higher grouping factors being represented for instance by geographical areas, administrative units, or study characteristics.[5] This configuration of repeated measures and/or hierarchical structures creates a

potentially complex pattern of dependence across effect sizes that must be accounted for. *Multilevel meta-analysis* has been proposed, in different forms, to extend the model defined in Section 4.1 by modeling the dependence through structured random effects.[5,9-11,13]

Analytic formulation

Here, we provide a general definition of multilevel random-effect meta-analysis that is applicable in various settings. The pattern of correlation is defined by aggregating effect sizes in *groups*, which can be defined both between and within studies. Nested grouping levels are used to express a hierarchical structure of random effects. For the sake of clarity, we start from a model for n effect sizes (units) aggregated in two nested grouping levels, written as

$$\begin{aligned} &y_{ijr} = \beta_0 + b_i + b_{ij} + \epsilon_{ijr}, i = 1, \dots, m, j = 1, \dots, m_i, r = 1, \dots, n_{ij}, \\ &b_i \sim \mathrm{N}\left(0, \tau_1^2\right), b_{ij} \sim \mathrm{N}\left(0, \tau_2^2\right), \epsilon_{ijr} \sim \mathrm{N}\left(0, s_{ijr}^2\right). \end{aligned} \tag{9}$$

Here, τ_1^2 is the variance of the random effects at the outer grouping level i, which includes m independent groups. In contrast, τ_2^2 is the random-effects variance within each of the m_i inner level groups nested in each outer-level group i. The units indexed by r represent measured effect sizes from the n_{ij} studies in the inner group j nested within the outer group i, each with known within-study variance s_{ijr}^2.

The model in Equation (9) can be extended to include an indefinite number L of grouping levels, with $\ell = 1, \dots, L$, and generally written in terms of the unified framework described in Section 2. First, we define $j = 1, \dots, g_i^\ell$ as the number of groups at level ℓ within each outer level i, with $g_i^1 = 1$ by definition. Each group includes $r = 1, \dots, n_{ij}^\ell$ units, with $\sum_{j\ell} n_{ij}^\ell = n_i$. The definition of the various elements in Equation (1) requires block-diagonal expansions and column binding consistent with repeated measures and grouping levels, respectively, with $\bigoplus_v a_v$ representing an operator that creates a block-diagonal matrix of elements a_v. We first define the design matrix for the fixed effects as $\mathbf{X}_i = \mathbf{1}_{n_i}$, and the known error structure at the outermost level as $\boldsymbol{\epsilon}_i = \bigoplus_{rj} \epsilon_{ijr}$. The random-effects part can be written by first defining design matrices for each group at various levels of random effects as $\mathbf{Z}_{ij}^\ell = \mathbf{1}_{n_{ij}^\ell}$, then expanding them at each level as $\mathbf{Z}_i^\ell = \bigoplus_j \mathbf{Z}_{ij}^\ell$, and finally binding them as $\mathbf{Z}_i = [\mathbf{Z}_i^1, \dots, \mathbf{Z}_i^L]$. Consistently, the between-group (co)variance matrix is defined as $\boldsymbol{\Psi}_i = \bigoplus_\ell \mathbf{I}_{g_i^\ell} \otimes \tau_\ell^2$. The model can be extended further to metaregression by replacing $\mathbf{X}_i$ with a $n_i \times p$ design matrix including p fixed-effects predictors. Similarly, q_ℓ random-effects predictors at any level ℓ can be included by replacing $\mathbf{Z}_{ij}^\ell$ with a $n_{ij}^\ell \times q_\ell$ design matrix, and τ_ℓ^2 with a random-effects (co)variance matrix $\boldsymbol{\Psi}^\ell$.

Illustrative examples

In a first example, we consider a meta-analysis of 56 studies that evaluate the effect of a modified school calendar on standardized reading achievement.[37] The studies were performed in 11 separate school districts, with at least three studies in each district, therefore providing a classic example of multilevel structure. Using the notation in Equation (9), in this example, the outer grouping level i are the school districts, which define $m = 11$ independent groups. Within each school district, a variable number of studies were performed, eg, four studies were performed in the first school district, ie, $m_1 = 4$. The study j, nested within the school district i, is the inner level in the multilevel structure with one single observation $r = 1$, with a single effect size in each inner group j nested within the outer group i. We fitted three models with different random-effects structures using a ML estimator: first, a traditional meta-analysis using the model in Section 4.1, with a single level of random effects assigned to each study; second, a single-level meta-analysis with random effects by district, therefore including repeated measures within each group; and third, a full two-level meta-analysis with nested random effects by study and district. The results, partly replicating the analysis of Konstantopoulos,[13] are reported in Table 2. The comparison makes clear the advantage of recognizing the multilevel structure of the data, with the pooled effect size β_0 increasing from 0.128 in the standard model to 0.184 in the two-level model. The latter, in addition, shows a better fit, as suggested by the lower AIC, and indicates the presence of heterogeneity at both district and study levels, with $\tau_1^2 = 0.058$ and $\tau_2^2 = 0.033$, respectively.

A second example of multilevel meta-analysis considers 20 randomized trials of thrombolytic therapy, which evaluated short-term mortality risks after a myocardial infarction.[38] The hypothesis is that the thrombolytic therapy reduces the risk and that the benefit is particularly substantial for very early treatment. Some of the trials report separate estimates of absolute risk change for sub-groups of treatment times, leading to a multilevel structure with 38 (potentially

TABLE 2 Example of multilevel meta-analysis of 56 studies that evaluate changes in standardized reading achievement after the implementation of a modified school calendar, with studies clustered within school districts. Previously reported by Konstantopoulos.[13]

Model	Grouping levels	Pooled estimate (Std error)	Random-effects variances		
		β_0	τ_1^2 (district)	τ_2^2 (study)	AIC
One-level (standard)	Study	0.128 (0.043)	-	0.087	37.292
One-level (repeated measures)	District	0.196 (0.086)	0.075	-	69.432
Two-level	Study within district	0.184 (0.080)	0.058	0.033	22.790

TABLE 3 Example of multilevel meta-analysis of 20 randomized trials of thrombolytic therapy for myocardial infarction, with multiple estimates of absolute risk change at different times of treatment. Previously reported by Thompson et al.[10]

Model	Fixed effects (Std error)		Random-effects variances	
	β_0 (intercept)	β_1 (treatment delay in hours)	τ_1^2 (trials)	τ_2^2 (times)
Standard meta-analysis	-0.02600 (0.00314)	-	-	0.00747
Two-level meta-analysis	-0.02600 (0.00314)	-	<0.00001	0.00747
Two-level meta-regression	-0.03494 (0.00421)	0.00161 (0.00049)	0.00216	0.00006

repeated) observations within 20 trials. We applied alternative models fitted by REML, partly reproducing the analysis by Thompson et al[10]: specifically, a standard meta-analysis that ignores clustering by trial, a two-level meta-analysis, and a two-level meta-regression that includes treatment delay as a metapredictor. The results, shown in Table 3, show that both standard and two-levels random-effects meta-analyses produce an estimate of absolute risk difference of -0.02600, suggesting a protective effects of thrombosis treatment, and that the second model indicates presence of heterogeneity within but not between the higher level of grouping represented by trials ($\tau_2^2 = 0.00747$ and $\tau_1^2 < 0.0000$). However, the inclusion of treatment delay in a meta-regression explains most of the variability at the inner level ($\tau_2^2 = 0.00006$), with a residual heterogeneity between trials of $\tau_1^2 = 0.00216$. These models (and other specifications) can be compared through AIC (when defined using the same fixed-effects structure) or using Wald tests for meta-predictors.

4.4 | Dose-response meta-analysis

Dose-response meta-analysis has been used to summarise linear and non-linear health associations across epidemiological studies.[39] The standard approach consists of a two-stage procedure. In the first stage, study-specific associations are determined using a set of parameters that represent estimates at different doses, usually retrieved from published data and relying on various methods to approximate their (co)variance matrix accounting for within-study correlations.[6] These estimates are then pooled in the second stage using standard meta-analytical models (see Section 4.1) for linear dose-response relationships or multivariate methods (see Section 4.2) for multiple parameters of functions representing nonlinear associations.[40,41] Recently, Crippa et al have proposed a one-stage approach for dose-response meta-analysis that provides important advantages and allows defining dose-response meta-analysis within the general framework proposed in Equation (1).[7]

Analytic formulation

The one-stage model for a linear dose-response random-effects meta-analysis can be written as follows:

$$\begin{aligned} y_{ij} &= \beta x_{ij} + b_i x_{ij} + \epsilon_{ij}, i = 1, \ldots, m, j = 1, \ldots, n_i, \\ b_i &\sim \mathrm{N}(0, \tau^2), [\epsilon_{i1}, \ldots, \epsilon_{in_i}] \sim \mathrm{N}(\mathbf{0}, \mathbf{S}_i). \end{aligned} \tag{10}$$

Here, for each study i, the n_i units y_{ij} represent the association measures (eg, log OR or log risk ratios) at different doses x_{ij}, commonly retrieved by using published estimates. The fixed-effects and random-effects parameters β and b_i represent the pooled linear dose-response association and its study specific deviations, respectively. Note the absence of an intercept for this model. Similarly to the standard model in Equation (7), τ^2 represents the variance of the random effects. Following Crippa et al,[7] this model can be written as a special version of the general framework in Section 2 by setting

TABLE 4 Example of dose-response meta-analysis of eight cohort studies on alcohol and colorectal cancer, with alternative model specifications. Previously partly reported by Orsini et al[6] and Crippa and Orsini[42]

Model	Within-study correlations	Fixed effects	Random effects	Degrees of freedom	AIC
L1	Zero correlation	Linear	Linear	2	-2.06
L2	Greenland and Longnecker	Linear	Linear	2	-6.13
NL1	Greenland and Longnecker	Non-linear	Non-linear	9	1.28
NL2	Greenland and Longnecker	Non-linear	Linear	4	-7.88
NL3	Greenland and Longnecker	Non-linear	None	3	-9.88

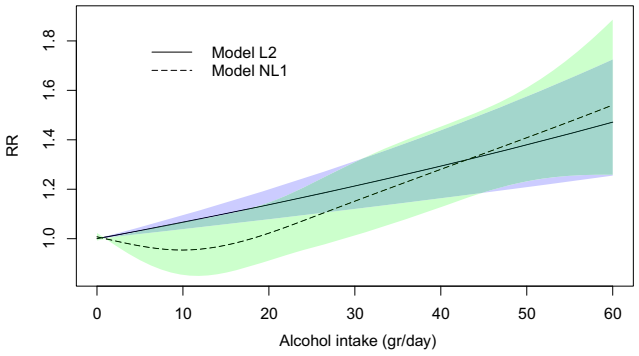


FIGURE 2 Dose-response relationships between alcohol intake and incidence relative rates of colorectal cancer assuming a linear and non-linear association (Models L2 and NL1 in Table 4), with 95% confidence intervals [Colour figure can be viewed at wileyonlinelibrary.com]

$\mathbf{X}_i = \mathbf{Z}_i = [x_{i1}, \ldots, x_{ij}]^{\mathrm{T}}$, $\boldsymbol{\beta} = \beta$, $\mathbf{b}_i = b_i$, and $\boldsymbol{\Psi} = \tau^2$. The within-study error structure is represented by a $n_i \times n_i$ matrix $\mathbf{S}_i$, usually approximated using alternative methods.[6,42] This model can be easily extended to the pooling of non-linear dose-responses by applying functions to transform x_{ij}, for instance quadratic terms or splines, thus obtaining a $n_i \times q$ matrices $\mathbf{X}_i$ and/or $\mathbf{Z}_i$, and a $q \times q$ matrix $\boldsymbol{\Psi}$.

Illustrative example

As an example of dose-response meta-analysis, we consider the data on eight cohort studies participating in the Pooling Project of Prospective Studies of Diet and Cancer.[43] Each study estimated the incidence relative rate (RR) of colorectal cancer in various categories of alcohol intake while controlling for a set of potential confounders, using non-drinkers as the reference. The categories were then converted in a dose by assigning the median value of individual consumptions, reporting log-RR estimates at multiple levels in a continuous scale. We fitted alternative models using a ML estimator, exploiting the flexibility of the extended framework in defining fixed effects and within and between-study correlations. Specifically, we specified linear and non-linear terms in both fixed and random parts, the latter by using natural cubic splines with internal knots at approximately the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of alcohol consumption. Within-study correlations were optionally reconstructed using the method of Berlin et al.[39]

The models are presented in Table 4, including number of parameters and AIC, and partly replicate and extend results presented by Orsini et al[6] and Crippa and Orsini.[6,42] Consistently, findings show that accounting for within-study correlation significantly improves the fit of the model, as indicated by the lower AIC of model L2 versus L1. The RR corresponding to 12 g/day of alcohol intake in the two models changes to 1.080 (95%CI: 1.047 to 1.115) from 1.048 (1.016 to 1.080), respectively. The inclusion of non-linear terms in both fixed and random parts does not improve the fit (NL1 versus L2). However, the simplification to linear random effects in NL2, allowed by the flexibility of the unified framework, indicates evidence of non-linearity (NL2 versus L2). The Cochran Q test for models NL1-NL2 suggests little evidence of heterogeneity (p-value = 0.25), as confirmed by the better fit of the fixed-effects non-linear dose-response meta-analysis in model NL3. The predicted RR for different doses obtained through linear and non-linear meta-analytic models L2 and NL1 are represented in Figure 2, with similar shapes to graphs previously presented.[6,42]

4.5 | Longitudinal meta-analysis

Another example of recent extensions of meta-analytical methods is for applications with studies where the same outcome is measured at several time points. *Longitudinal meta-analysis* have been proposed in this context to account for the intrinsic within-study and between-study correlations.[8,44] The common procedure is to apply meta-analytical methods for multivariate meta-analysis (see Section 4.2), treating effect sizes estimated at different times as separate outcomes,[8,45,46]

although this representation poses important constraints, as explained below. Alternative approaches define the longitudinal design as a special case of the multilevel setting described in Section 4.3, with repeated measures within each study.[44,47] This provides a way to formulate longitudinal meta-analysis within the unified framework in Equation (1), offering a general, flexible, and efficient modeling structure.

Analytic formulation

As mentioned above, traditional methods defines longitudinal meta-analysis as a multivariate model, where the effect sizes $\mathbf{y}_i = (y_{i,t_1}, \ldots, y_{i,t_k})^{\mathrm{T}}$ measured at k times in study i are treated as separate outcomes, and modeled as in Equation (8). However, this approach requires that the measurements are taken at common time points across studies, and while it may account for the sequential order, for instance by imposing autoregressive structures to the within and/or between-study correlations for evenly-spaced measures, it ignores most of the information provided by the longitudinal setting. A more flexible definition for a set of effect sizes measures at n_i times in study i is derived directly from LME models as

$$\begin{aligned} y_{it} &= (\alpha + a_i) + (\beta + b_i)t_{ij} + \epsilon_{it}, i = 1, \ldots, m, j = 1, \ldots, n_i, \\ [a_i, b_i] &\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}), [\epsilon_{it_1}, \ldots, \epsilon_{it_{n_i}}] \sim \mathrm{N}(\mathbf{0}, \mathbf{S}_i), \end{aligned} \tag{11}$$

with α, β, a_i, and b_i as fixed and random coefficients for intercepts and slopes. This formulation treats time as a continuous predictor that can be modeled through both fixed and random terms, and allows studies to report estimates at different times. The traditional multivariate approach can be defined as a special case by using indicators for a common set of time points. The model in Equation (11) can be written as the general framework in Equation (1) by setting $\mathbf{t}_i = [t_{i1}, \ldots, t_{in_i}]^{\mathrm{T}}$, and $\mathbf{X}_i = \mathbf{Z}_i = [\mathbf{1}_{n_i}, \mathbf{t}_i]$. $\boldsymbol{\Psi}$ and $\mathbf{S}_i$ define the random-effects and within-study error (co)variance matrices, respectively, optionally with specific structures, such as diagonal or (continuous) autoregressive of first order (AR_1). The model can allow nonlinear trends by specifying smooth functions of time (see Section 4.4), or include additional metapredictors, in both cases either as fixed or random effects by extending $\mathbf{X}_i$ and $\mathbf{Z}_i$, respectively.

Example of longitudinal meta-analysis

We illustrate an application of longitudinal meta-analysis using data on 17 randomized controlled trials comparing treatments of malignant gliomas. Each study measured the survival OR at 6, 12, 18, and 24 months since the start of the treatment.[48] Musekiwa et al[45] have previously analyzed the data using multivariate models fitted with REML, defining various longitudinal meta-analytic models with different specification of the within and between-study (co)variance structures. Here, we replicate and extend the results using the more flexible general framework and adopting alternative specifications.

The first set of results using multivariate models with indicators for the four time points are reported in Table 5. Consistently with the original analysis,[45] the first three options (Models 1-3) do not allow correlations in the within-study errors, while the other options (Models 4-6) assume a heterogeneous AR_1 structure with correlation fixed at 0.61. Different structures were used for the random-effects (co)variance, leading to different total degrees of freedom. The best-fitting option in terms of AIC is Model 4, with independent random effects and AR_1 within-study errors. The analysis can be extended by defining time as a continuous variable, specifying an additional Model 7 as in Equation (11). This random-slope model specifies a diagonal structure for intercept and (centered) time as random effects, and keeps the AR_1 within-study errors, using only four degrees of freedom. Models 7 and 4 were (re)fitted using ML, that allows comparison between different fixed-effects specifications (in this case, linear and through indicators, respectively). The results are graphically illustrated in Figure 3, showing the pooled OR along time after treatment. AIC indicates a better fit of Model 7 (101.4 versus 107.5, respectively), suggesting a linear trend and actually little evidence of changes in survival along time, with a p-value of 0.92 for the coefficient β of time (not shown). This example highlights the advantages offered by the modeling flexibility of the general modeling framework.

5 | SOFTWARE

The unified random-effects framework for meta-analysis and the frequentist inferential procedure described in the previous sections are implemented in the R package mixmeta. The main function of the program is `mixmeta()`, which uses

Model	(Co)variance structures		Degrees of freedom	AIC
	Within-study errors	Random effects		
Model 1	Diagonal	Diagonal	8	121.6
Model 2	Diagonal	Compound symmetry	5	117.0
Model 3	Diagonal	Heterogeneous AR_1	9	120.9
Model 4	Heterogeneous AR_1	Diagonal	8	107.5
Model 5	Heterogeneous AR_1	Heterogeneous AR_1	9	107.7
Model 6	Heterogeneous AR_1	Unstructured	14	117.3

TABLE 5 Example of longitudinal meta-analysis of 17 randomized controlled trials comparing treatments of malignant gliomas, reporting survival odds ratio at multiple times after treatment. Previously partly reported by Musekiwa et al[45]

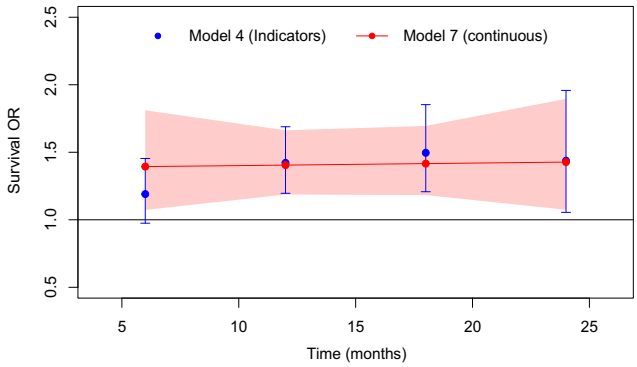


FIGURE 3 Survival odds ratio after start of the treatment of gliomas (Models 4 and 7 fitted using maximum likelihood), with 95% confidence intervals [Colour figure can be viewed at wileyonlinelibrary.com]

a simple syntax to fit a wide range of meta-analytical models. For instance, the following code:

```
mixmeta(cbind(y1 + y2)~ x1 + x2, S, data, random=list(~ z1 | g1, ~ 1 | g2),
method="reml")
```

performs a bivariate two-level meta-regression using a REML estimator. In this example, effect sizes for two outcomes `y1` and `y2`, with unit-level errors `S`, are modeled in terms of fixed-effect predictors `x1` and `x2`. Random effects are specified by intercept plus variable `z1` and intercept only, for nested grouping levels `g1` (outer) and `g2` (inner), respectively. The flexible formula syntax, similar to that applied in the R package nlme for LME models, allows the definition of all the various versions illustrated in Section 4. Other functions are available for hypothesis testing, predictions, model assessment, and simulations, among other regression tasks.

At the time of writing, the package implements a hybrid estimation procedure, with few runs of a (R)IGLS algorithm followed by quasi-Newton iterations. The former is robust to initial values and quickly moves close to the ML, while the latter provides a fast convergence within this region. As mentioned in Section 3, these algorithms adopt a profiled approach, where the likelihood functions in Equations (2) are defined in terms of random-effect parameters $\boldsymbol{\xi}$ only, with a parameterization that ensures positive definiteness and allows different structures for any of the (co)variance matrices $\boldsymbol{\Psi}^{(\ell)}$. Computationally, the estimation algorithms exploit the block-diagonal form of the design and (co)variance matrices defined in Equations (1)-6, which is particularly convenient in the presence of a high number of studies or outer groups. A QR decomposition is applied internally in the GLS routine, providing numerical stability even in not well-conditioned least squares problems.

6 | SIMULATIONS

We performed a simulation study to explore the validity and inferential properties of the software implementation of the unified framework. We considered a complex case represented by a multivariate multilevel meta-analysis, combining the applications described in Sections 4.2 and 4.3 within the general model defined in Equation (1). Specifically, we simulated $k = 3$ outcomes and $L = 2$ grouping levels, with m groups at the (outer) level 1, each including $g_i^2 = 10$ groups at the (inner) level 2. All the fixed-effects $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]^{\mathrm{T}}$, representing the three pooled intercepts, were simulated as 0. We assumed a compound-symmetry structure for both the 3×3 random-effects (co)variance matrices $\boldsymbol{\Psi}^1$ and $\boldsymbol{\Psi}^2$, and a heterogeneous compound-symmetry structure for the residual error matrix $\mathbf{S}_i$. Random-effects variances τ_1^2 and τ_2^2 were set to 1, while the residual error variances s_{ijr}^2 were sampled from a uniform distribution with range $[0.1, 2]$. Various

TABLE 6 Simulation study: multivariate multilevel meta-analysis with $k = 3$ outcomes and $L = 2$ grouping levels, with m groups at the (outer) level 1, each including $g_i^2 = 10$ groups at the (inner) level 2. Eight simulation scenarios are defined by the number of outer-level groups m and correlations ρ_{b1} and ρ_{b2} for each level of random effects

Parameters			β_1			τ_1	τ_2	ρ_{b1}	ρ_{b2}
m	ρ_w	ρ_b	Bias	RMSE	Coverage	Bias	Bias	Bias	Bias
10	0.00	0.00	0.000	0.349	0.940	-0.006	0.004	0.004	-0.004
50	0.00	0.00	-0.001	0.157	0.948	-0.001	0.001	0.002	-0.001
10	0.80	0.00	-0.005	0.351	0.930	-0.005	0.005	-0.045	-0.003
50	0.80	0.00	0.001	0.154	0.951	-0.000	0.001	-0.007	0.000
10	0.00	0.80	0.003	0.357	0.934	-0.007	0.004	-0.019	-0.030
50	0.00	0.80	-0.004	0.155	0.949	0.001	0.000	-0.002	-0.007
10	0.80	0.80	-0.006	0.347	0.928	-0.006	-0.001	-0.058	-0.015
50	0.80	0.80	0.001	0.153	0.951	0.000	0.000	-0.008	-0.003

simulation scenarios are represented by combinations of number of outer-level groups m (10 or 50), correlation ρ_{b1} and ρ_{b2} for each level of random effects (0 or 0.8), and residual correlation ρ_w (0 or 0.8). For each combination, we simulated 10 000 datasets using the function `mixmetaSim()`, and fitted the general model with `mixmeta()` assuming the correct random-effects (co)variance structures.

Results are reported in Table 6, showing the bias, root mean square error (RMSE), and coverage for the (first) fixed-effects coefficient, and the bias for the four random-effects parameters. Simulations indicate a negligible amount of bias for both the fixed and random-effects parameters in all scenarios. As expected, the RMSE decreases when increasing number of outer-level units. The coverage is slightly below the nominal value, especially for scenarios with lower number ($m = 10$) of outer-level groups. Inferential properties do not seem affected by the presence of within or between units correlation.

7 | DISCUSSION

In this contribution, we have presented an extended mixed-effects framework that provides a common modeling and inferential setting for meta-analysis. It includes traditional applications but also non-standard extensions for which common meta-analytical methods are not appropriate. The unified approach proposed here generally characterizes these extensions as patterns of dependence between effect sizes, modeled through fixed and random effects defined by meta-predictors and grouping structures. This modeling approach allows a flexible specification of variety of meta-analytical models and facilitates the design and implementation of non-standard pooling studies.

The LME structure adopted in the definition of the general model in Equation (1) provides substantial modeling flexibility, through which important constraints in design and modeling aspects can be relaxed. For instance, analyses of longitudinal data are traditionally performed using models for multivariate meta-analysis that consider repeated measurements from the same study as multiple outcomes.[6,8] However, this approach requires a limited set of measurements to be taken at the same doses/times across studies and prevents their analysis as continuous variables. In the examples in Sections 4.4 and 4.5, we showed how more flexible models can be defined within our general framework, allowing studies to provide an indefinite number of measurements taken at any point and the modeling of continuously dose-response shapes and trends through linear or smooth functions. Similarly, the flexible definition of multilevel models in Section 4.3 allows the specification of complex hierarchical structures and the inclusion of random-effects meta-predictors.

The extended framework presented in this paper is well placed for a two-stage analytical setting, where the estimated effect sizes are derived from published studies or previously obtained from separate study-specific analyses. One-stage formulations have been proposed for individual patient data meta-analysis, when data from the original studies are available and can be directly modeled.[49] However, in many applications the one-stage approach provides little advantages, and two-stage procedures offer a valid, computationally stable, and efficient alternative.[9,50] In addition, the flexible framework proposed here allows extensions of the two-stage design to address specific limitations, for example with the pooling of multiple study-specific parameters of main and interaction terms to evaluates effect modification from participant-level variables. The two-stage approach relies on the assumption of normal distribution of estimated effect sizes and random effects, thus requiring approximations, in particular for outcomes measured in a binary scale. One-stage methods based on generalized linear mixed models (GLMMs) have been developed in this setting, including versions with alternative distributional assumptions.[11,51,52] While these have theoretical and inferential advantages, they present considerable computational problems, and simulations show improvements only in the presence of small and sparse data.[53,54] An additional requirement of the two-stage procedure, when applied in multi-parameter meta-analyses, is the knowledge of the within-study covariances. Methods for estimating them from published data were developed in

multivariate meta-analysis,[5,31,55] and in dose-response meta-analysis, and can be applied in this general model.[56-58] In addition, interestingly, the unknown correlations can alternatively be merged in a marginal random-effects structure that includes within and between-study dependencies.[16,59] One of the advantage of using an LME formulation in the extended framework is that it does not require balanced data where the full set of effect sizes is measured (or reported) for each study. The extended framework can in fact deal with unbalanced data and more generally deal with the presence of missing effect sizes. However, the analysis requires the assumption of missing at random (MAR) to provide unbiased estimates.[60]

The methodology is implemented in the freely available and fully documented R package mixmeta, which complements standard software for meta-analysis and additional tools for specific extensions. For instance, some analysts have proposed the use of general LME programs for fitting complex meta-analytical models, such as the procedure `PROC MIXED` in SAS,[17] the program GLLAMM in Stata,[15] MLwiN,[11] or the package nlme in R.[61] However, the use of general LME software requires advanced knowledge of statistical and computational aspects and can be difficult for more applied users. Dedicated routines are available for specific meta-analytical extensions, such as the Stata command and R package `mvmeta` for multivariate meta-analysis,[3,35] or the R package dosresmeta and drmeta Stata module for dose-response meta-analysis,[42,62] while metafor in R can offer a set of general tools for standard models and various extensions.[63] Our implementation in mixmeta offers a flexible platform where the full range of models presented in Section 4, and their combinations, can be defined through a simple syntax, fitted using an efficient computational structure, and estimated following a common underlying statistical theory. This software can complement existing packages and modules for the specific meta-analytic extensions presented in Section 4.

The simulation study in Section 6 demonstrates the validity and good performance of the modeling framework and software, even in a relatively complex application represented by a trivariate multilevel meta-analysis. However, some limitations of the inferential approaches described in Section 3 must be acknowledged. The Wald test procedure for fixed effects is based on asymptotic distributional approximations, and it ignores the uncertainty related to the estimation of the random-effects components. This explains the small undercoverage of confidence intervals in Table 6, which however can be non-negligible in small-sample studies. Similarly, hypotheses on random effects are evaluated through LR tests and AIC/BIC, but these can have poor performances and problems with boundary conditions. Solutions can be found in the LME models literature, such as the use of t or F distributions,[18] adjustments for standard errors and degrees of freedom,[64] and use of mixture distributions.[18,23] Some of these have also been defined for meta-analytical models,[27,29,65-68] but still need to be fully developed for this extended framework. Alternative methods can also be developed in a Bayesian framework, which offers advantages in accounting for various sources of uncertainty, although requiring appropriate parameterizations and priors specification.[10,27,69-71]

There is an increasing interest in developing meta-analysis for applications in more complex pooling studies, beyond the now established extensions described in Section 4.[4,72] Emerging areas include investigations that apply two-stage designs for the analysis of large datasets, where either the complexity of the first-stage regression or the computational demand prevent the definition of a one-stage model, and the partition of the analysis in two steps provides a feasible and efficient approach.[3,50,73] However, the limitations of traditional meta-analytical methods, requiring the estimation of single independent parameters from each subset, poses important constraints in this setting. In contrast, the model in Equation (1) offers flexibility in the definition of the two-stage analysis, allowing for instance repeated measurements in time or subgroups, hierarchies, and spatial or temporal clustering, and complex multiparameter effect estimates. The definition of a unified framework for meta-analysis, complemented with a full software implementation, provides researchers with a flexible tool for defining and applying flexible meta-analytical models in a variety of pooling problems.

ACKNOWLEDGEMENTS

This research was supported by funding from the Medical Research Council UK (grants MR/M022625/1 and MR/R013349/1).

DATA AVAILABILITY STATEMENT

R code and data for replicating examples and simulation results in Sections 4 and 6 are available at the personal website and GitHub page of the last author.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

FINANCIAL DISCLOSURE

None reported.

ORCID

Francesco Sera https://orcid.org/0000-0002-8890-6848
Benedict Armstrong https://orcid.org/0000-0003-4407-0409
Marta Blangiardo https://orcid.org/0000-0002-1621-704X
Antonio Gasparrini https://orcid.org/0000-0002-2271-3568

REFERENCES

1. Borenstein M, Hedges LV, Higgins J, Rothstein HR. *Introduction to Meta-Analysis*. Hoboken, NJ: John Wiley & Sons; 2009.
2. Jackson D, Riley R, White IR. Multivariate meta-analysis: potential and promise. *Statist Med*. 2011;30(20):2481-2498.
3. Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statist Med*. 2012;31(29):3821-3839.
4. Riley RD, Jackson D, Salanti G, et al. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ*. 2017;358:j3932.
5. Stevens JR, Taylor AM. Hierarchical dependence in meta-analysis. *J Educ Behav Stat*. 2009;34(1):46-73.
6. Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol*. 2011;175(1):66-73.
7. Crippa A, Discacciati A, Bottai M, Spiegelman D, Orsini N. One-stage dose–response meta-analysis for aggregated data. *Stat Methods Med Res*. 2019;28(5):1579-1596.
8. Ishak K, Platt RW, Joseph L, Hanley JA, Caro JJ. Meta-analysis of longitudinal studies. *Clinical Trials*. 2007;4(5):525-539.
9. Goldstein H, Yang M, Omar R, Turner R, Thompson S. Meta-analysis using multilevel models with an application to the study of class size effects. *J R Stat Soc Ser C Appl Stat*. 2000;49(3):399-412.
10. Thompson SG, Turner RM, Warn DE. Multilevel models for meta-analysis, and their application to absolute risk differences. *Stat Methods Med Res*. 2001;10(6):375-392.
11. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statist Med*. 2000;19(24):3417-3432.
12. Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Statist Med*. 1998;17(22):2537-2550.
13. Konstantopoulos S. Fixed effects and variance components estimation in three-level meta-analysis. *Res Synth Methods*. 2011;2(1):61-76.
14. Stram DO. Meta-analysis of published data using a linear mixed-effects model. *Biometrics*. 1996;52(2):536-544.
15. Bagos PG. Meta-analysis in Stata using gllamm. *Res Synth Methods*. 2015;6(4):310-332.
16. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Three-level meta-analysis of dependent effect sizes. *Behav Res Methods*. 2013;45(2):576-594.
17. Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statist Med*. 2002;21(4):589-624.
18. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-Plus*. New York, NY: Springer Verlag; 2000.
19. Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. *J Am Stat Assoc*. 1977;72(358):320-338.
20. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. New York, NY: Springer; 1997.
21. Goldstein H. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*. 1986;73(1):43-56.
22. Goldstein H. Restricted unbiased iterative generalized least-squares estimation. *Biometrika*. 1989;76(3):622-623.
23. Goldstein H. *Multilevel Statistical Models*. Vol 922. Hoboken, NJ: John Wiley & Sons; 2011.
24. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177-188.
25. Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statist Med*. 2002;21(11):1539-1558.
26. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
27. Higgins J, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137-159.
28. Rice K, Higgins J, Lumley T. A re-evaluation of fixed effect (s) meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2018;181(1):205-227.
29. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for meta-analysis. *Statist Med*. 1995;14(4):395-411.
30. Colditz GA, Brewer TF, Berkey CS, et al. Efficacy of BCG vaccine in the prevention of tuberculosis: meta-analysis of the published literature. *JAMA*. 1994;271(9):698-702.
31. Olkin I, Gleser L. Stochastically dependent effect sizes. In: Cooper H, Hedges LV, Valentine JC, eds. *The Handbook of Research Synthesis and Meta-Analysis*. New York, NY: Russell Sage Foundation; 2009:357-376.
32. Ma X, Nie L, Cole SR, Chu H. Statistical methods for multivariate meta-analysis of diagnostic tests: an overview and tutorial. *Stat Methods Med Res*. 2016;25(4):1596-1619.

33. White IR, Barrett JK, Jackson D, Higgins J. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-125.
34. Fiore MC, Bailey WC, Cohen SJ, et al. Smoking cessation: clinical practice guideline no. 18. AHCPR Publication No. 96-0692, Agency for Health Care Policy and Research, U.S. Department of Health and Human Services; 1996.
35. White IR. Multivariate random-effects meta-regression: updates to mvmeta. *Stata J*. 2011;11(2):255-270.
36. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods*. 2012;3(2):98-110.
37. Cooper H, Valentine JC, Charlton K, Melson A. The effects of modified school calendars on student achievement and on school and community attitudes. *Rev Educ Res*. 2003;73(1):1-52.
38. Boersma E, Maas ACP, Deckers JW, Simoons ML. Early thrombolytic treatment in acute myocardial infarction: reappraisal of the golden hour. *Lancet*. 1996;348(9030):771-775.
39. Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology*. 1993;4(3):218-228.
40. Liu Q, Cook NR, Bergström A, Hsieh CC. A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose–response data. *Comput Stat Data Anal*. 2009;53(12):4157-4167.
41. Rota M, Bellocco R, Scotti L, et al. Random-effects meta-regression models for studying nonlinear dose–response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statist Med*. 2010;29(26):2679-2687.
42. Crippa A, Orsini N. Multivariate dose-response meta-analysis: the dosresmeta R package. *J Stat Softw*. 2016;72(1):1-15.
43. Cho E, Smith-Warner SA, Ritz J, et al. Alcohol intake and colorectal cancer: a pooled analysis of 8 cohort studies. *Ann Intern Med*. 2004;140(8):603-613.
44. Peters JL, Mengersen KL. Meta-analysis of repeated measures study designs. *J Eval Clin Pract*. 2008;14(5):941-950.
45. Musekiwa A, Manda SOM, Mwambi HG, Chen DG. Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *PloS One*. 2016;11(10):e0164898.
46. Trikalinos TA, Olkin I. Meta-analysis of effect sizes reported at multiple time points: a multivariate approach. *Clinical Trials*. 2012;9(5):610-620.
47. Ahn JE, French JL. Longitudinal aggregate data model-based meta-analysis with NONMEM: approaches to handling within treatment arm correlation. *J Pharmacokinet Pharmacodyn*. 2010;37(2):179-201.
48. Fine HA, Dear KBG, Loeffler JS, Mc Black PL, Canellos GP. Meta-analysis of radiation therapy with and without adjuvant chemotherapy for malignant gliomas in adults. *Cancer*. 1993;71(8):2585-2597.
49. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*. 2010;340:c221.
50. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Statist Med*. 2017;36(5):855-875.
51. Platt RW, Leroux BG, Breslow N. Generalized linear mixed models for meta-analysis. *Statist Med*. 1999;18(6):643-654.
52. Stijnen T, Hamza TH, Özdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statist Med*. 2010;29(29):3046-3067.
53. Jackson D, Law M, Stijnen T, Viechtbauer W, White IR. A comparison of seven random-effects models for meta-analyses that estimate the summary odds ratio. *Statist Med*. 2018;37(7):1059-1085.
54. Bakbergenuly I, Kulinskaya E. Meta-analysis of binary outcomes via generalized linear mixed models: a simulation study. *BMC Med Res Methodol*. 2018;18(1):70.
55. Wei Y, Higgins J. Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statist Med*. 2013;32(7):1191-1205.
56. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol*. 1992;135(11):1301-1309.
57. Orsini N, Bellocco R, Greenland S. Generalized least squares for trend estimation of summarized dose–response data. *Stata J*. 2006;6(1):40-57.
58. Hamling J, Lee P, Weitkunat R, Ambuhl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Statist Med*. 2008;27(7):954-970.
59. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. Meta-analysis of multiple outcomes: a multilevel approach. *Behav Res Methods*. 2015;47(4):1274-1294.
60. Molenberghs G, Kenward M. *Missing Data in Clinical Studies*. Vol 61. Hoboken, NJ: John Wiley & Sons; 2007.
61. Heisterkamp SH, van Willigen E, Diderichsen PM, Maringwa J. Update of the NLME package to allow a fixed standard deviation of the residual error. *R J*. 2017;9(1):239-251.
62. Orsini N. DRMETA: Stata module for dose-response meta-analysis. Statistical Software Components, Boston College Department of Economics; 2018.
63. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36(3):1-48.
64. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. 1997:983-997.
65. Kalaian HA, Raudenbush SW. A multivariate mixed linear model for meta-analysis. *Psychol Methods*. 1996;1(3):227.
66. Riley RD, Abrams KR, Lambert PC, Sutton AJ, Thompson JR. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Statist Med*. 2007;26(1):78-97.
67. Follmann DA, Proschan MA. Valid inference in random effects meta-analysis. *Biometrics*. 1999;55(3):732-737.

68. Morris TP, Fisher DJ, Kenward MG, Carpenter JR. Meta-analysis of Gaussian individual patient data: two-stage or not two-stage? *Statist Med*. 2018;37(9):1419-1438.
69. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res*. 2001;10(4):277-303.
70. Nam IS, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Statist Med*. 2003;22(14):2309-2333.
71. Wei Y, Higgins J. Bayesian multivariate meta-analysis with multiple outcomes. *Statist Med*. 2013;32(17):2911-2934.
72. Sutton AJ, Higgins J. Recent developments in meta-analysis. *Statist Med*. 2008;27(5):625-650.
73. Rhodes KM, Turner RM, Payne RA, White IR. Computationally efficient methods for fitting mixed models to electronic health records data. *Statist Med*. 2018;37(29):4557-4570.

SUPPORTING INFORMATION

The following supporting information is available as part of the online article: R code and data for replicating examples and simulation results in Sections 4 and 6, with an updated version available at the personal website and GitHub page of the last author.

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Sera F, Armstrong B, Blangiardo M, Gasparrini A. An extended mixed-effects framework for meta-analysis. *Statistics in Medicine*. 2019;38:5429–5444. https://doi.org/10.1002/sim.8362

The following supporting information is available as part of the online article: R code and data for replicating examples and simulation results in Sections 4 and 6.

```
##############################################################################
#######
# Updated R code for the analysis in:
#
#   "An extended mixed-effects framework for meta-analysis"
#   Francesco Sera, Ben Armstrong, Marta Blangiardo & Antonio Gasparrini
#   Statistics in Medicine - 2019
#   http://www.ag-myresearch.com/2019_sera_statmed.html
#
# Update: 22 Aug 2019
# * an up-to-date version of this code is available at:
#   https://github.com/gasparrini/2019_sera_StatMed_Rcode
##############################################################################
#######

##############################################################################
#######
# STANDARD META-ANALYSIS (SECTION 4.1)
##############################################################################
#######

# LOAD THE PACKAGES
library(mixmeta); library(Epi)

# STANDARD RANDOM-EFFECTS META-ANALYSIS WITH MAXIMUM LIKELIHOOD
uniran <- mixmeta(logor, logorvar, data=bcg, method="ml")

# RESULTS
print(summary(uniran), digits=3, report="var")
print(ci.exp(uniran), digits=3)

# EXTRACT LOG-OR AND CALCULATE 95% CONFIDENCE INTERVALS
pred <- with(bcg, cbind(logor, logor-1.96*sqrt(logorvar),
  logor+1.96*sqrt(logorvar)))

# BEST-LINEAR UNBIASED PREDICTIONS, WITH PREDICTION INTERVALS
blup <- blup(uniran, pi=TRUE)

# FOREST PLOT
plot(pred[,1], rev(bcg$trial)+0.2, xlim=c(-3,3), ylim=c(0,14), pch=18,
  axes=FALSE, xlab="Log odds ratio", ylab="Trial", main="Forest plot")
axis(1)
axis(2, at=bcg$trial, labels=rev(bcg$trial), lty=0, las=1)
abline(v=coef(uniran))
segments(pred[,2], rev(bcg$trial)+0.2, pred[,3], rev(bcg$trial)+0.2,
lty=5)
points(blup[,1], rev(bcg$trial)-0.2, pch=19)
segments(blup[,2], rev(bcg$trial)-0.2, blup[,3], rev(bcg$trial)-0.2)
legend("right", c("Original","BLUPs"),lty=c(5,1), pch=c(18,19), lwd=1,
bty="n")

# META-REGRESSION MODEL TO EVALUATE THE EFFECT OF LATITUDE
uniranlat <- update(uniran, .~. + ablat)

# LIKELIHOOD RATIO TEST (ALLOWED WITH ML)
drop1(uniranlat, test="Chisq")

# RESULTS
print(summary(uniranlat), digits=3, report="var")
```

```
# SEE help(bcg) FOR FURTHER INFO

#########################################################################
#######
# MULTIVARIATE (NETWORK) META-ANALYSIS (SECTION 4.2, TABLE 1)
#########################################################################
#######

# LOAD THE PACKAGE
library(mixmeta)

# INSPECT THE DATA
head(smoking)
names(smoking)

# CONSISTENCY MODEL, UNSTRUCTURED BETWEEN-STUDY (CO)VARIANCE
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod1 <- mixmeta(y, S)
summary(mod1)

# CONSISTENCY MODEL, STRUCTURED BETWEEN-STUDY (CO)VARIANCE (PROPORTIONAL)
mod2 <- mixmeta(y, S, bscov="prop", control=list(Psifix=diag(3)+1))
summary(mod2)

# TRANSFORM IN LONG FORMAT, WITH S AS LIST (EXCLUDING MISSING)
long <- na.omit(reshape(smoking[,c(1,2,11:13)], varying=list(3:5),
idvar="study",
  v.names="y", timevar="outcome", times=colnames(y), direction="long"))
Slist <- lapply(lapply(seq(nrow(S)), function(i) xpndMat(S[i,])),
function(x)
  x[!is.na(diag(x)), !is.na(diag(x)), drop=F])

# THE MODELS ABOVE CAN BE REPLICATED IN THE LONG FORMAT
mod2b <- mixmeta(y ~ 0 + factor(outcome), random= ~ 0 +
factor(outcome)|study,
  data=long, bscov="prop", control=list(addS=Slist, Psifix=diag(3)+1))
summary(mod2b)

# DEFINE AND ADD INDICATORS FOR OUTCOME AND DESIGN
dummy <- cbind(model.matrix(~0+outcome, long), model.matrix(~0+design,
long))
colnames(dummy) <- c(levels(factor(long$outcome)), levels(long$design))
long <- cbind(long, data.frame(dummy))

# INCONSISTENCY MODEL (SPECIAL PARAMETERIZATION OF OUTCOME-BY-DESIGN
INTERACTION)
formula <- y ~ 0 + yB + yC + yC:acd + yC:bc + yC:bcd + yD + yD:acd +
yD:bcd +
  yD:bd + yD:cd
mod3 <- update(mod2b, formula=formula)
summary(mod3)

# WALD TEST
fwald <- function(model,var) {
  ind <- grep(var,names(coef(model)))
  coef <- coef(model)[ind]
  vcov <- vcov(model)[ind,ind]
```

```
  waldstat <- coef%*%solve(vcov)%*%coef
  df <- length(coef)
  return(1-pchisq(waldstat,df))
}
fwald(mod3, c(":"))

# SEE help(smoking) FOR FURTHER INFO

#########################################################################
#######
# MULTILEVEL META-ANALYSIS (SECTION 4.3 - EXAMPLE 1, TABLE 2)
#########################################################################
#######

# LOAD THE PACKAGE
library(mixmeta)

# STUDY AS SINGLE LEVEL: STANDARD META-ANALYSIS
mod1 <- mixmeta(effect, var, random= ~ 1|study, data=school, method="ml")
print(summary(mod1), digits=3, report="var")

# DISTRICT AS SINGLE LEVEL: META-ANAYSIS WITH REPEATED MEASURES
mod2 <- mixmeta(effect, var, random= ~ 1|district, data=school,
method="ml")
print(summary(mod2), digits=3, report="var")

# NESTED LEVELS OF STUDY AND DISTRICT: TWO-LEVEL META-ANALYSIS
mod3 <- mixmeta(effect, var, random= ~ 1|district/study, data=school,
  method="ml")
print(summary(mod3), digits=3, report="var")

# COMPARISON
AIC(mod1, mod2, mod3)

# SEE help(school) FOR FURTHER EXAMPLES


#########################################################################
#######
# MULTILEVEL META-ANALYSIS (SECTION 4.3 - EXAMPLE 2, TABLE 3)
#########################################################################
#######

# STANDARD META-ANALYSIS: IGNORING CLUSTERING OF TRIALS
subtrial <- seq(nrow(thrombolytic))
mod1 <- mixmeta(absrisk, var, random= ~ 1|subtrial, data=thrombolytic)
print(summary(mod1), digits=5)

# STANDARD META-REGRESSION
mod2 <- mixmeta(absrisk~time2treat, var, random= ~ 1|subtrial,
  data=thrombolytic)
print(summary(mod2), digits=5)

# TWO-LEVEL META-ANALYSIS
mod3 <- mixmeta(absrisk, var, random= ~ 1|trial/subtrial,
data=thrombolytic)
print(summary(mod3), digits=5)

# TWO-LEVEL META-REGRESSION
```

```
mod4 <- mixmeta(absrisk~time2treat, var, random= ~ 1|trial/subtrial,
  data=thrombolytic)
print(summary(mod4), digits=5)

# SEE help(thrombolytic) FOR FURTHER INFO

#########################################################################
#######
# DOSE-RESPONSE META-ANALYSIS (SECTION 4.4, TABLE 4 AND FIGURE 2)
#########################################################################
#######

# LOAD THE PACKAGES
library(mixmeta); library(dosresmeta); library(splines)

# INSPECT THE DATA
head(alcohol)

# COMPUTE THE WITHIN-STUDY CORRELATIONS EXCLUDING THE REFERENCE
addS <- lapply(split(alcohol, alcohol$id), function(x)
  covar.logrr(y=logrr, v=se^2, cases=cases, n=peryears, type=type,
data=x))
sub <- subset(alcohol, !is.na(se))

# LINEAR FIXED AND RANDOM EFFECTS NOT ACCOUNTING FOR WITHIN-STUDY
CORRELATIONS
modL1 <- mixmeta(logrr ~ 0 + dose, S=se^2, random= ~ 0 + dose|id,
data=sub,
  method="ml")
summary(modL1)

# LINEAR FIXED AND RANDOM EFFECTS ACCOUNTING FOR WITHIN-STUDY
CORRELATIONS
modL2 <- mixmeta(logrr ~ 0 + dose, random= ~ 0 + dose|id, data=sub,
method="ml",
  control=list(addSlist=addS))
summary(modL2)

# NON-LINEAR FIXED AND RANDOM EFFECTS
modNL1 <- mixmeta(logrr ~ 0 + ns(dose, knots=c(10,25)), data=sub,
  random= ~ 0 + ns(dose, knots=c(10,25))|id, method="ml",
  control=list(addSlist=addS))
summary(modNL1)

# SIMPLIFY THE MODEL BY ALLOWING NON-LINEARITY ONLY IN FIXED EFFECTS
modNL2 <- update(modNL1, random= ~ 0 + dose|id)
summary(modNL2)

# FIXED-EFFECTS MODEL (TRICK: random TO DEFINE THE GROUPING, THEN FIX IT
TO 0)
modNL3 <- mixmeta(logrr ~ 0 + ns(dose, knots=c(10,25)), random= ~ 1|id,
  data=sub, method="ml",bscov="fixed", control=list(addSlist=addS,
Psifix=0))
summary(modNL3)

# COMPARE WITH AIC
AIC(modL1, modL2, modNL1, modNL2, modNL3)

# PREDICT THE RR FOR 12g/day FOM TWO MODELS
```

```
exp(predict(modL1, newdata=data.frame(dose=12), ci=TRUE))
exp(predict(modL2, newdata=data.frame(dose=12), ci=TRUE))

# PREDICT THE RR ALONG THE DOSE RANGE
predlin <- exp(predict(modL2, newdata=data.frame(dose=0:60), ci=TRUE))
prednonlin <- exp(predict(modNL1, newdata=data.frame(dose=0:60),
ci=TRUE))

# PLOT
par(mar=c(5,4,1,0.5))
col1 <- do.call(rgb,c(as.list(col2rgb("blue")/255), list(0.2)))
col2 <- do.call(rgb,c(as.list(col2rgb("green")/255), list(0.2)))
plot(0:60,predlin[,1], type="l", ylim=c(0.85,1.9), ylab="RR",
  xlab="Alcohol intake (gr/day)")
polygon(c(0:60,60:0), c(predlin[,2],rev(predlin[,3])), col=col1
,border=NA)
lines(0:60,prednonlin[,1], lty=5)
polygon(c(0:60,60:0), c(prednonlin[,2], rev(prednonlin[,3])), col=col2
,border=NA)
legend("topleft", c("Model L2","Model NL1"), lty= c(1,5), bty="n",
inset=0.1)

# SEE help(alcohol) FOR FURTHER INFO

#########################################################################
#######
# LONGITUDINAL META-ANALYSIS (SECTION 4.5, TABLE 5 AND FIGURE 3)
#########################################################################
#######

# LOAD THE PACKAGE
library(mixmeta)

data(gliomas)
# THE gliomas DATASET IS ARRANGED IN A LONG FORMAT
head(gliomas)

# INDEPENDENT RANDOM EFFECTS, NO WITHIN-STUDY CORRELATION (MODEL 1)
mod1 <- mixmeta(logOR~0+factor(time), S=logORvar,
random=~0+factor(time)|study,
  bscov="diag", data=gliomas)
print(summary(mod1), digits=3, report="var")

# COMPOUND-SYMMETRY RANDOM EFFECTS, NO WITHIN-STUDY CORRELATION (MODEL 2)
# NB: THIS REQUIRES A TWO-LEVEL MODEL WITH THE INNER-LEVEL VARIANCE FIXED
TO 0
unit <- factor(seq(nrow(gliomas)))
mod2 <- mixmeta(logOR~0+factor(time), S=logORvar, random=~1|study/unit,
  bscov=c("unstr","fixed"), data=gliomas,
control=list(Psifix=list(unit=0)))
print(summary(mod2), digits=3, report="var")

# HETEROGENEOUS AR1 RANDOM EFFECTS, NO WITHIN-STUDY CORRELATION (MODEL 3)
mod3 <- update(mod1, bscov="ar1")
print(summary(mod3), digits=3, report="var")

# BUILD THE HETEROGENEOUS AR1 WITHIN-STUDY ERRORS (CORRELATION AT 0.61)
cormat <- 0.61^abs(col(matrix(1,4,4)) - row(col(matrix(1,4,4))))
```

```
addS <- lapply(split(sqrt(gliomas$logORvar), gliomas$study), inputcov,
cormat)
addS <- lapply(addS, function(x) x[apply(!is.na(x),1,any),
  apply(!is.na(x),2,any)])

# INDEPENDENT RANDOM EFFECTS, HAR1 WITHIN-STUDY CORRELATION (MODEL 4)
mod4 <- mixmeta(logOR~0+factor(time), random=~0+factor(time)|study,
  bscov="diag", data=gliomas, control=list(addSlist=addS))
print(summary(mod4), digits=3, report="var")

# HAR1 RANDOM EFFECTS, HAR1 WITHIN-STUDY CORRELATION (MODEL 5)
mod5 <- update(mod4, bscov="ar1")
print(summary(mod5), digits=3, report="var")

# UNSTRUCTURED RANDOM EFFECTS, HAR1 WITHIN-STUDY CORRELATION (MODEL 6)
mod6 <- update(mod4, bscov="unstr")
print(summary(mod6), digits=3, report="var")

# COMPARE THE FIT WITH AIC
AIC(mod1, mod2, mod3, mod4, mod5, mod6)

# RE-RUN BEST FITTING MODEL WITH ML (ALLOWS TESTING OF FIXED EFFECTS)
mod4ml <- update(mod4, method="ml")
print(summary(mod4ml), digits=3, report="var")

# RANDOM-SLOPE MODEL WITH TIME AS CONTINUOUS AND CENTERED
mod7ml <- mixmeta(logOR~time, random=~I(time-15)|study, bscov="diag",
  method="ml", data=gliomas, control=list(addSlist=addS, maxiter=200))
print(summary(mod7ml), digits=3, report="var")

# PREDICT
times <- unique(gliomas$time)
predmod4ml <- exp(predict(mod4ml, data.frame(time=times), ci=TRUE))
predmod7ml <- exp(predict(mod7ml, data.frame(time=times), ci=TRUE))

# PLOT
par(mar=c(5,4,1,0.5))
plot(c(0.5,2.5)~c(4,26), gliomas, type="n", xlab="Time (months)",
  ylab="Survival OR")
abline(h=1)
colnew <- do.call(rgb,c(as.list(col2rgb("red")/255),list(0.2)))
polygon(c(times,rev(times)), c(predmod7ml[,2], rev(predmod7ml[,3])),
col=colnew,
  border=NA)
arrows(times, predmod4ml[,2], times, predmod4ml[,3], col=4, angle=90,
code=3,
  length=0.05)
points(predmod4ml[,1]~times, pch=19, col=4)
lines(predmod7ml[,1]~times, type="o", pch=19, col=2)
legend("top",c("Model 4 (Indicators)","Model 7 (continuous)"),
col=c(4,2),
  lty=c(NA,1),pch=19, cex=0.8, ncol=2, bty = "n", inset=0.05)

# WE COMPARE THE TWO MODELS
AIC(mod4ml, mod7ml)

# SEE help(gliomas) AND help(dbs) FOR FURTHER INFO
```

```
#############################################################################
#######
# SIMULATION STUDY (SECTION 6, TABLE 6)
#############################################################################
#######

# LOAD THE PACKAGES
library(mixmeta)

# DEFINE FIXED PARAMETERS
g2 <- 10
beta <- 0
tau <- 1

# DATA FRAME WITH COMBINATIONS OF SCENARIOS
comb <- data.frame(m=c(10,50), rhob=rep(c(0,0.8),each=2),
  rhow=rep(c(0,0.8),each=4))
comb$rhob <- ifelse(comb$rhob <= -1/(3-1), -1/3, comb$rhob)
comb$rhow <- ifelse(comb$rhow <= -1/(3-1), -1/3, comb$rhow)

# MATRIX WITH FINAL RESULTS
res <- matrix(NA, nrow(comb), 8)
stats <- c("bias","rmse","cov")
ran <- c("tau1","tau2","rhob1","rhob2")
colnames(res) <- c(paste("beta",stats,sep="-"),paste(ran,"bias",sep="-
"),"conv")

# NOMINAL VALUE
qn <- qnorm(0.975)

# NUMBER OF SIMULATIONS
nsim <- 1000

#############################################################################
#######

# START THE LOOP BY SCENARIO
for(i in seq(nrow(comb))) {

  # PRINT
  cat("\n\n ",paste("Combination",i),"\n")

  # DEFINE THE DATA FROM WHICH TO SIMULATE
  n <- comb$m[i] * g2
  y <- matrix(0,n,3)
  S <- inputcov(matrix(runif(n*3, 0.5, 2), n, 3), cor=comb$rhow[i])
  Psi1 <- Psi2 <- inputcov(rep(tau, 3), cor=comb$rhob[i])
  level1 <- rep(seq(comb$m[i]), each=g2)
  level2 <- rep(seq(g2), comb$m[i])

  # SIMULATE (WITH SEED)
  set.seed(13041975+i)
  simlist <-
mixmetaSim(y,S,Psi=list(Psi1,Psi2),random=~1|level1/level2,nsim=nsim)

  # BUILD THE TEMP0RARY OBJECT TO STORE THE ESTIMATES FROM EACH MODEL
  temp <- matrix(NA, nsim, 6,
dimnames=list(NULL,c("beta","beta.se",ran)))
```

```
#############################################################################
#######

  # START THE LOOP BY ITERATION
  for(j in seq(nsim)) {

    # PRINT
    cat(j,"")

    # FIT THE MODEL (PREVENT ERRORS DUE TO NON-CONVERGENCE)
    arglist <- list(simlist[[j]]~1, S=S, random=~1|level1/level2,
bscov="cs")
    model <- tryCatch(do.call("mixmeta",arglist), error=function(x) NULL)

    # STORE THE ESTIMATES (SET TO NA IF NON-CONVERGENCE)
    temp[j,] <- if(!is.null(model)) c(coef(model)[1],
sqrt(vcov(model)[1,1]),
      model$Psi[[1]][1,1], model$Psi[[2]][1,1],
cov2cor(model$Psi[[1]])[1,2],
      cov2cor(model$Psi[[2]])[1,2]) else NA
  }

#############################################################################
#######

  # COMPUTE THE STATS (REMOVING THE MISSINGS)
  res[i,8] <- sum(!is.na(temp[,1]))/nsim
  temp2 <- na.omit(temp)
  res[i,-c(2,3,8)] <- colMeans(temp2[,-2]) -
c(beta,tau,tau,rep(comb$rhob[i],2))
  res[i,2] <- sqrt(mean((temp2[,1]-beta)^2))
  res[i,3] <- mean(temp2[,1]-qn*temp2[,2]<=beta &
temp2[,1]+qn*temp2[,2]>=beta)

}

#############################################################################
#######

# SAVE THE RESULTS
#save.image("simul.RData")

#############################################################################
#######
# SIMULATION STUDY (SECTION 6, TABLE 6)
# WITH PARALLELIZATION: THIS ROUTINE MAY NOT WORK IN SOME OS AND MACHINE
#############################################################################
#######

# LOAD THE PACKAGES
library(mixmeta) ; library(foreach) ; library(doParallel)
library(iterators) ; library(parallel)

# DEFINE FIXED PARAMETERS
g2 <- 10
beta <- 0
tau <- 1

# DATA FRAME WITH COMBINATIONS OF SCENARIOS
```

```
comb <- data.frame(m=c(10,50), rhob=rep(c(0,0.8),each=2),
  rhow=rep(c(0,0.8),each=4))
comb$rhob <- ifelse(comb$rhob <= -1/(3-1), -1/3, comb$rhob)
comb$rhow <- ifelse(comb$rhow <= -1/(3-1), -1/3, comb$rhow)

# NAMES
stats <- c("bias","rmse","cov")
ran <- c("tau1","tau2","rhob1","rhob2")

# NOMINAL VALUE
qn <- qnorm(0.975)

# NUMBER OF SIMULATIONS
nsim <- 10000

##############################################################################
#######

# PREPARE THE PARALLELIZATION
ncores <- detectCores()
cl <- makeCluster(max(1,ncores-2))
registerDoParallel(cl)

##############################################################################
#######

# START THE NESTED LOOP BY SCENARIO/ITERATIONS
temp <- foreach(combi=iter(comb,by="row"), .packages=c("mixmeta")) %:%
  foreach(i=icount(nsim), .combine=rbind) %dopar% {

  # DEFINE THE DATA FROM WHICH TO SIMULATE
  n <- combi$m * g2
  y <- matrix(0,n,3)
  S <- inputcov(matrix(runif(n*3, 0.5, 2), n, 3), cor=combi$rhow)
  Psi1 <- Psi2 <- inputcov(rep(tau, 3), cor=combi$rhob)
  level1 <- rep(seq(combi$m), each=g2)
  level2 <- rep(seq(g2), combi$m)

  # SIMULATE (WITH SEED)
  set.seed(13041975+i)
  sim <- mixmetaSim(y,S,Psi=list(Psi1,Psi2),random=~1|level1/level2)

  # FIT THE MODEL (PREVENT ERRORS DUE TO NON-CONVERGENCE)
  arglist <- list(sim~1, S=S, random=~1|level1/level2, bscov="cs")
  model <- tryCatch(do.call("mixmeta",arglist), error=function(x) NULL)

  # RETURN THE VECTOR OF ESTIMATES (SET TO NA IF NON-CONVERGENCE)
  if(!is.null(model)) c(coef(model)[1], sqrt(vcov(model)[1,1]),
    model$Psi[[1]][1,1], model$Psi[[2]][1,1],
cov2cor(model$Psi[[1]])[1,2],
    cov2cor(model$Psi[[2]])[1,2]) else NA
}

##############################################################################
#######

# REMOVE PARALLELIZATION
stopCluster(cl)
```

```
##############################################################################
#######

# COMPUTE THE STATS (REMOVING THE MISSINGS)
res <- t(sapply(seq(temp), function(i) {
  nan <- sum(!is.na(temp[[i]][,1]))
  x <- na.omit(temp[[i]])
  bias <- colMeans(x[,-2]) - c(beta,tau,tau,rep(comb[i,"rhob"],2))
  c(bias[1],
    sqrt(mean((x[,1]-beta)^2)),
    mean(x[,1]-qn*x[,2]<=beta & x[,1]+qn*x[,2]>=beta),
    bias[-1],
    nan/nsim)
}))
colnames(res) <- c(paste("beta",stats,sep="-"),paste(ran,"bias",sep="-
"),"conv")

##############################################################################
#######

# SAVE THE RESULTS
#save.image("simul.RData")
```

Chapter 4

Research paper II

Title: Extended Two-Stage Designs for Environmental Research.

Author(s): Francesco Sera Antonio Gasparrini.

Journal/Publisher: Environmental Health.

Type of publication: Research paper.

Stage of publication: Published online on April 19, 2022 as doi.org/10.1186/s12940-022-00853-z.

URL: `https://ehjournal.biomedcentral.com/articles/10.1186/s12940-022-00853-z`.

Academic peer-reviewed: Yes.

Candidate's role: See Section 2.3.

Senior author: (Prof. Antonio Gasparrini)

Sera and Gasparrini *Environmental Health* (2022) 21:41
https://doi.org/10.1186/s12940-022-00853-z

Environmental Health

METHODOLOGY Open Access

Extended two-stage designs for environmental research

Francesco Sera[1,2*] and Antonio Gasparrini[2,3,4]

Abstract

Background: The two-stage design has become a standard tool in environmental epidemiology to model multi-location data. However, its standard form is rather inflexible and poses important limitations for modelling complex risks associated with environmental factors. In this contribution, we illustrate multiple design extensions of the classical two-stage method, all implemented within a unified analytic framework.

Methods: We extended standard two-stage meta-analytic models along the lines of linear mixed-effects models, by allowing location-specific estimates to be pooled through flexible fixed and random-effects structures. This permits the analysis of associations characterised by combinations of multivariate outcomes, hierarchical geographical structures, repeated measures, and/or longitudinal settings. The analytic framework and inferential procedures are implemented in the R package mixmeta.

Results: The design extensions are illustrated in examples using multi-city time series data collected as part of the National Morbidity, Mortality and Air Pollution Study (NMMAPS). Specifically, four case studies demonstrate applications for modelling complex associations with air pollution and temperature, including non-linear exposure–response relationships, effects clustered at multiple geographical levels, differential risks by age, and effect modification by air conditioning in a longitudinal analysis.

Conclusions: The definition of several design extensions of the classical two-stage design within a unified framework, along with its implementation in freely-available software, will provide researchers with a flexible tool to address novel research questions in two-stage analyses of environmental health risks.

Keywords: Environmental epidemiology, Two-stage design, Meta-analysis, Temperature, Pollution

Introduction

In environmental epidemiological studies, it is common practice to investigate short-term associations between environmental exposures and health outcomes by analysing data collected from multiple locations. An analytical approach applied in this setting is based on the two-stage design, which has become the standard method for the analysis of multi-location data [1–12]. The design is based on the separation of the analysis into two steps: in the first stage, location-specific exposure–response associations are estimated while adjusting for various confounders; then, in the second stage, the estimates are pooled using meta-analytic methods, which can potentially incorporate location-specific meta-predictors.

The two-stage design offers several advantages. First, the pooling of data collected in multiple locations increases the statistical power, thus facilitating the detection of small risks usually associated with environmental stressors [13]. At the same time, the separation in two steps provides a flexible and computationally efficient analytical framework compared to one-stage approaches [2, 14, 15]. This allows analyses of large datasets collected across multiple populations, increasing the

*Correspondence: francesco.sera@unifi.it
[1] Department of Statistics, Computer Science and Applications "G. Parenti", University of Florence, Florence, Italy
Full list of author information is available at the end of the article

representativeness of the findings. Finally, an important advantage of the two-stage design is the enhanced ability to examine heterogeneity in risk across populations, which can be linked to contextual characteristics.

However, there are known limitations of this analytical method. For instance, the standard two-stage design requires the association of interest to be represented by a single effect summary (*e.g.*, a relative risk or odds ratio) for being pooled in the second stage. However, in the context of modelling exposure–response associations, this step requires the simplification of potentially complex relationships and/or the adoption of strong functional assumptions (*e.g.*, linearity). Similarly, this restriction prevents combining multiple estimates of the association of interest from the same location, for example when collected from different age groups or periods. Finally, the standard two-stage analytic design does not take into account potential geographical dependencies, often occurring in the presence of clustering. These limitations represent important barriers to the application of the two-stage framework for addressing more complex research questions about environmental health risks.

In this contribution, we illustrate a unified framework that combines multiple design extensions of the classical two-stage method for environmental health studies, some of which were described independently in published analyses [6, 16–19]. This extended two-stage framework is based on linear mixed-effects meta-analytical models, previously developed and published by our research group [20], that can combine multivariate outcomes, longitudinal settings, multilevel structures, and/or repeated measurement [20]. This framework relaxes the constraints described above and offers a flexible and generally applicable tool to implement more advanced study designs using multi-location data.

The article is organized as follows. Firstly, we introduce the extended two-stage design and its features, including the design structure and related modelling framework. Then, after presenting the specific example and the related dataset, we will demonstrate applications of the various design extensions in multiple case studies using multi-location analyses of health risks of temperature and air pollution. In a final discussion section, we describe the epidemiological context, strengths and limitations, and area of further research. An up-to-date version of the notes, data, and R scripts for reproducing the examples are available on a GitHub repository (see Availability of data and material).

Methods

Extended two-stage design

In the classical two-stage design, the data are organised and analysed in first-stage models that provide independent estimates of a single parameter representing the association of interest in each study area, for instance, a city. These effect summaries are then pooled in the second stage using meta-analytic techniques to combine the information and compute an overall estimate. As discussed above, these requirements pose important analytical constraints. The extended two-stage described here overcomes these limitations, first allowing different estimates of single or multiple parameters to be computed in each location, and then relaxing the assumption of independence of estimates within and between locations.

This extended framework provides a flexible setting that allows designing more complex epidemiological studies to address more elaborated research questions. For example, in each study area, multiple parameters could be used to represent complex exposure–response dependencies, such as non-linear and lagged temperature-health associations of temperature [21], or alternatively correlated effects of multiple exposures, such as different pollutants included in the same first-stage model [22]. At the same time, relaxing the independence assumption allows accounting for correlations arising when the locations are nested within higher geographical levels (*e.g.*, cities within countries), therefore modelling patterns of similarities and differences [19]. Moreover, in each study area, the first-stage model can be applied multiple times to obtain repeated measures of the same association, for instance longitudinally at different times or for different sub-groups, such as by age or sex. This structure allows the investigation of temporal variations in risk [17] and the flexible pooling of effect modifications [16].

These analytic features, namely complex multivariate exposure–response relationships, geographical hierarchies, and longitudinal or repeated-measure structures can be incorporated individually or simultaneously in the extended two-stage framework, offering a flexible analytic context for modern environmental research studies.

Statistical framework

The extension of the two-stage design is made possible by the development of a unified statistical framework, previously developed and published by our research group [20], that specifies the second-stage meta-analysis as a mixed-effects linear model [20], as described below. Here we assume that estimates of the association of interest $\widehat{\boldsymbol{\theta}}_i$ have been obtained from each of the $i = 1, \ldots, n$ locations. Here $\widehat{\boldsymbol{\theta}}_i$ generally represents the output of the first-stage analysis (see appendix A), and it can include single or multiple coefficients obtained by single or repeated measurements across times or groups, depending on the specific application. In addition, without loss of generality, such estimates can be obtained from various types of first-stage models, such as time series for aggregated data

[23] or survival analysis of individual-level records [24], among others.

The first-stage estimates $\widehat{\boldsymbol{\theta}}_i$ can be combined in the second stage using an extended random-effects meta-analysis that flexibly models potentially complex dependence structures. This extended meta-analytical model can be written as a linear mixed-effects model:

$$\widehat{\boldsymbol{\theta}}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i \quad (1)$$

with $\boldsymbol{b}_i \sim N(0, \boldsymbol{\Psi})$, and $\boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{S}_i)$.

The design matrix $\boldsymbol{X}_i$, potentially expanded to account for multivariate outcomes, includes fixed-effect predictors and associated coefficients $\boldsymbol{\beta}$. Random terms are represented by the design matrix $\boldsymbol{Z}_i$ with coefficients $\boldsymbol{b}_i$, and by the errors $\boldsymbol{\varepsilon}_i$. The random terms have (co)variance matrices $\boldsymbol{\Psi}$ and $\boldsymbol{S}_i$, representing the deviations and errors between and within locations, respectively.

It is important to note that the association parameter $\widehat{\boldsymbol{\theta}}_i$ could have a general nested design with L level inducing possible non-independence of the estimates, e.g. associations estimated at multiple times, or in cities nested within a country. The extended framework naturally considers the nested design with a hierarchy of the random-effects effects vector$\boldsymbol{b}_i$, then $\boldsymbol{b}_i$ consists of the random coefficients operating on the levels (from outer to inner)$l = 1, \ldots, L$: $\boldsymbol{b}_i^T = \left(\boldsymbol{b}_{i1}^T, \ldots, b_{iL}^T\right)$, and the design matrix $\mathbf{Z}_i$ of the random terms has the corresponding $\mathrm{partitioning} \boldsymbol{Z}_i = (\boldsymbol{Z}_{i1}| \ldots |\boldsymbol{Z}_{iL}), \boldsymbol{Z}_{il} = (\boldsymbol{Z}_{il1}| \ldots |\boldsymbol{Z}_{ilnl})$. Note that every matrix $\boldsymbol{Z}_{ilj}$ has nonzero entries only in the rows that correspond to units in the group j ($j = 1, \ldots, n_l$) of level$l$.

The (co)variance matrix of the random terms has then the following structure:

$$\boldsymbol{\Psi} = \sum_{j=1}^{n_l} \boldsymbol{Z}_{ilj}\boldsymbol{\Psi}_l\boldsymbol{Z}_{ijl}^T$$

where $\boldsymbol{\Psi}_l$ is the covariance of the random terms operating at level l.

Example and data

The various extensions of the two-stage design will be illustrated using the same analytical example of multi-city time-series data collected as part of the National Morbidity, Mortality and Air Pollution Study (NMMAPS) [25]. This database contains, among other information, daily series of mortality counts and weather and pollution measurements totalling 5114 observations for the period 1987–2000 in each of 108 cities in the USA. This data resource has been used in several epidemiological analyses to assess health risks associated with air pollution and later with temperature [5, 26–30].

The NMMAPS data consisted of daily series of all-cause and cause-specific mortality, also stratified by age groups (0–64, 65–74, 65 and older), and various indices of daily levels of several pollutants and weather variables. In addition, the database included city-level metadata with several variables on geographical, climatological, demographic and socio-economic characteristics. The original datasets were collected on the 15th of May, 2013 through the package NMMAPSdata in the R software [31]. The package is now archived and the mortality series are not provided anymore. The data are here complemented with information on air conditioning use, collected longitudinally for a subset of cities and obtained from different sources [17].

The database is used in a series of case studies described in the next sections to illustrate the various extensions of the two-stage design. In each of them, we assume that first-stage models have been performed separately in the 108 locations, collecting summary estimates of association parameter(s) $\widehat{\boldsymbol{\theta}}_i$ and their (co) variance matrix $V\left(\widehat{\boldsymbol{\theta}}_i\right)$, and optionally location-specific metadata. These data are made available in a GitHub repository, together with the R code for the first stage to produce these quantities from the original data, and for the second stage to reproduce the results of the case studies (see Availability of data and material). Methodological and analytical details, in particular related to the first-stage modelling, are omitted to focus on specific aspects of the extensions of the two-stage design, with additional information provided in the Supplementary Material. As methodological case studies, these analyses should be considered illustrative examples and are not meant to offer substantive epidemiological evidence.

Results

Case study 1: modelling complex multi-parameter associations

Motivation

As mentioned earlier, an important limitation of the standard two-stage design is the need to simplify the relationship estimated in the first stage in a single effect summary, for it to be pooled in the second stage. This prevents the modelling of more complex associations represented by multiple parameters.

This limitation can be addressed by extending the two-stage design so that multiple quantities can be jointly combined in the second stage, using meta-analytic models that take into account their multivariate structure and their covariance (correlation) within and between locations. The meta-analytical methods can be further extended to multivariate meta-regression models that include specific predictors to explain (part of) the observed heterogeneity. This extension of the two stage

design has been known as a multivariate meta-analysis or multivariate meta-regression [15], and it can be represented as a specific parametrisation of the linear mixed effects meta-analytic framework presented above. These extensions can be implemented with the R package mvmeta [32] or with the updated and more general R package mixmeta [33].

In this case study, we offer an example of this extension to assess health risks associated with outdoor temperature, often characterised by marked non-linearity and heterogeneity of the effects across locations. In particular, we will investigate the association between heat and all-cause mortality during the summer months and the potential role of city-specific characteristics in modifying the risk. This extension of the two-stage design has been previously used in published analyses which evaluated the short-term health impacts of temperature [6, 16, 34].

Brief description of the data, model, and analysis

We assume that summer-only time series models have been fitted in each of the 108 NMMAPS cities to estimates temperature-mortality relationships using spline functions (see Supplementary Material B1), obtaining sets of four coefficients and their (co)variance matrices that represent the multi-parameter non-linear associations. In the second stage, we use these estimates as multivariate outcomes in the extended meta-analytical framework.

First, we fit a multivariate meta-analysis using a maximum likelihood (ML) estimator to pool the first-stage results and obtain an estimate of the average heat-mortality exposure–response curve. We then attempt to identify possible contextual characteristics that explain a quota of heterogeneity. Among potential factors, we consider population size, education (% of people with high-school degree) and unemployment (% of unemployed). These variables are included as predictors in multivariate meta-regressions, and their effects tested through likelihood ratio (LR) tests. Finally, a stepwise procedure is applied to select the best set among univariable and multivariable models. See Supplementary Material B1 for details.

Results

The basic multivariate meta-analytic model (with no predictors and only intercepts) produces pooled estimates of the set of coefficients representing the average heat-mortality association across the 108 cities. These coefficients can be used to compute the non-linear exposure–response curve expressed as relative risk (RR) by applying the same spline transformations on an average summer temperature distribution represented in a relative percentile scale [15]. The results are displayed in Fig. 1, showing a minimum mortality risk at low summer temperatures (MMT) and the sharp increase of the RR beyond the 90th percentile.

The simple meta-analysis shows a substantial heterogeneity in heat-mortality associations across cities, with an I^2 of 61.5% and a highly significant Cochran Q test (p-value < 0.001). Therefore, we assess if some of this heterogeneity was explained by some city characteristics, specifically population size, education, and

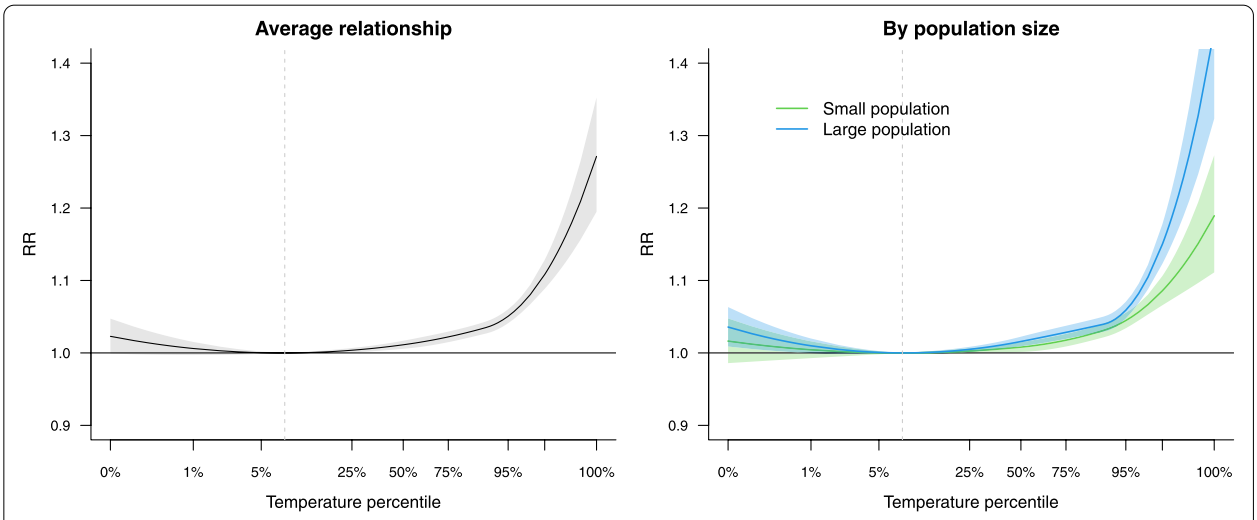


Fig. 1 Pooled association between relative temperature (percentiles) and all-cause mortality in 108 US cities during the summer period in 1987–2000 in Case Study 1. The x-axis is scaled so that the summer temperature distribution match the average percentiles of all the cities. The left panel shows the average heat-mortality curve estimated by the multivariate meta-analysis. The right panel illustrate the effect modification from population size, predicted from the full multivariate-meta-regression at the 10th-90th percentile values of the city-specific meta-variable

unemployment, by adding them as predictors in multivariate meta-regressions. Results are reported in Table 1. When tested separately in univariable models, each predictor is significantly associated with modification of the heat-mortality association. The full multivariable model identifies instead independent associations only for population size and unemployment, and these results are consistent with the selection of the forward stepwise procedure.

The tests above demonstrate an effect modification by specific city-level meta-variables, but provide little information on its direction. This can be identified by using the parameters of the multivariate meta-regression models to predict the multivariate outcome, namely the coefficients of the spline function representing the heat-mortality relationship, for given values of the meta-predictors. As an example, we used this method to isolate the effect modification of population size, keeping the other meta-predictors constant. The results, shown in the right panel of Fig. 1, indicate a higher mortality risk of heat in larger cities.

This case study demonstrates an extension of the two-stage design to pool multi-parameter associations. The specific example illustrates an application for complex exposure–response relationships, but the multi-parameter definition can be generalised, and the method is applicable for instance also to pool effects of multiple pollutants or multiple health outcomes [22].

Case study 2: modelling complex hierarchical structures

Motivation

Another important limitation of the standard two-stage design is the assumption of conditional independence between first-stage estimates. In environmental epidemiological associations, this assumption is invalid in the presence of geographical clustering, occurring when estimates are more similar in locations within the same region than between regions.

The two-stage design can be extended accordingly by modelling the dependencies among estimates through a hierarchical structure (*e.g.*, cities within countries, or countries within states). This extension can be implemented through a second-stage multilevel meta-analysis that defines multiple sets of random effects at different geographical levels.

In this case study, we provide an example in an analysis of the association between air pollution and non-accidental mortality in a multi-city time series study. Specifically, we assess the increased risk associated with exposure to ozone in a sample of NMMAPS cities accounting for clustering within states. We previously applied this extended two-stage design in a study evaluating the short-term health effects of pollutants [19].

Brief description of the model, data, and analysis

As in the previous case study, we assume that first-stage time series models have been performed in each city, collecting estimates of the log-RR for an increase in ozone of 10 $\mu g/m^3$, along with its variance as a measure of the uncertainty (see Supplementary material B2). Estimates for cities with no or limited daily measurements of ozone were set to missing, leaving a sample of 98 cities within 38 states.

We start the analysis by fitting a standard meta-analysis with city-specific random effects. Then, in order to account for potential geographical differences, we first perform a standard meta-regression with state indicators as fixed-effects predictors, and then the extended model including two levels of random effects by cities nested within states. Finally, we compute state-level fixed-effects predictions from the meta-regression, and best linear unbiased predictions at both city and state level from the multilevel model [20]. See Supplementary material B2 for details.

Table 1 Degrees of freedom (df), I^2, information criteria, and likelihood ratio (LR) tests for meta-predictors in second-stage multivariate regression models of Case Study 1. The last model selected by forward stepwise procedure includes only population size and unemployment

		df	I^2 (%)	AIC	BIC	LR test (*p*-value)
Model 0	Intercepts	14	61.5	-520.60	-463.64	
Model 1	+ population size	18	53.3	-529.81	-456.57	0.002
Model 2	+ education	18	58.1	-530.26	-456.80	0.002
Model 3	+ unemployment	18	55.7	-536.24	-463.11	< 0.0001
Model 4	Full model	26	48.3	-539.60	-433.82	
Model 5	Stepwise-selected model	22	49.7	-543.67	-454.16	

Results

The standard meta-analytic model with single-level random effects for cities returns a pooled RR of non-accidental mortality of 1.0037 (95%CI: 1.0027 to 1.0047), corresponding to a percentage increase of 0.37%, with a between-city variance equal to 0.0049^2. The inclusion of state indicators in the meta-regression suggests that there are significant geographical differences (LR test with a p-value < 0.001). Two drawbacks of this fixed-effects approach are the lack of a pooled effect estimate, and the high uncertainty in state-level predictions given the low number of cities within states and an highly-parameterised model.

The multilevel random-effects model addresses these limitations. First, this model provides a pooled relative risk of 1.0038 (95%CI: 1.0024 to 1.0051), with a similar point estimate and slighter higher confidence intervals than the standard meta-analysis. The between-group heterogeneity is split between states (0.0030^2) and cities (0.0040^2), suggesting variation at both levels. Figure 2 displays these geographical differences by mapping the city-level best linear unbiased predictions (BLUPs) of the RR for a 10 $\mu g/m^3$ increase in ozone.

Second, the multilevel model can improve the state-specific estimates by computing BLUPs at this geographical level. Figure 3 compares these quantities with fixed-effects predictions obtained from the standard meta-regression model. The results reveal the gain in precision of the BLUPs resulting from the shrinkage and borrowing of information across states [20]. These estimates are more reliable than fixed-effects predictions, where only the within-state information is used.

This case study illustrates how to extend the classical two-stage design by accounting for hierarchical dependencies between estimates from different locations. This flexible multilevel structure offers the possibility to separate the heterogeneity across geographical levels and to obtain more reliable and informative association estimates. The approach can be seamlessly extended to multi-parameter associations, combining multilevel and multivariate models [18].

Case study 3: sub-groups analysis, and dose–response relationships

Motivation

Common applications of the two-stage design entail the provision of single effect summaries from each location. However, the analysis can sometimes be repeated by sub-groups of the population defined by specific characteristics, such as sex or age, resulting in repeated measures and dependencies that the standard two-stage design is not able to handle.

The extended framework addresses this limitation, offering an adaptable grouping structure that allows multiple association estimates within a location. Moreover, the role of sub-groups characteristics can be flexibly examined in a

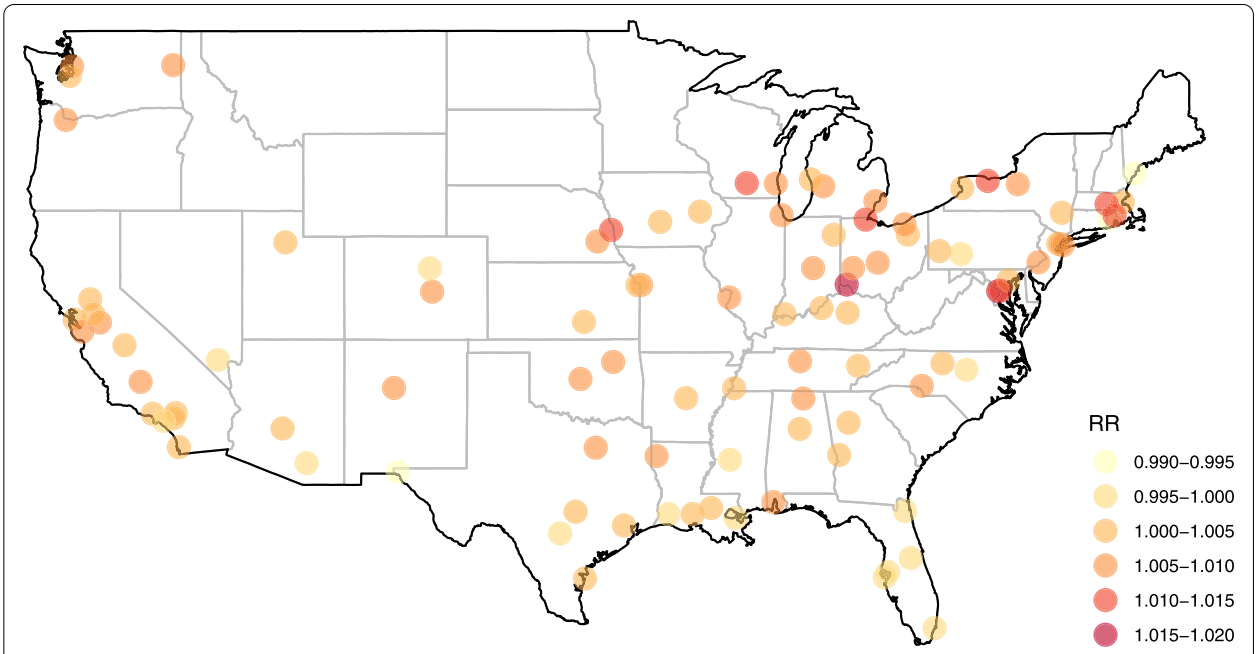


Fig. 2 City-level best linear unbiased predictions of the RR of non-accidental mortality for 10 $\mu g/m^3$ increase in ozone in 97 US cities (Honolulu not shown) during 1987–2000, as computed from the two-level random-effects meta-analysis in Case Study 2

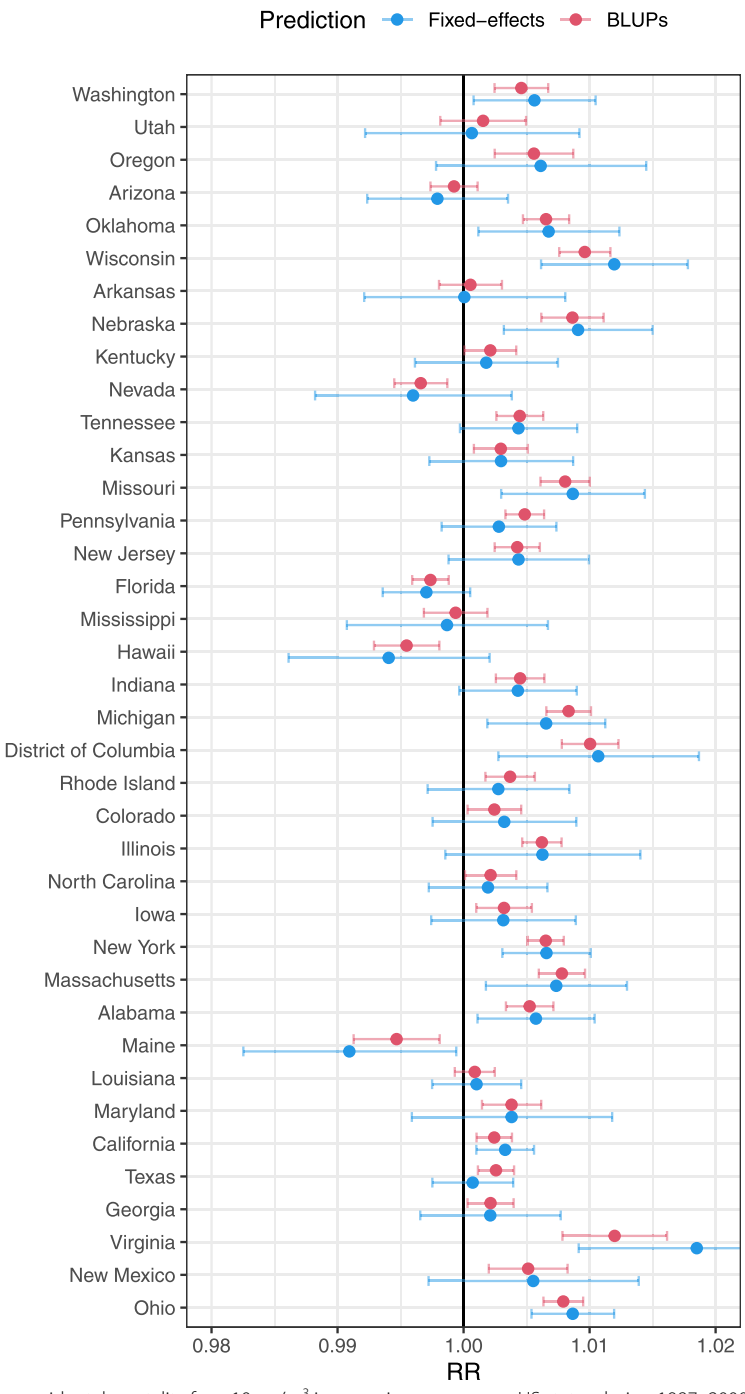


Fig. 3 Relative risk (RR) of non-accidental mortality for a 10 μg/m^3 increase in ozone across US states during 1987–2000 in Case Study 2. Estimates were obtained as state-level fixed-effects predictions from a standard meta-regression model (blue) and as best linear unbiased predictions (BLUPs) from a two-level random-effects model (red)

dose–response fashion by including either categorical and continuous variables in the fixed-part component. As for the extensions presented in the previous case studies, this framework is also applicable to multivariate outcomes.

In this case study, we extend further the investigation of the association between heat and all-cause mortality illustrated in Case Study 1 by stratifying the analysis by age. This provides repeated estimates for each of the 108

NMMAPS cities and the opportunity to apply flexible models to examine patterns of risk varying by age.

Brief description of the model, data, and analysis

The stratified analysis involves the fitting of the same first-stage regression model as in Case Study 1, but this time repeated separately for the three age groups (0–64, 65–74, 65 and older) using age-specific mortality series (see Supplementary material B3). We assume that this step has been performed and that we have obtained 324 sets of coefficients and associated (co)variance matrices representing age-specific heat-mortality associations in three age groups and 108 cities.

In the second stage, we first fit a standard meta-regression that ignores the city-level clustering and models the 324 multivariate outcomes using categorical indicators for age groups and unit-specific random effects. This model is first extended to account for clustering by defining the random-effect grouping structure at the city level. Then, we specify a continuous age variable by assigning specific values to the groups (60, 70, and 85 years) and finally we model it using either a linear or non-linear spline parametrisation. See Supplementary Material B3 for details.

Results

Table 2 offer a comparison between the different modelling strategies. All the models indicate evidence for an effect modification of age, but those correctly accounting for clustering by defining city-level groups (Models 1–3) demonstrate a better fit. The comparison of the more flexible models that define a continuous dose–response parametrisation (Models 2 and 3) suggests the presence of non-linearity. Note that the spline model (Model 3) has virtually an identical fit of the model with categorical indicators (Model 1), given that the number of groups/values equals the spline terms. However, the more flexible option defining the effect modification on a continuous scale has still some advantages, as illustrated below.`

The analysis has similarities to Case Study 1, which illustrated the effect modification related to city-specific variables, but, in this case, modelling within-city variations in risk. Still, the direction of the effect is difficult to ascertain when applying complex multi-parameter functions. Therefore, we rely on the same approach to predict average heat-mortality exposure–response curves for specific age values, taking advantage of the continuous dose–response parametrisation of the repeated-measure multivariate model. The results are reported in Fig. 4, suggesting a clear age pattern with the risk of heat increasing at older ages.

This case study shows how to extend the classical two-stage design to account for repeated measures originating, for instance, in the presence of multiple estimates from population sub-groups in the same location. This design extension also offers the possibility of modelling effect modifications by specific characteristics using flexible dose–response parametrisations on a continuous scale. It is interesting to note that this approach relaxes the requirement of defining fixed sub-groups (e.g., by age), as different values can be attributed across locations.

Case study 4: modelling longitudinal patterns of risk

Motivation

A different setting in which repeated measures can arise in two-stage analyses is when multiple estimates are collected at different times for the same location. This situation poses methodological problems that, similarly to the previous case study, standard designs are not equipped to handle.

The development of the two-stage methods to address these limitations requires accounting for the longitudinal structure of the data and modelling temporal trends in the exposure–response association. This extension provides environmental epidemiologists with the possibility of studying longitudinal patterns of risk, and considering potential time-varying factors explaining the variability of the estimated association over time.

Table 2 Comparison of various second-stage repeated-measure meta-analytical models to examine age-specific associations between heat and all-cause mortality in Case Study 3. The table report if clustering is accounted for, the parametrisation of age, the I^2 index and information criteria

	Clustering	Age parametrisation	I^2 (%)	AIC	BIC	LR test for age (*p*-value)
Model 0	No	Categorical	36.0	-480.99	-367.82	0.004
Model 1	Yes	Categorical	36.0	-553.06	-439.38	< 0.001
Model 2	Yes	Linear	36.9	-543.27	-450.26	< 0.001
Model 3	Yes	Non-Linear	36.0	-553.06	-439.38	< 0.001

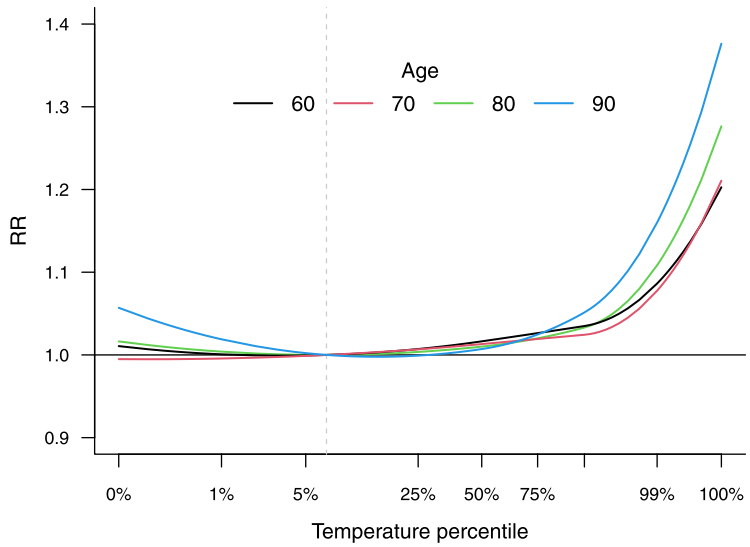


Fig. 4 Average temperature-mortality relationships across 108 US cities during the summer period in 1987–2000 predicted at different ages (in years) from the extended model with a continuous spline parametrisation (Model 3) in Case Study 3

In this case study, we again revise the analysis of heat-mortality relationships described in Case Study 1 by fitting the model in multiple sub-periods in each city. This step offers the opportunity to study temporal changes in the exposure–response curve and to assess the role of air conditioning (AC) in attenuating the risk. This case study is an illustrative example of a published analysis by our research group [17].

Brief description of the model, data, and analysis

We assume that in the first stage the data for the subset of 89 NMMAPS cities with information on AC data were split into five sub-periods (1987–98, 1990–92, 1993–95, 1996–98, and 1999–2000), and that separate time series models were fitted in each city/period combination, deriving a total of 445 sets of coefficients (co)variance matrices representing the multivariate association. Each city/period combination can be assigned a measure of AC prevalence use (%) reconstructed from an external database [17] (see Supplementary Material B4).

In a second step, we apply a longitudinal multivariate random-effect meta-regression to evaluate changes in heat-related mortality risks, accounting for both within and between-city variations. We include in the model a smooth spline function of calendar year and a linear term for AC as time-varying predictors, assessing their contribution with LR tests. As in the previous case study, this flexible continuous parametrisation allows the prediction of non-linear exposure–response curves for any given year and potential scenarios of AC use. See Supplementary Material B4 for details.

Results

The longitudinal meta-regression model suggests an independent effect of both calendar year (LR test p-value = 0.038) and air conditioning (p-value = 0.008). We evaluate their role by predicting the exposure–response associations in RR scale for different AC prevalence levels (80% vs 20%) in the year 1990. The curves are displayed in Fig. 5 (left panel), indicating how increasing AC has a protective effect at hot temperatures.

In order to assess the joint contributions of trends and AC use, we depict two scenarios to represent longitudinal changes in risk along years: a factual scenario using the observed trend in average AC prevalence, and a counterfactual scenario with AC use kept constant in time at the value of 1987. The right panel of Fig. 5 shows the results, summarising the heat effects as the RR computed at the 99th percentile versus the MMT along the period 1987–2000. The predicted risk under the counterfactual scenario (in blue) reveals a decreasing trend independent from AC use. Nonetheless, the comparison with the factual scenario (in red) suggests that the increase in AC prevalence during the period contributed somehow to attenuate the risk.

This last case study demonstrates the extension of the two-stage design to study longitudinal associations, evaluating changes in risk across both spatial and temporal dimensions. The flexibility of the extended framework allows parametrising effects on a continuous scale and performing second-stage meta-analysis with balanced and unbalanced data, with important design advantages.

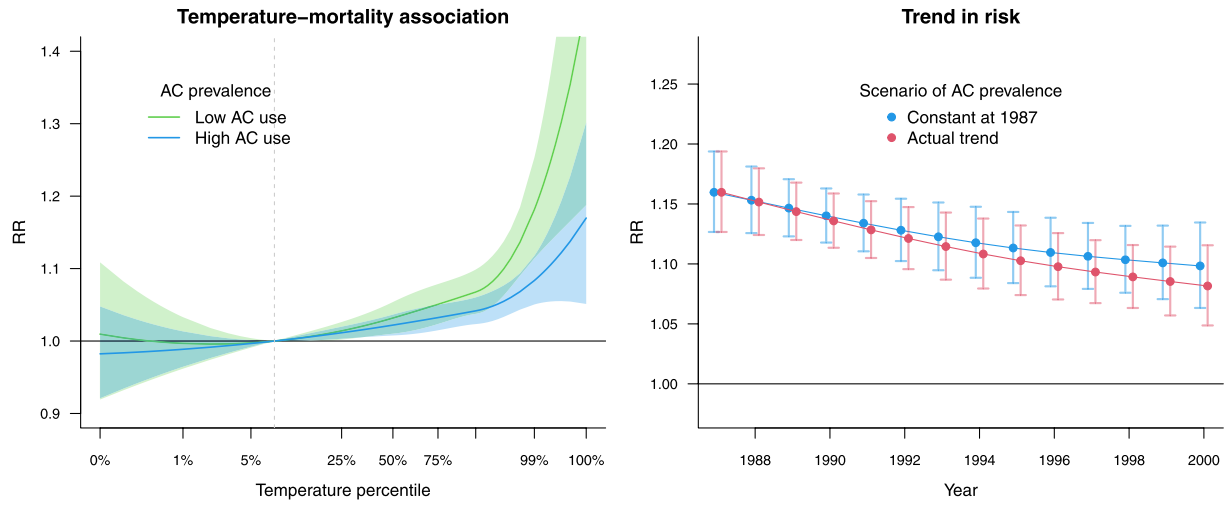


Fig. 5 *Left panel*: predicted average heat-mortality association (in RR) during the summer predicted for different air conditioning (AC) prevalence (20% and 80%) in Case Study 4. *Right panel*: trends in RR at 99th summer temperature predicted under two scenarios of AC use, corresponding to the observed average and a constant 1987 value

Discussion

In this contribution, we presented several design extensions of classical two-stage studies, and introduced several examples that illustrate how the flexibility of this modelling tool can improve the investigation of the effect of environmental exposures on health outcomes. Specifically, we showed how the extended two-stage design can be applied to investigate complex exposure–response dependencies, multilevel longitudinal structures, and repeated-measure dose–response associations. The analytic framework can be applied using classical inferential procedures and can be easily implemented using the R package mixmeta.

The two-stage design was proposed for the analysis of multi-location data. The methodology has been popularised by multi-city time series studies investigating short-term risk associations with environmental stressors [2, 5, 10], and it has become a common tool to assess the acute effects of pollutants [4, 7–9, 11] and temperature [1, 3]. The two-stage design has been also implemented in multi-cohort studies (e.g. ESCAPE project) to evaluate to long-term effect of pollutants [12, 24, 35], and in genetic epidemiology studies [36, 37]. Several extensions of a standard design have been proposed over the years, all of which can be represented as specific applications of the unified framework proposed here.

The most straightforward extension considers multiple estimates obtained in the first stage and the application of multivariate meta-analytic models in the second stage. This approach was originally developed to pool lagged effects [2], multiple pollutants [22], and non-linear dependencies [15], or more complex distributed lag non-linear associations [38].

Early applications of the two-stage design considered a small number of locations within a country, but the increased availability of environmental measures and health data now allows studies that include hundreds of locations within several countries [18, 19, 39]. In this setting, the locations can have a hierarchical structure that can be directly incorporated into the extended two-stage design. This extension has been proposed to obtain global, country, and city-level estimates of the associations by combining information within and between locations [18, 19, 39].

Environmental risk factors are often associated with risks that vary according to some individual or contextual characteristic [28, 40, 41]. The comparison of association measures across sub-groups was originally performed qualitatively and/or without consideration of the possible non-independence of multiple estimates collected within a location [42]. The extended two-stage design can directly model dependencies between the stratified estimates within each location, and appropriate inferential procedures can be used to evaluate differences across sub-group estimates.

In addition, such differences can be linked with measurable characteristics that can be included as categorical and continuous fixed-effects terms in the extended second-stage meta-regression. This extension allows modelling risks varying both within locations (*e.g.*, age in Case Study 3) and between locations (*e.g.*, population size and unemployment rate in Case Study 1). This effect

modification patterns can be modelled linearly or non-linearly using flexible parametric functions, representing a further extension of dose–response pooling methods applied in observational studies [43, 44].

With the availability of longer time series of environmental exposures and health outcomes, researchers have started to investigate the temporal variation in associations of short term environmental exposures and health outcomes [3, 17, 45–49]. In particular, modelling approaches have proposed time-varying extensions of distributed lag non-linear models [47, 48], Bayesian hierarchical models [3, 46], and functional meta-regression [49]. The extended two-stage design naturally accommodates balanced and unbalanced association parameters longitudinally directly accounting for possible non-independences, and it provides the possibility to parametrise trends through linear and non-linear functions. It is important to note that the longitudinal setting can incorporate other extensions, such as multivariate outcomes and multilevel structures, modelling potentially complex structures of longitudinal associations [17].

The data example and the four case studies are consistent with the most common application of the two-stage design in time series analysis of short-term effects of environmental exposures. However, it is worth noting that the framework proposed here is not restricted to the time series setting, and first-stage estimates can be obtained by any other approach such as case-crossover or time-to-event Cox models. Therefore, the extended two-stage design can similarly be applied in environmental epidemiological studies investigating either short or long-term effects of environmental exposure, using either individual-level or aggregated cross-sectional, case–control, and cohort data [12, 24, 35–37, 50].

An important advantage of the proposed development is the fact that it is grounded on a unified likelihood-based inferential framework and implemented in freely available and easy-to-use software. All the analyses illustrated in the four case studies can be performed using the R package mixmeta, which offers a simple syntax to define all the different models and combinations of them. Similar extensions of the two-stage design were proposed based on Bayesian hierarchical models, for instance for multivariate [22], multilevel [14] and longitudinal data [46], but they usually require advanced statistical and programming skills and can be computationally more demanding. Nonetheless, the Bayesian framework offers more flexibility in accommodating random-effects and correlations, for instance spatial structures that are not yet available and generally more difficult to implement in our likelihood-based development.

Conclusions

Technological developments in environmental monitoring, coupled with advancements in data linkage and collaborative tools, offer new opportunities for researchers to collect large multi-locations databases. The development of a general and extended framework for two-stage designs is therefore timely and offers a flexible and generally applicable tool for modern environmental epidemiological studies.

Abbreviations

NMMAPS: National Morbidity, Mortality and Air Pollution Study; ML: Maximum likelihood; LR: Likelihood ratio; RR: Relative risk; MMT: Minimum mortality risk temperature; BLUPs: Best linear unbiased predictions; AC: Air conditioning; AIC: Akaike information criterion; BIC: Bayesian. information criterion.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12940-022-00853-z.

Additional file 1.

Acknowledgements

Not applicable

Authors' contributions

FS and AG conceptualise the research goals and aims, FS and AG developed the methods and the software, FS and AG analysed the data, FS was a major contributor in writing the manuscript that was revised and approved by AG.

Funding

AG was funded by the Medical Research Council-UK (Grant ID: MR/M022625/1), the Natural Environment Research Council UK (Grant ID: NE/R009384/1) and the European Union's Horizon 2020 Project Exhaustion (Grant ID: 820655).

Availability of data and materials

An up-to-date version of the R scripts and data to fully reproduce the examples described in the four case studies are added in a GitHub repository, available at https://github.com/gasparrini/extended2stage.

Declarations

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Statistics, Computer Science and Applications "G. Parenti", University of Florence, Florence, Italy. [2]Department of Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, London, UK. [3]Centre On Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, UK. [4]Centre for Statistical Modelling, London School of Hygiene & Tropical Medicine, London, UK.

Received: 8 November 2021 Accepted: 8 April 2022
Published: 19 April 2022

References

1. Baccini M, Biggeri A, Accetta G, Kosatsky T, Katsouyanni K, Analitis A, et al. Heat effects on mortality in 15 European cities. Epidemiology. 2008;19(5):711–9.
2. Berhane K, Thomas DC. A two-stage model for multiple time series data of counts. Biostatistics. 2002;3(1):21–32.
3. Bobb JF, Peng RD, Bell ML, Dominici F. Heat-Related Mortality and Adaptation to Heat in the United States. Environ Health Perspect. 2014;122(8):811–6.
4. Chen R, Kan H, Chen B, Huang W, Bai Z, Song G, et al. Association of particulate air pollution with daily mortality: the China Air Pollution and Health Effects Study. Am J Epidemiol. 2012;175(11):1173–81.
5. Dominici F, Samet JM, Zeger SL. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. J R Stat Soc Ser. 2000;163:263–84.
6. Gasparrini A, Guo YM, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet. 2015;386(9991):369–75.
7. Liu C, Chen R, Sera F, Vicedo-Cabrera AM, Guo YM, Tong SL, et al. Ambient Particulate Air Pollution and Daily Mortality in 652 Cities. N Engl J Med. 2019;381(8):705–15.
8. Romieu I, Gouveia N, Cifuentes LA, de Leon AP, Junger W, Vera J, et al. Multicity study of air pollution and mortality in Latin America (the ESCALA study). Res Rep Health Eff Inst. 2012;171:5–86.
9. Samoli E, Analitis A, Touloumi G, Schwartz J, Anderson HR, Sunyer J, et al. Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. Environ Health Perspect. 2005;113(1):88–95.
10. Schwartz J. Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths. Environ Health Perspect. 2000;108(6):563–8.
11. Wong CM, Vichit-Vadakan N, Kan H, Qian Z. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. Environ Health Perspect. 2008;116(9):1195–202.
12. Basagana X, Pedersen M, Barrera-Gomez J, Gehring U, Giorgis-Allemand L, Hoek G, et al. Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. Int J Epidemiol. 2018;47(4):1343–54.
13. Armstrong BG, Gasparrini A, Tobias A, Sera F. Sample size issues in time series regressions of counts on environmental exposures. Bmc Med Res Methodol. 2020;20(1):1–9.
14. Dominici F, Daniels M, Zeger SL, Samet JM. Air pollution and mortality: Estimating regional and national dose-response relationships. J Am Stat Assoc. 2002;97(457):100–11.
15. Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. Stat Med. 2012;31(29):3821–39.
16. Sera F, Armstrong B, Tobias A, Vicedo-Cabrera AM, Astrom C, Bell ML, et al. How urban characteristics affect vulnerability to heat and cold: a multi-country analysis. Int J Epidemiol. 2019;48(4):1101–12.
17. Sera F, Hashizume M, Honda Y, Lavigne E, Schwartz J, Zanobetti A, et al. Air Conditioning and Heat-related Mortality A Multi-country Longitudinal Study. Epidemiology. 2020;31(6):779–87.
18. Vicedo-Cabrera AM, Scovronick N, Sera F, Roye D, Schneider R, Tobias A, et al. The burden of heat-related mortality attributable to recent human-induced climate change. Nat Clim Chang. 2021;11(6):492–500.
19. Vicedo-Cabrera AM, Sera F, Liu C, Armstrong B, Milojevic A, Guo YM, et al. Short term association between ozone and mortality: global two stage time series study in 406 locations in 20 countries. BMJ. 2020;368. https://www.bmj.com/content/368/bmj.m108.
20. Sera F, Armstrong B, Blangiardo M, Gasparrini A. An extended mixed-effects framework for meta-analysis. Stat Med. 2019;38(29):5429–44.
21. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. Stat Med. 2014;33(5):881–99.
22. Dominici F, Zanobetti A, Zeger SL, Schwartz J, Samet JM. Hierarchical bivariate time series models: a combined analysis of the effects of particulate matter on morbidity and mortality. Biostatistics. 2004;5(3):341–60.
23. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. Int J Epidemiol. 2013;42(4):1187–95.
24. Cesaroni G, Forastiere F, Stafoggia M, Andersen ZJ, Badaloni C, Beelen R, et al. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. BMJ. 2014;348. https://www.bmj.com/content/348/bmj.f7412.full.
25. Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. The National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and methodologic issues. Res Rep Health Eff Inst. 2000;94 Pt 1:5–14. discussion 75-84.
26. Daniels MJ, Dominici F, Zeger SL, Samet JM. The National Morbidity, Mortality, and Air Pollution Study. Part III: PM10 concentration-response curves and thresholds for the 20 largest US cities. Res Rep Health Eff Inst. 2004;94 Pt 3:1–21. discussion 3–30.
27. Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: Mortality among residents of 90 cities. J Toxicol Environ Health-Part a-Current Issues. 2005;68(13–14):1071–92.
28. Huang Y, Dominici F, Bell ML. Bayesian hierarchical distributed lag models for summer ozone exposure and cardio-respiratory mortality. Environmetrics. 2005;16(5):547–62.
29. Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, et al. The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and mortality from air pollution in the United States. Res Rep Health Eff Inst. 2000;94(Pt 2):5–70 discussion 1-9.
30. Zhang YQ, Xiang QQ, Yu Y, Zhan ZY, Hu KJ, Ding Z. Socio-geographic disparity in cardiorespiratory mortality burden attributable to ambient temperature in the United States. Environ Sci Pollut Res. 2019;26(1):694–705.
31. Peng RD, Welty LJ. The nmmapsdata package. R news. 2004;4(2):10–4.
32. Gasparrini A, Gasparrini MA. Package 'mvmeta.' 2019.
33. Gasparrini A, Sera F, Gasparrini MA. Package 'mixmeta.' 2021.
34. Scovronick N, Sera F, Acquaotta F, Garzena D, Fratianni S, Wright CY, et al. The association between ambient temperature and mortality in South Africa: A time-series analysis. Environ Res. 2018;161:229–35.
35. Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet. 2014;383(9919):785–95.
36. Raimondi S, Gandini S, Fargnoli MC, Bagnardi V, Maisonneuve P, Specchia C, et al. Melanocortin-1 receptor, skin cancer and phenotypic characteristics (M-SKIP) project: study design and methods for pooling results of genetic epidemiological studies. Bmc Med Res Methodol. 2012;12:1–13.
37 Surendran P, Feofanova EV, Lahrouchi N, Ntalla I, Karthikeyan S, Cook J, et al. Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. Nat Genet. 2020;52(12):1314–32.
38. Gasparrini A, Armstrong B. Reducing and meta-analysing estimates from distributed lag non-linear models. Bmc Med Res Methodol. 2013;13:1–10.
39. Meng X, Liu C, Chen RJ, Sera F, Vicedo-Cabrera AM, Milojevic A, et al. Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities. BMJ. 2021;372. https://www.bmj.com/content/372/bmj.n534.
40. Son JY, Liu JC, Bell ML. Temperature-related mortality: a systematic review and investigation of effect modifiers. Environ Res Letters. 2019;14(7):073004.
41. Zeka A, Zanobetti A, Schwartz J. Individual-level modifiers of the effects of particulate matter on daily mortality. Am J Epidemiol. 2006;163(9):849–59.
42. Nordio F, Zanobetti A, Colicino E, Kloog I, Schwartz J. Changing patterns of the temperature-mortality association by time and location in the US, and implications for climate change. Environ Int. 2015;81:80–6.
43. Crippa A, Discacciati A, Bottai M, Spiegelman D, Orsini N. One-stage dose-response meta-analysis for aggregated data. Stat Methods Med Res. 2019;28(5):1579–96.
44. Orsini N, Li RF, Wolk A, Khudyakov P, Spiegelman D. Meta-Analysis for Linear and Nonlinear Dose-Response Relations: Examples, an Evaluation of Approximations, and Software. Am J Epidemiol. 2012;175(1):66–73.

45. Carugno M, Consonni D, Bertazzi PA, Biggeri A, Baccini M. Temporal trends of PM10 and its impact on mortality in Lombardy. Italy Environ Pollut. 2017;227:280–6.
46. Chen C, Warrington JA, Dominici F, Peng RD, Esty DC, Bobb JF, et al. Temporal variation in association between short-term exposure to fine particulate matter and hospitalisations in older adults in the USA: a long-term time-series analysis of the US Medicare dataset. Lancet Planetary Health. 2021;5(8):E534–41.
47. Chung Y, Yang D, Gasparrini A, Vicedo-Cabrera AM, Ng CFS, Kim Y, et al. Changing Susceptibility to Non-Optimum Temperatures in Japan, 1972–2012: The Role of Climate, Demographic, and Socioeconomic Factors. Environ Health Perspect. 2018;126(5):057002.
48. Gasparrini A, Guo YM, Hashizume M, Kinney PL, Petkova EP, Lavigne E, et al. Temporal Variation in Heat-Mortality Associations: A Multicountry Study. Environ Health Perspect. 2015;123(11):1200–7.
49. Yu J, Park J, Choi T, Hashizume M, Kim Y, Honda Y, et al. Nonparametric Bayesian Functional Meta-Regression: Applications in Environmental Epidemiology. J Agric Biol Environ Stat. 2021;26(1):45–70.
50. Sera F, Ferrari P. A Multilevel Model to Estimate the Within- and the Between-Center Components of the Exposure/Disease Association in the EPIC Study. Plos One. 2015;10(3):e0117815.

Publisher's Note

On-line Appendix

A. First-stage time-series models

In the examples described in the four case studies, the first-level data are time series, and were modelled accordingly. In each first-level unit *i*, a set of N_i observations are defined at equally-spaced time intervals (here corresponding to days) t=1,…,N_i. In the first stage, the aim is to estimate the association between the exposure x_{it} and the outcome Y_{it} after controlling for the set of the time-varying confounders $\mathbf{c}_{it}$, in addition to time trends. Usually, this aim is achieved by fitting generalized linear models (usually with Poisson or quasi-Poisson family) (1):

$$g[\mathrm{E}(Y_{it})] = \alpha + f(x_{it}, \ell; \boldsymbol{\theta}_i) + \sum_{q=1}^{Q} s_{qi}(t; \boldsymbol{\delta}_i) + \sum_{p=1}^{P} h_{pi}(\boldsymbol{c}_{ipt}; \boldsymbol{\gamma}_i) \qquad (1)$$

Where the function $f(x_{it}, \ell; \boldsymbol{\theta}_i)$ specifies the association with the exposure of interest x, allowing non-linearity and complex temporal dependencies along the lag dimension ℓ. These complex relationships can be modelled through distributed lag linear and non-linear models (DLMs and DLNMs), which can flexibly define cumulative effects of multiple exposure episodes (2). The term(s) s_{qi} represent functions expressed at different timescales to model temporal variations in risk associated with underlying trends or seasonality, among others.

B. Modelling details about the four case studies

B1. Pooling complex multi-parameter associations

In the first stage, we fitted distributed lag non-linear time-series models for each of the 108 US cities in order to estimate the association between mean temperature and overall mortality in summer months (June to September). Briefly, the bi-dimensional cross-basis

function was composed of a quadratic B-spline system with two internal knots used to represent non-linearity in the temperature dimension, and an unconstrained parameterization in the lag space within 0-3 days (2). The coefficients of the cross-basis were then "reduced" over the lag dimension to obtain parameters representing the non-linear net association between mean temperature and overall mortality (2) obtaining a set of four coefficients vector $\widehat{\boldsymbol{\theta}}_i$ (4×1) and their (co)variance matrix $\boldsymbol{S}_i$ (4×4) for each of the 108 cities (3). The confounding effect of long-term trends and seasonality was modelled through interaction of a smooth function of day of the year (natural splines with four degrees of freedom) and year.

In the second-stage, the extended random-effects meta-analysis can be used to obtain the pooled set of coefficients taking into account the dependencies and uncertainty of the coefficients as measured by the covariance matrix of the coefficient set for each city.

In terms of the extended framework, this represents an example of multivariate meta-analysis, where the design matrices of the fixed and random effects are identity matrices with dimensions equal to the number of coefficients: $\boldsymbol{X}_i = \boldsymbol{Z}_i = \boldsymbol{I}_4$. Note that the random coefficient vector $\boldsymbol{b}_i$ is a (4×1) row vector.

We then fitted meta-regression models to identify contextual factors that could explain a quota of the heterogeneity. Among several potential factors, we considered population size, percentage of people with high-school degree, and percentage of unemployment. These variables will be included in the fixed-effects part of the meta-analytic model by setting $\boldsymbol{X}_i = \boldsymbol{x}_i^T \otimes \boldsymbol{I}_4$, with $\otimes$ is the Kronecker product operator, where $\boldsymbol{x}_i^T$ are the variable values measured in first-level unit *i*. With this parametrisation, in equation (1), $\boldsymbol{\beta}$ is the (12 X 1)

dimensional coefficient vector that defines the association of the 4 outcomes with the 3 predictors.

B2. More complex hierarchical structure

In the first-stage, we fitted quasi-Poisson time-series models for each of the 85 US cities with available data in order to estimate the association between ozone and non-accidental mortality. Ozone was modelled linearly over the moving average of 0-1 lags. The parameters of the DLM model were "reduced" over the lag dimension, similarly to the previous case stury, obtaining a single coefficient (log-RR for an increase of 10 μg/m3) and its variance for each of the 85 cities. We considered the confounding effect of temperature and long-term trend/seasonality. Briefly, temperature was modelled using a cross-basis composed of quadratic B-spline function with three internal knots (at the 10th, 75th, and 90th percentiles) to represent non-linear exposure-response relationships, and and unconstrained parametrisation within 0-3 days to model the lag-response shape. The confounding effect of long-term trends was modelled with a smooth function of day of the year (natural splines with four 7 degrees of freedom per year).

We estimated the log-RR for an increase of 10 μg/m3 in each of the 85 cities, along with its variance as a measure of the precision (uncertainty). These estimates are nested within 35 states.

Within the extended meta-analytic framework, we structured the first stage estimated coefficients $\widehat{\boldsymbol{\theta}}_i \ (\mathrm{n}_i \times 1)$ considering two hierarchical levels $L = 2$, with cities nested within US states. Note that $\widehat{\boldsymbol{\theta}}_i$ is defined at the highest hierarchical level (US states) with n_i cities, and that cities and US states have specific random effects (e.g. random intercepts) $\boldsymbol{b}_{i1}^T, \boldsymbol{b}_{i2}^T$ with

design matrices $\boldsymbol{Z}_{i1}$ $(n_i \times 1)$, a column of ones, and $\boldsymbol{Z}_{i2}$ $(n_i \times n_i)$ with ones on the specific row and column identifying the city. Here $\boldsymbol{X}_i = \mathbf{1}$, $(n_i \times 1)$, a column of ones and specifies the intercept term. This is an example of multilevel meta-analytic model.

B3. Sub-group analysis, modelling dependencies and dose-response relationship

The NMMAPS dataset provides daily time-series of overall mortality data for three age classes: [0, 65), [65, 75), and [75+) for each of the 108 US cities. The analysis follows the same scheme as in Section B1, only repeated by each age class. Thus, we obtained 324 sets of coefficients and associated (co)variance matrix describing the age-specific temperature mortality association nested within 108 cities (i.e., each city has three coefficients).

When considering the clustering effect of cities, the first stage estimates $\widehat{\boldsymbol{\theta}}_i$ is a (12×1) vector (4 splines coefficients estimated for the three age classes) and the within-location (co)variance matrix $\boldsymbol{S}_i = \oplus_a \boldsymbol{S}_{i,a}$ is a block diagonal matrix where diagonal blocks $\boldsymbol{S}_{i,a}$ are the covariance matrices of the coefficients calculated in each age class (a). We modelled age as a fixed-effects variable by stacking by rows the matrices $\boldsymbol{X}_{i,a}$, with $\boldsymbol{X}_{i,a} = \boldsymbol{x}_{i,a}^T \otimes \boldsymbol{I}_4$, where $x_{i,a}^T$ is the vector representing the age class (a) in the different parametrisations: (categorical; 2 dummies variables), (linear; 1 variable), (non-linear; natural splines with two degrees of freedom). Note that in this case the parametrisation of age as categorical and non-linear with the chosen cut-offs of 70 years are equivalent, as there are three age groups. However, the two model would be different with a higher number of categories. The design matrix of the random terms can be defined by stacking by rows the three matrices $\boldsymbol{Z}_{i,a} = \boldsymbol{I}_4$.

B4. Longitudinal

For each of the 108 US cities the study interval (1987-2000) was split into five periods: [1987, 1989], [1990, 1992], [1993, 1995], [1996, 1998] and [1999, 2000]. For each period, we fitted distributed lag non-linear time-series models during summer months (June to September). The analysis follows the same scheme as in Section B1, only repeated by period. Thus, we obtained 540 sets of coefficients and associated (co)variance matrix describing the period-specific temperature mortality association nested within 108 cities.

Here $\widehat{\boldsymbol{\theta}}_i$ is a (20×1) vector (four splines coefficients estimated in the five periods) and the within location (co)variance matrix $\boldsymbol{S}_i = \oplus_p \boldsymbol{S}_{i,p}$ is a block diagonal matrix where diagonal blocks $\boldsymbol{S}_{i,p}$ are the (co)variance matrices of the coefficients calculated in each period (p).

In addition, we reconstructed location-specific AC air conditioning trends and assigned AC prevalence estimates to each location/period unit (4).

We modelled AC prevalence and calendar year as fixed-effects variables by stacking by rows the matrices $\boldsymbol{X}_{i,p}$, with $\boldsymbol{X}_{i,p} = \boldsymbol{x}_{i,p}^T \otimes \boldsymbol{I}_4$, where $\boldsymbol{x}_{i,p}^T$ is the vector representing the AC prevalence (linear term) and calendar year (natural spline with one internal knot) in period (p). The design matrix of the random terms can be defined by stacking by rows the five matrices $\boldsymbol{Z}_{i,p} = \boldsymbol{I}_4$.

C. Backward or step-forward procedures

The analyst could use different procedures to select relevant predictors (e.g. backward or forward stepwise). The backward selection approach starts from the full model (with the 3 predictors), and a likelihood ratio test is performed to compare the full model and the models without each predictor. The step-forward approach starts instead from the model

containing only the intercept (Model 0), and at each step the predictor with highest reduction of the Akaike information criterion (AIC) is included.

References

1. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. International Journal of Epidemiology. 2013;42(4):1187-95.
2. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. Statistics in Medicine. 2014;33(5):881-99.
3. Gasparrini A, Armstrong B. Reducing and meta-analysing estimates from distributed lag non-linear models. Bmc Medical Research Methodology. 2013;13.
4. Sera F, Hashizume M, Honda Y, Lavigne E, Schwartz J, Zanobetti A, et al. Air Conditioning and Heat-related Mortality A Multi-country Longitudinal Study. Epidemiology. 2020;31(6):779-87.

R code to reproduce the four case studies presented in the article:

Sera F, Gasparrini A. Extended two-stage designs for environmental research. *Environmental Health*. 2022;21:41. DOI: 10.1186/s12940-022-00853-z. [[freely available here](http://www.ag-myresearch.com/2022_sera_envhealth.html)]

The five R scripts prepare the data and perform the analyses in each of the four case studies. Specifically:

00.prepdata.R loads the original NMMAPS data and performs the first-stage time series models to obtain exposure summaries and estimates of short-term risk associations with air pollution and temperature. These data are used in the following R scripts to perform the second-stage meta-analytical models for each design extension (Note: to run this scripts, the user needs to gather the original NMMAPS data, which is not made available here).

01.heterogeneity.R reproduces the first case study, illustrating the extension of the two-stage design for pooling multi-parameter associations. Specifically, the case study investigates the non-linear and delayed relationship between heat and all-cause mortality during the summer months and the potential role of city-specific characteristics in modifying the risk.

02.multilevel.R reproduces the second case study, describing the extension of the two-stage design for the analysis of complex hierarchical structures and geographical clustering. Specifically, the case study demonstrates how to combine estimates of associations between ozone and mortality in multiple cities nested within states.

03.doseresp.R performs the analysis of the third case study, showing the extension of the two-stage design for sub-groups analysis and dose–response relationships. Specifically, the case study illustrates how the heat-mortality association can be estimated when there are repeated measures from the same city, resulting from multiple first-stage models fitted by different age groups. The case study also shows how to flexibly model the age effect in a dose-response fashion.

04.longitudinal.R reproduces the fourth case study, illustrating the extension of the two-stage design for longitudinal analysis of estimates collected along time. Specifically, the case study examines the temporal changes in the exposure–response curve between heat and mortality, and assesses the role of air conditioning (AC) in attenuating the risk.

```
# 00.prepdata.R

################################################################################
# R code for the analysis in:
#
#  Sera F, Gasparrini A. Extended two-stage designs for environmental research.
#    Environmental Health. 2022;21:41.
#  https://doi.org/10.1186/s12940-022-00853-z
#
# * an updated version of this code, compatible with future versions of the
#   software, is available at:
#   https://github.com/gasparrini/Extended2stage
################################################################################

################################################################################
# RUN FIRST-STAGE MODELS AND SAVE DATA
################################################################################

# LOAD THE PACKAGES
library(dlnm) ; library(mixmeta)
library(tsModel) ; library(splines)
library(lubridate)

# PATH TO NMMAPS CITY-SPECIFIC SERIES
# NB: TO BE REPLACED BY THE PATH TO THE ORIGINAL NMMAPS DATA
path <- ""

# LOAD THE DATASET WITH METADATA FOR THE CITIES
cities <- read.csv("data/cities.csv")

################################################################################
# LOOP

# CREATE OBJECTS TO STORE THE RESULTS
tmeanparlist <- tmeanperparlist <- tmeanageparlist <- o3parlist <-
  tmeansumlist <- vector("list", nrow(cities))

# RUN THE LOOP
for(i in seq(nrow(cities))) {

  # PRINT CITY
  cat(cities$city[i],"")

  # LOAD THE DATA NOT COLLAPSED BY AGE
  load(paste0(path, "/cities/", cities$city[i]))

  # SELECT VARIABLES
  vars <-
c("date","dow","death","accident","tmpd","o3tmean","o3mtrend","agecat")
  data <- subset(data, select=vars)

  # CONVERT ENVIRONMENTAL VARIABLES
  # - TEMPERATURE FROM FAHRENHEIT TO CELSIUS
  # - OZONE FROM DETRENDED AND CONVERTED IN MICROGR/M3
  data <- transform(data,
    tmean = (tmpd-32)*5/9,
      o3 = (o3tmean+o3mtrend) * 1.96
  )

  # CREATE COLLAPSED DATA (SELECT ONE AGE GROUP AND COLLAPSE DEATH CAUSES)
  dataggr <- data[data$agecat=="under65",]
  dataggr$death <- tapply(data$death, data$date, sum)
```

```
  dataggr$accident <- tapply(data$accident, data$date, sum)

  # CREATE ALL-CAUSE
  data$all <- data$death + data$accident
  dataggr$all <- dataggr$death + dataggr$accident

  # SUBSET FOR SUMMER-ONLY
  datasum <- subset(data, month(date) %in% 6:9)
  dataggrsum <- subset(dataggr, month(date) %in% 6:9)

#################################################################################
# ANALYSIS OF TEMPERATURE - ALL-CAUSE MORTALITY (SUMMER-ONLY)

     # DEFINE CROSS-BASIS FOR TEMPERATURE
  cbtmean <- crossbasis(dataggrsum$tmean, lag=3, argvar=list(fun="bs", degree=2,
    knots=quantile(dataggrsum$tmean, c(50,90)/100, na.rm=T)),
    arglag=list(fun="integer"), group=year(dataggrsum$date))

  # RUN THE MODEL AND EXTRACT REDUCED PRED (RE-CENTRED LATER)
  model <- glm(all ~ cbtmean + dow + ns(yday(date), df=4)*factor(year(date)),
    data=dataggrsum, family=quasipoisson)
  redpred <- crossreduce(cbtmean, model, cen=mean(dataggrsum$tmean, na.rm=T))

  # STORE PARAMETERS (COEF + VECTORIZED VCOV)
  ncoef <- length(coef(redpred))
  par <- c(coef(redpred), vechMat(vcov(redpred)))
  names(par) <- c(paste0("coef", seq(ncoef)),
    paste0("vcov", seq(ncoef*(ncoef+1)/2)))
  tmeanpar <- data.frame(cities[i, c("city", "cityname", "state", "statename")],
    t(par), row.names=i)
  tmeanparlist[[i]] <- tmeanpar

#################################################################################
# ANALYSIS OF TEMPERATURE - ALL-CAUSE MORTALITY (SUMMER-ONLY, BY PERIOD)

  # DEFINE THE PERIODS
  yearlist <- list(1987:1989, 1990:1992, 1993:1995, 1996:1998, 1999:2000)

  # PERFORM MODEL BY PERIOD
  parlist <- lapply(yearlist, function(ysub) {
    model <- glm(all ~ cbtmean + dow + ns(yday(date), df=4)*factor(year(date)),
      data=dataggrsum, family=quasipoisson, subset=year(date) %in% ysub)
    redpred <- crossreduce(cbtmean, model, cen=mean(dataggrsum$tmean, na.rm=T))
    t(c(coef(redpred), vechMat(vcov(redpred))))
  })

  # STORE PARAMETERS (COEF + VECTORIZED VCOV)
  par <- do.call(rbind, parlist)
  colnames(par) <- names(tmeanpar)[-c(1:4)]
  tmeanperpar <- data.frame(
    cities[i, c("city", "cityname", "state", "statename")],
    period = sapply(yearlist, function(x) paste(range(x), collapse="-")),
    year = sapply(yearlist, mean),
    par,
    row.names=paste0(i, ".", seq(yearlist))
  )
  tmeanperparlist[[i]] <- tmeanperpar

#################################################################################
# ANALYSIS OF TEMPERATURE - ALL-CAUSE MORTALITY (SUMMER-ONLY, BY AGE)

  # DEFINE AGE GROUPS
```

```
  agecat <- as.character(unique(data$agecat))

  # PERFORM MODEL BY AGE (SELECTING FROM NON-COLLAPSED DATA)
  parlist <- lapply(agecat, function(cat) {
    y <- data$all[data$agecat==cat & month(data$date) %in% 6:9]
    model <- glm(y ~ cbtmean + dow + ns(yday(date), df=4)*factor(year(date)),
      data=dataggrsum, family=quasipoisson)
    redpred <- crossreduce(cbtmean, model, cen=mean(dataggrsum$tmean, na.rm=T))
    t(c(coef(redpred), vechMat(vcov(redpred))))
  })

  # STORE PARAMETERS (COEF + VCOV FOR 10-UNIT INCREASE)
  par <- do.call(rbind, parlist)
  colnames(par) <- names(tmeanpar)[-c(1:4)]
  tmeanagepar <- data.frame(
    cities[i, c("city", "cityname", "state", "statename")],
    agecat=agecat,
    par,
    row.names=paste0(i, ".", seq(agecat))
  )
  tmeanageparlist[[i]] <- tmeanagepar

##############################################################################
# ANALYSIS OF OZONE - NON-EXTERNAL MORTALITY (FULL YEAR)

  # DEFINE MOVING AVERAGE OF OZONE AT LAG 0-1
  o301 <- runMean(dataggr$o3, 0:1)

  # DEFINE CROSS-BASIS FOR TEMPERATURE
  cbtmean <- crossbasis(dataggr$tmean, lag=3, argvar=list(fun="bs", degree=2,
    knots=quantile(dataggrsum$tmean, c(10,50,90)/100, na.rm=T)),
    arglag=list(fun="strata"))

  # RUN THE MODEL AND EXTRACT PAR (ONLY IF ENOUGH NON-MISSING)
  par <- if(nrow(na.omit(cbind(o301, cbtmean))) < 500 ) c(NA,NA) else {
    model <- glm(death ~ o301 + cbtmean + dow + ns(date, df=14*7),
      data=dataggr, family=quasipoisson)
    c(coef(model)["o301"]*10, vcov(model)["o301","o301"]*10)
  }
  names(par) <- c("coef", "vcov")

  # STORE PARAMETERS (COEF + VCOV FOR 10-UNIT INCREASE)
  o3par <- data.frame(cities[i, c("city", "cityname", "state", "statename")],
    t(par), row.names=i)
  o3parlist[[i]] <- o3par

##############################################################################
# TEMPERATURE DISTRIBUTION (SUMMER ONLY)

  # DEFINE PERCENTILES
  per <- c(0:9/10, 1:99, 991:1000/10)/100
  tmeansumlist[[i]] <- quantile(dataggrsum$tmean, per, na.rm=T)
}

##############################################################################
# PREPARE AND STORE

# RBIND COEF/VCOV TOGETHER IN DATAFRAMES
tmeanpar <- do.call(rbind, tmeanparlist)
tmeanperpar <- do.call(rbind, tmeanperparlist)
tmeanagepar <- do.call(rbind, tmeanageparlist)
o3par <- do.call(rbind, o3parlist)
```

```
# CREATE COUNTRY-AVERAGE SUMMER TEMPERATURE DISTRIBUTION
avgtmeansum <- data.frame(perc=names(tmeansumlist[[1]]),
  tmean=apply(do.call(cbind, tmeansumlist), 1, mean))

# STORE
write.csv(tmeanpar, file="data/tmeanpar.csv", row.names=F)
write.csv(tmeanperpar, file="data/tmeanperpar.csv", row.names=F)
write.csv(tmeanagepar, file="data/tmeanagepar.csv", row.names=F)
write.csv(o3par, file="data/o3par.csv", row.names=F)
write.csv(avgtmeansum, file="data/avgtmeansum.csv", row.names=F)
```

```
# 01.heterogeneity.R

################################################################################
# R code for the analysis in:
#
#  Sera F, Gasparrini A. Extended two-stage designs for environmental research.
#    Environmental Health. 2022;21:41.
#  https://doi.org/10.1186/s12940-022-00853-z
#
# * an updated version of this code, compatible with future versions of the
#   software, is available at:
#   https://github.com/gasparrini/Extended2stage
################################################################################

################################################################################
# POOLING COMPLEX MULTI-PARAMETER ASSOCIATIONS
################################################################################

# LOAD PACKAGES
library(mixmeta) ; library(dlnm) ; library(scales)

# LOAD COEF/VCOV FROM FIRST-STAGE MODELS
tmeanpar <- read.csv(file="data/tmeanpar.csv")
coef <- as.matrix(tmeanpar[,grep("coef", names(tmeanpar))])
vcov <- as.matrix(tmeanpar[,grep("vcov", names(tmeanpar))])

# LINK WITH CENSUS DATA
cityind <- tmeanpar[,1:4,]
citycensus <- read.csv("data/citycensus.csv")
cityind <- merge(cityind, citycensus, by="city")

################################################################################
# RUN THE MODELS

# MODEL WITH NO META-PREDICTOR
model0 <- mixmeta(coef~1, vcov, data=cityind, method="ml")

# SUMMARY AND HETEROGENEITY TEST
summary(model0)
qtest(model0)

# MODELS WITH A SINGLE META-PREDICTOR
# PREDICTORS: TOTAL POP, % OF PEOPLE WITH HIGH-SCHOOL DEGREE, % OF UNEMPLOYED
model1 <- update(model0, .~pop100)
model2 <- update(model0, .~Phigh)
model3 <- update(model0, .~Punem)

# FULL MODEL
model4 <- update(model0, .~pop100+Phigh+Punem)
summary(model4)

# MODEL COMPARISON AND TESTS
AIC(model0, model1, model2, model3, model4)
BIC(model0, model1, model2, model3, model4)
drop1(model4, test="Chisq")

# MODEL SELECTION (STEP FORWARD)
step(model0, .~pop100+Phigh+Punem)

################################################################################
# PLOT THE AVERAGE EXPOSURE-RESPONSE RELATIONSHIPS
```

```
# LOAD AVERAGE TEMPERATURE DISTRIBUTION ACROSS CITIES
avgtmeansum <- read.csv("data/avgtmeansum.csv")
tmean <- avgtmeansum$tmean

# DEFINE SPLINE TRANSFORMATION ORIGINALLY USED IN FIRST-STAGE MODELS
knots <- tmean[avgtmeansum$perc %in% paste0(c(50,90), ".0%")]
bvar <- onebasis(tmean, fun="bs", degree=2, knots=knots)

# DEFINE THE CENTERING POINT (AT POINT OF MINIMUM RISK)
cen <- tmean[which.min(bvar%*%coef(model0))]

# PREDICT THE ASSOCIATION
cp <- crosspred(bvar, coef=coef(model0), vcov=vcov(model0), model.link="log",
  at=tmean, cen=cen)

# PLOTTING LABELS
xperc <- c(0,1,5,25,50,75,95,99,100)
xval <- tmean[avgtmeansum$perc %in% paste0(xperc, ".0%")]

# PLOT
plot(cp, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, xaxt="n", mgp=c(2.5,1,0), cex.axis=0.8,
  main="Temperature-mortality association")
axis(1, at=xval, labels=paste0(xperc, "%"))
abline(v=cen, lty=2, col=grey(0.8))

################################################################################
# PREDICT AND PLOT FOR GIVEN VALUES OF META-PREDICTORS

# DEFINE THE VALUES (SMALL/LARGE POP, WITH AVERAGE OF OTHERS)
datapred <- data.frame(
  pop100 = quantile(cityind$pop100, c(5,95)/100),
  Phigh = mean(cityind$Phigh),
  Punem = mean(cityind$Punem),
  row.names=c("Small population", "Large population")
)

# PREDICT COEF/VCOV
pred <- predict(model4, datapred, vcov=T)

# PREDICT ASSOCIATIONS
cp1 <- crosspred(bvar, coef=pred[[1]]$fit, vcov=pred[[1]]$vcov,
  model.link="log", at=tmean, cen=cen)
cp2 <- crosspred(bvar,coef=pred[[2]]$fit,vcov=pred[[2]]$vcov,
  model.link="log", at=tmean, cen=cen)

# PLOT
plot(cp1, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), cex.axis=0.8, col=3,
  ci.arg=list(col=alpha(3,0.3)), main="Temperature-mortality association")
axis(1, at=xval, labels=paste0(xperc, "%"), cex.axis=0.8)
lines(cp2, lwd=1.5, col=4, ci="area", ci.arg=list(col=alpha(4,0.3)))
abline(v=cen, lty=2, col=grey(0.8))
mtext("By population size")
legend("topleft", c(rownames(datapred)), lwd=1.5, col=c(3,4), bty="n",
inset=0.1)

################################################################################
# SAVE ARTICLE-STYLE PLOT

# GRAPHICALS PARAMETERS
layout(t(1:2))
```

```
oldpar <- par(no.readonly = TRUE)
par(mar=c(4,4,2,0.5), cex.axis=0.8)

# PLOTS
plot(cp, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, xaxt="n", mgp=c(2.5,1,0))
axis(1, at=xval, labels=paste0(xperc, "%"))
abline(v=cen, lty=2, col=grey(0.8))
title("Average relationship")

plot(cp1, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), col=3,
  ci.arg=list(col=alpha(3,0.3)))
axis(1, at=xval, labels=paste0(xperc, "%"))
lines(cp2, lwd=1.5, col=4, ci="area", ci.arg=list(col=alpha(4,0.3)))
abline(v=cen, lty=2, col=grey(0.8))
legend("topleft", c(rownames(datapred)), lwd=1.5, col=c(3,4), bty="n",
inset=0.1)
title("By population size")

# RESET
par(oldpar)
layout(1)

# PRINT
dev.print(pdf, file="figures/multipar.pdf", height=5, width=12)
```

```
# 02.multilevel.R

################################################################################
# R code for the analysis in:
#
#  Sera F, Gasparrini A. Extended two-stage designs for environmental research.
#    Environmental Health. 2022;21:41.
#  https://doi.org/10.1186/s12940-022-00853-z
#
# * an updated version of this code, compatible with future versions of the
#   software, is available at:
#   https://github.com/gasparrini/Extended2stage
################################################################################

################################################################################
# MULTILEVEL ANALYSIS BY CITY AND STATE
################################################################################

# LOAD THE PACKAGES
library(mixmeta) ; library(Epi) ; library(maps) ; library(scales)
library(ggplot2)

# LOAD COEF/VCOV FROM FIRST-STAGE MODELS (FOR OZONE)
o3par <- read.csv(file="data/o3par.csv")
coef <- o3par[,grep("coef", names(o3par))]
vcov <- o3par[,grep("vcov", names(o3par))]

# LOAD CITY-LEVEL METADATA AND LINK WITH LAT/LONG DATA
cityinfo <- o3par[,1:4,]
latlong <- read.csv("data/latlong.csv")
cityinfo <- merge(cityinfo, latlong, by="city")

################################################################################
# RUN MODELS

# STANDARD META-ANALYSIS (SINGLE RANDOM-EFFECTS LEVEL)
model0 <- mixmeta(coef, vcov)
summary(model0)

# ADD STATE AS FIXED EFFECTS (PARAMETERIZED TO SHOW STATE-SPECIFIC EFFECTS)
model1 <- mixmeta(coef~0+state, vcov, data=cityinfo)
summary(model1)

# TWO-LEVEL RANDOM EFFECTS (CITY NESTED WITHIN STATE)
model2 <- mixmeta(coef, vcov, data=cityinfo, random=~1|state/city)
summary(model2)

# TEST HETEROGENEITY
qtest(model0)
qtest(model1)
qtest(model2)

# SIGNIFICANCE TEST OF BETWEEN-STATE DIFFERENCES
# NB: REQUIRES STANDARD PARAMETERIZATION AND USE OF ML ESTIMATOR FOR COMPARISON
drop1(update(model1, .~state, method="ml"), test="Chisq")

# ESTIMATED EFFECTS IN RR SCALE FOR 10-UNIT INCREASE
# - POOLED EFFECTS FOR MODELS WITH NO FIXED EFFECTS (WITH CI)
# - RANGE OF EFFECTS FOR MODEL WITH FIXED EFFECTS
ci.exp(model0)
ci.exp(model2)
range(exp(predict(model1)))
```

```
################################################################################
# PREDICTIONS

# FIXED-EFFECTS PREDICTION OF STATE-AVERAGE EFFECTS
statefix <- exp(unique(predict(model1, ci=T)))

# BLUPS AT CITY AND STATE LEVEL FROM TWO-LEVEL MODEL
cityblup <- exp(blup(model2, pi=T))
stateblup <- exp(unique(blup(model2, pi=T, level=1)))

# NAMES
rownames(cityblup) <- cityinfo$cityname[!is.na(coef)]
rownames(stateblup) <- rownames(statefix) <-
  unique(cityinfo$statename[!is.na(coef)])

################################################################################
# MAP OF CITY-SPECIFIC BLUPS

# EFFECTS, COLOURS, COORDINATES (NB: REVERSE LONG)
cutoff <- pretty(cityblup[,1], 8)
labmap <- paste(format(cutoff)[-length(cutoff)], format(cutoff)[-1], sep="-")
rrcat <- cut(cityblup[,1], cutoff,  labels=labmap)
colmap <- colorRampPalette(c("yellow","red"))(length(cutoff))
lat <- as.numeric(as.character(cityinfo$lat[!is.na(coef)]))
long <- -as.numeric(as.character(cityinfo$long[!is.na(coef)]))

# MAP OF CITY-SPECIFIC BLUPS (VERSION 1)
map("state", interior=F)
map("state", lty=2, add=T)
points(long, lat, col=alpha(colmap[rrcat], 0.6), pch=19, cex=1.8)
legend("bottomright", labmap, pch=19, col=colmap, bty="n", pt.cex=1.5, cex=0.8,
  title="RR", inset=0.02)
title("Map of risk associated to ozone")
mtext("From a multi-level meta-analysis")

# MAP OF CITY-SPECIFIC BLUPS (VERSION 2)
mapstate <- map_data("state")
ggplot(mapstate, aes(long, lat, group=group)) +
  geom_polygon(fill=NA, col="grey") +
  borders("usa", col="black") +
  geom_point(aes(group=NULL, col=rrcat), alpha=0.6, size=5,
    data=data.frame(long=long,lat=lat)) +
  scale_color_brewer(name="RR", palette="YlOrRd") +
  xlim(min(mapstate$long), max(mapstate$long)) +
  ylim(min(mapstate$lat), max(mapstate$lat)) +
  theme_void() +
  theme(legend.position=c(1,0), legend.justification=c(1.2,-0.1)) +
  labs(title="Map of risk associated to ozone",
     subtitle="From a multi-level meta-analysis") +
  coord_quickmap()

################################################################################
# FOREST PLOT

# SEQUENCES AND LABELS
yseq <- seq(nrow(stateblup))
ylab <- rownames(stateblup)

# FOREST PLOT (VERSION 1)
par(mar=c(5,8.2,4,2)+0.1)
plot(yseq, type="n", xlim=c(0.98,1.02), yaxt="n", xlab="RR", ylab="",
```

```
  mgp=c(2.5,1,0), main="Comparison of state-specific risk estimates")
axis(2, at=yseq, las=1, tick=F, labels=ylab, cex.axis=0.8)
grid()
abline(v=1)
arrows(stateblup[,2], yseq-0.2, stateblup[,3], yseq-0.2, col=alpha(2,0.5),
  code=3, angle=90, length=0.01, lwd=2)
points(stateblup[,1], yseq-0.2, col=2, pch=19, cex=1.2)
arrows(statefix[,2], yseq+0.2, statefix[,3], yseq+0.2, col=alpha(4,0.5),
  code=3, angle=90, length=0.01, lwd=2)
points(statefix[,1], yseq+0.2, col=4, pch=19, cex=1.2)
legend("topright", c("Fixed-effects","BLUPs"), pch=19, col=c(4,2), pt.cex=1.2,
  bty="n")
par(mar=c(5,4,4,1)+0.1)

# CREATE DATAFRAME
eststate <- data.frame(rbind(statefix, stateblup), row.names=NULL)
eststate$state <- rep(rownames(statefix), 2)
eststate$model <- rep(c("Fixed-effects","BLUPs"), each=nrow(statefix))
eststate$model <- factor(eststate$model, levels=unique(eststate$model))

# FOREST PLOT (VERSION 2)
ggplot(eststate, aes(fit, state, col=model)) +
  geom_vline(xintercept=1) +
  geom_errorbar(aes(xmin=ci.lb, xmax=ci.ub), width=0.6, alpha=0.5,
    position=position_dodge(width=0.75)) +
  geom_point(size=2, position=position_dodge(width=0.75)) +
  scale_color_manual(values=c(4,2), name="Prediction")+
  theme_bw() +
  scale_y_discrete(limits=unique(eststate$state)) +
  coord_cartesian(xlim=c(0.98,1.02)) +
  theme(legend.position="top") +
  labs(title="Comparison of state-specific risk estimates") +
  xlab("RR") + ylab("")

################################################################################
# SAVE ARTICLE-STYLE PLOTS

# MAP OF CITY-SPECIFIC BLUPS (VERSION 2)
ggplot(mapstate, aes(long, lat, group=group)) +
  geom_polygon(fill=NA, col="grey") +
  borders("usa", col="black") +
  geom_point(aes(group=NULL, col=rrcat), alpha=0.6, size=5,
    data=data.frame(long=long,lat=lat)) +
  scale_color_brewer(name="RR", palette="YlOrRd") +
  xlim(min(mapstate$long), max(mapstate$long)) +
  ylim(min(mapstate$lat), max(mapstate$lat)) +
  theme_void() +
  theme(legend.position=c(1,0), legend.justification=c(1.2,-0.1)) +
  coord_quickmap()
ggsave("figures/multilevmap.pdf", height=5, width=9)

# FOREST PLOT (VERSION 2)
ggplot(eststate, aes(fit, state, col=model)) +
  geom_vline(xintercept=1) +
  geom_errorbar(aes(xmin=ci.lb, xmax=ci.ub), width=0.6, alpha=0.5,
    position=position_dodge(width=0.75)) +
  geom_point(size=2, position=position_dodge(width=0.75)) +
  scale_color_manual(values=c(4,2), name="Prediction")+
  theme_bw() +
  scale_y_discrete(limits=unique(eststate$state)) +
  coord_cartesian(xlim=c(0.98,1.02)) +
  theme(legend.position="top") +
```

```
  xlab("RR") + ylab("")
ggsave("figures/multilevforest.pdf", height=9, width=5)
```

```
# 03.doseresp.R

#############################################################################
#######
# R code for the analysis in:
#
#  Sera F, Gasparrini A. Extended two-stage designs for environmental research.
#    Environmental Health. 2022;21:41.
#  https://doi.org/10.1186/s12940-022-00853-z
#
# * an updated version of this code, compatible with future versions of the
#   software, is available at:
#   https://github.com/gasparrini/Extended2stage
#############################################################################

#############################################################################
# MODELLING REPEATED MEASURES AND DOSE-RESPONSE RELATIONSHIPS
#############################################################################

# LOAD PACKAGES
library(mixmeta) ; library(dlnm) ; library(splines) ; library(scales)

# LOAD COEF/VCOV FROM FIRST-STAGE MODELS
tmeanagepar <- read.csv(file="data/tmeanagepar.csv")
coef <- as.matrix(tmeanagepar[,grep("coef", names(tmeanagepar))])
vcov <- as.matrix(tmeanagepar[,grep("vcov", names(tmeanagepar))])

# CITY-SPECIFIC META-DATA
cityinfo <- tmeanagepar[,1:5,]

#############################################################################
# RUN MODELS

# MODEL IGNORING CLUSTERING
model0 <- mixmeta(coef~agecat, vcov, data=cityinfo, method="ml")
summary(model0)

# MODEL CONSIDERING CLUSTERING
model1 <- update(model0, random=~1|city)
summary(model1)

# DEFINE AGE AS CONTINUOUS VARIABLE
cityinfo$age <- c(60,70,85)[factor(cityinfo$agecat,
c("under65","65to74","75p"))]

# MODEL WITH LINEAR EFFECT OF AGE
model2 <- update(model1, .~age)
summary(model2)

# MODEL WITH NON-LINEAR EFFECT OF AGE
model3 <- update(model1, .~ns(age, df=2))
summary(model3)

# MODEL COMPARISON AND TESTS
AIC(model0, model1, model2, model3)
BIC(model0, model1, model2, model3)
drop1(model0, test="Chisq")
drop1(model1, test="Chisq")
drop1(model2, test="Chisq")
drop1(model3, test="Chisq")
```

```
################################################################################
# PREDICT AND PLOT FOR GIVEN AGES

# DEFINE THE AGE VALUES
datapred <- data.frame(age=6:9*10)

# PREDICT COEF/VCOV
pred <- predict(model3, datapred, vcov=T)

# LOAD AVERAGE TEMPERATURE DISTRIBUTION ACROSS CITIES
avgtmeansum <- read.csv("data/avgtmeansum.csv")
tmean <- avgtmeansum$tmean

# DEFINE SPLINE TRANSFORMATION ORIGINALLY USED IN FIRST-STAGE MODELS
knots <- tmean[avgtmeansum$perc %in% paste0(c(50,90), ".0%")]
bvar <- onebasis(tmean, fun="bs", degree=2, knots=knots)

# PREDICT FOR 80-YEAR-OLD
cen <- tmean[which.min((bvar%*%pred[[3]]$fit))]
cp <- crosspred(bvar, coef=pred[[3]]$fit, vcov=pred[[3]]$vcov,
  model.link="log", at=tmean, cen=cen)

# PLOTTING LABELS
xperc <- c(0,1,5,25,50,75,90,99,100)
xval <- tmean[avgtmeansum$perc %in% paste0(xperc, ".0%")]

# PLOT
plot(cp, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), cex.axis=0.8, col=3,
  ci="n", main="Temperature-mortality association")
axis(1, at=xval, labels=paste0(xperc, "%"), cex.axis=0.8)
abline(v=cen, lty=2, col=grey(0.8))

# ADD PREDICTIONS FOR OTHER AGES
for(i in c(1,2,4)) {
  cp <- crosspred(bvar, coef=pred[[i]]$fit, vcov=pred[[i]]$vcov,
    model.link="log", at=tmean, cen=cen)
  lines(cp, lwd=1.5, col=i)
}
legend("top", as.character(datapred$age), lwd=1.5, col=1:4, bty="n", inset=0.1,
  title="Age", ncol=4)
mtext("Age patterns")

################################################################################
# SAVE ARTICLE-STYLE PLOT

# GRAPHICALS PARAMETERS
oldpar <- par(no.readonly = TRUE)
par(mar=c(4,4,1,0.5), cex.axis=0.8)

# PREDICT FOR 80-YEAR-OLD
cp <- crosspred(bvar, coef=pred[[3]]$fit, vcov=pred[[3]]$vcov,
  model.link="log", at=tmean, cen=cen)

# PLOT
plot(cp, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), col=3, ci="n")
axis(1, at=xval, labels=paste0(xperc, "%"))
abline(v=cen, lty=2, col=grey(0.8))

# ADD PREDICTIONS FOR OTHER AGES
for(i in c(1,2,4)) {
```

```
  cp <- crosspred(bvar, coef=pred[[i]]$fit, vcov=pred[[i]]$vcov,
    model.link="log", at=tmean, cen=cen)
  lines(cp, lwd=1.5, col=i)
}
legend("top", as.character(datapred$age), lwd=1.5, col=1:4, bty="n", inset=0.1,
  title="Age", ncol=4)

# RESET
par(oldpar)

# PRINT
dev.print(pdf, file="figures/dosresp.pdf", height=4.5, width=6)
```

```
# 04.longitudinal.R

################################################################################
# R code for the analysis in:
#
#  Sera F, Gasparrini A. Extended two-stage designs for environmental research.
#    Environmental Health. 2022;21:41.
#  https://doi.org/10.1186/s12940-022-00853-z
#
# * an updated version of this code, compatible with future versions of the
#   software, is available at:
#   https://github.com/gasparrini/Extended2stage
################################################################################

################################################################################
# LONGITUDINAL ANALYSIS OF EFFECT MODIFICATION BY AIR CONDITIONING
################################################################################

# LOAD PACKAGES
library(mixmeta) ; library(dlnm) ; library(splines)
library(nlme) ; library(scales)

# LOAD FIRST-STAGE DATA
tmeanperpar <- read.csv(file="data/tmeanperpar.csv")

# LOAD AIR CONDITIONING DATA, RESHAPE TO LONG, REMOVE MISSING
acdata <- read.csv(file="data/acdata.csv")
acdata <- reshape(acdata, varying=seq(ncol(acdata))[-1], idvar="city", sep="",
  timevar="year", direction="long")
acdata <- na.omit(acdata[with(acdata, order(city, year)),])

# SUBSET ORIGINAL DATA TO CITIES WITH AC MEASURES
tmeanperpar <- tmeanperpar[tmeanperpar$city %in% acdata$city,]

# ESTRACT COEF/VCOV FROM FIRST-STAGE MODELS
coef <- as.matrix(tmeanperpar[,grep("coef", names(tmeanperpar))])
vcov <- as.matrix(tmeanperpar[,grep("vcov", names(tmeanperpar))])

# CITY-SPECIFIC META-DATA
cityinfo <- tmeanperpar[,1:6,]

# PERFORM RANDOM-EFFECTS MODEL TO OBTAIN SMOOTH FIT OF AC TREND
mlme <- lme(ac ~ bs(year, degree=2, df=4), data=acdata,
  random=list(city=pdDiag(~ bs(year, degree=2, df=4))))

# PREDICT AC FOR EACH PERIOD IN EACH CITY
cityinfo$ac <- predict(mlme, cityinfo)
summary(cityinfo$ac)
# PROBLEM: ABOVE 100%

################################################################################

# MODEL WITH NO META-PREDICTOR
model0 <- mixmeta(coef, vcov, data=cityinfo, method="ml",
  random=~1|city, bscov="diag")

# MODEL WITH AC AND TIME (TAKES SOME TIME)
# NB: ADD control=list(showiter=TRUE) TO INSPECT THE ITERATIVE OPTIMIZATION
model1 <- mixmeta(coef ~ac+ns(year, knots=1995), vcov, data=cityinfo,
  method="ml", random=~ns(year, knots=1995)|city, bscov="diag")

# TEST (TAKES EVEN LONGER)
```

```
drop1(model1, test="Chisq")

################################################################################
# PLOT THE AVERAGE EXPOSURE-RESPONSE AT LOW HIGH AC IN YEAR 2000

# LOAD AVERAGE TEMPERATURE DISTRIBUTION ACROSS CITIES
avgtmeansum <- read.csv("data/avgtmeansum.csv")
tmean <- avgtmeansum$tmean

# DEFINE SPLINE TRANSFORMATION ORIGINALLY USED IN FIRST-STAGE MODELS
knots <- tmean[avgtmeansum$perc %in% paste0(c(50,90), ".0%")]
bvar <- onebasis(tmean, fun="bs", degree=2, knots=knots)

# DEFINE THE CENTERING POINT (AT POINT OF MINIMUM RISK)
cen <- tmean[which.min(bvar%*%coef(model0))]

# PLOTTING LABELS
xperc <- c(0,1,5,25,50,75,90,99,100)
xval <- tmean[avgtmeansum$perc %in% paste0(xperc, ".0%")]

# DEFINE THE VALUES (AC AT 20-80%, YEAR 2000)
datapred <- data.frame(ac=c("Low AC use"=20,"High AC use"=80), year=2000)

# PREDICT COEF/VCOV
pred <- predict(model1, datapred, vcov=T)

# PREDICT ASSOCIATIONS
cp1 <- crosspred(bvar, coef=pred[[1]]$fit, vcov=pred[[1]]$vcov,
  model.link="log", at=tmean, cen=cen)
cp2 <- crosspred(bvar,coef=pred[[2]]$fit,vcov=pred[[2]]$vcov,
  model.link="log", at=tmean, cen=cen)

# EXPOSURE-RESPONSE PLOT
plot(cp1, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), cex.axis=0.8, col=3,
  ci.arg=list(col=alpha(3,0.3)), main="Temperature-mortality association")
axis(1, at=xval, labels=paste0(xperc, "%"), cex.axis=0.8)
lines(cp2, lwd=1.5, col=4, ci="area", ci.arg=list(col=alpha(4,0.3)))
abline(v=cen, lty=2, col=grey(0.8))
mtext("By AC prevalence")
legend("topleft", c(rownames(datapred)), lwd=1.5, col=c(3,4), bty="n",
inset=0.1)

################################################################################
# PLOT THE TREND IN RR AT 99TH TMEAN PERC UNDER DIFFERENT AC SCENARIOS

# PREDICT TREND IN AVERAGE AC USE (NB: USE level ARGUMENT IN predict)
acpred <- predict(mlme, data.frame(year=1987:2000), level=0)

# AVERAGE TMEAN AT 99TH PERCENTILE
tmean99 <- tmean[avgtmeansum$perc=="99.0%"]

# DEFINE THE SCENARIOS: CONSTANT AC IN 1987 AND ACTUAL AC TREND
datapred1 <- data.frame(ac=acpred[1], year=1987:2000)
datapred2 <- data.frame(ac=acpred, year=1987:2000)

# PREDICT COEF/VCOV
pred1 <- predict(model1, datapred1, vcov=T)
pred2 <- predict(model1, datapred2, vcov=T)

# PREDICT ASSOCIATIONS
rr99 <- t(sapply(seq(nrow(datapred1)), function(i) {
```

```
  # PREDICT ASSOCIATIONS AT 99TH PERCENTILE
  cp1 <- crosspred(bvar, coef=pred1[[i]]$fit, vcov=pred1[[i]]$vcov,
    model.link="log", at=tmean99, cen=cen)
  cp2 <- crosspred(bvar,coef=pred2[[i]]$fit,vcov=pred2[[i]]$vcov,
    model.link="log", at=tmean99, cen=cen)

  # EXTRACT RR ANC CI
  est <- c(with(cp1, c(allRRfit, allRRlow, allRRhigh)),
    with(cp2, c(allRRfit, allRRlow, allRRhigh)))
  names(est) <- c(t(outer(c("rr1","rr2"), c("","low","high"), paste0)))

  # RETURN
  est
}))

# PLOT
plot(1987:2000, seq(1987:2000), type="n", ylim=range(rr99)*c(0.93,1.07),
  ylab="RR", xlab="Year", bty="l", las=1, mgp=c(2.5,1,0), cex.axis=0.8,
  main="Trend in risk")
arrows(1987:2000-0.1, rr99[,2], 1987:2000-0.1, rr99[,3], col=alpha(4,0.5),
  code=3, angle=90, length=0.05, lwd=2)
points(1987:2000-0.1, rr99[,1],  type="o", col=4, pch=19)
arrows(1987:2000+0.1, rr99[,5], 1987:2000+0.1, rr99[,6], col=alpha(2,0.5),
  code=3, angle=90, length=0.05, lwd=2)
points(1987:2000+0.1, rr99[,4], type="o", col=2, pch=19)
abline(h=1)
mtext("By scenario of trends in AC prevalence")
legend("top", c("Constant at 1987","Actual trend"), pch=19, col=c(4,2), bty="n",
  inset=0.1)

################################################################################
# SAVE ARTICLE-STYLE PLOT

# GRAPHICALS PARAMETERS
layout(t(1:2))
oldpar <- par(no.readonly = TRUE)
par(mar=c(4,4,2,0.5), cex.axis=0.8)

# PLOTS
plot(cp1, ylim=c(0.9,1.4), xlab="Temperature percentile", ylab="RR",
  lab=c(6,5,7), las=1, lwd=1.5, xaxt="n", mgp=c(2.5,1,0), col=3,
  ci.arg=list(col=alpha(3,0.3)), main="Temperature-mortality association")
axis(1, at=xval, labels=paste0(xperc, "%"))
lines(cp2, lwd=1.5, col=4, ci="area", ci.arg=list(col=alpha(4,0.3)))
abline(v=cen, lty=2, col=grey(0.8))
legend("topleft", c(rownames(datapred)), lwd=1.5, col=c(3,4), bty="n",
  inset=0.1, title="AC prevalence")

plot(1987:2000, seq(1987:2000), type="n", ylim=range(rr99)*c(0.93,1.07),
  ylab="RR", xlab="Year", bty="l", las=1, mgp=c(2.5,1,0), main="Trend in risk")
arrows(1987:2000-0.1, rr99[,2], 1987:2000-0.1, rr99[,3], col=alpha(4,0.5),
  code=3, angle=90, length=0.05, lwd=2)
points(1987:2000-0.1, rr99[,1],  type="o", col=4, pch=19)
arrows(1987:2000+0.1, rr99[,5], 1987:2000+0.1, rr99[,6], col=alpha(2,0.5),
  code=3, angle=90, length=0.05, lwd=2)
points(1987:2000+0.1, rr99[,4], type="o", col=2, pch=19)
abline(h=1)
legend("top", c("Constant at 1987","Actual trend"), pch=19, col=c(4,2), bty="n",
  inset=0.1, title="Scenario of AC prevalence")

# RESET
```

```
par(oldpar)
layout(1)

# PRINT
dev.print(pdf, file="figures/longitudinal.pdf", height=5, width=12)
```

Chapter 5

Research paper III

Title: How urban characteristics affect vulnerability to heat and cold: a multi-country analysis.

Author(s): Francesco Sera, Ben Armstrong, Aurelio Tobias, Ana Maria Vicedo-Cabrera, Christofer Åström, Michelle L. Bell, Bing-Yu Chen, Micheline de Sousa Zanotti Stagliorio Coelho, Patricia Matus Correa, Julio Cesar Cruz, Tran Ngoc Dang, Magali Hurtado-Diaz, Dung Do Van, Bertil Forsberg, Yue Leon Guo1, Yuming Guo, Masahiro Hashizume, Yasushi Honda, Carmen Iñiguez, Jouni J. K. Jaakkola, Haidong Kan, Ho Kim, Eric Lavigne, Paola Michelozzi, Nicolas Valdes Ortega, Samuel Osorio, Mathilde Pascal, Martina S. Ragettli, Niilo R. I. Ryti, Paulo Hilario Nascimento Saldiva, Joel Schwartz, Matteo Scortichini, Xerxes Seposo, Shilu Tong, Antonella Zanobetti, Antonio Gasparrini.

Journal/Publisher: International Journal of Epidemiology.

Type of publication: Research paper.

Stage of publication: Published in Volume 48, Issue 3, August 2019, Pages 1101–1112.

URL: `http://https://doi.org/10.1093/ije/dyz008`.

Academic peer-reviewed: Yes.

Copyright: Permission obtained from the publisher.

Candidate's role: See Section 2.3.

Senior author: (Prof. Antonio Gasparrini)

International Journal of Epidemiology, 2019, 1101–1112
doi: 10.1093/ije/dyz008
Advance Access Publication Date: 27 February 2019
Original article

Environment, Green Space and Pollution

How urban characteristics affect vulnerability to heat and cold: a multi-country analysis

Francesco Sera,[1]* Ben Armstrong,[1] Aurelio Tobias,[2] Ana Maria Vicedo-Cabrera,[1] Christofer Åström,[3] Michelle L Bell,[4] Bing-Yu Chen,[5] Micheline de Sousa Zanotti Stagliorio Coelho,[6] Patricia Matus Correa,[7] Julio Cesar Cruz,[8] Tran Ngoc Dang,[9,10] Magali Hurtado-Diaz,[8] Dung Do Van,[9] Bertil Forsberg,[3] Yue Leon Guo,[5,11] Yuming Guo,[12,13] Masahiro Hashizume,[14] Yasushi Honda,[15] Carmen Iñiguez,[16] Jouni JK Jaakkola,[17,18] Haidong Kan,[19] Ho Kim,[20] Eric Lavigne,[21,22] Paola Michelozzi,[23] Nicolas Valdes Ortega,[7] Samuel Osorio,[24] Mathilde Pascal,[25] Martina S Ragettli,[26,27] Niilo Rl Ryti,[17,18] Paulo Hilario Nascimento Saldiva,[6] Joel Schwartz,[28] Matteo Scortichini,[23] Xerxes Seposo,[29] Shilu Tong,[30,31,32] Antonella Zanobetti[28] and Antonio Gasparrini[1]

[1]Department of Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, London, UK, [2]Institute of Environmental Assessment and Water Research (IDAEA), Spanish Council for Scientific Research (CSIC), Barcelona, Spain, [3]Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden, [4]School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA, [5]National Institute of Environmental Health Science, National Health Research Institutes, Zhunan, Taiwan, [6]Institute of Advanced Studies, University of São Paulo, São Paulo, Brazil, [7]Department of Public Health, Universidad de los Andes, Santiago, Chile, [8]Department of Environmental Health, National Institute of Public Health, Cuernavaca, Morelos, Mexico, [9]Department of Environmental Health, Faculty of Public Health, University of Medicine and Pharmacy of Ho Chi Minh City, Ho Chi Minh City, Vietnam, [10]Institute of Research and Development, Duy Tan University, Da Nang, Vietnam, [11]Environmental and Occupational Medicine, National Taiwan University (NTU) and NTU Hospital, Taipei, Taiwan, [12]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia, [13]Climate, Air Quality Research Unit, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia, [14]Department of Pediatric Infectious Diseases, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan, [15]Faculty of Health and Sport Sciences, University of Tsukuba, Tsukuba, Japan, [16]Department of Statistics and Computational Research, Environmental Health Research Joint Reseaech Unit FISABIO-UV-UJI CIBERESP, University of València, València, Spain, [17]Medical Research Center Oulu (MRC Oulu), Oulu University Hospital and University of Oulu, Oulu, Finland, [18]Center for Environmental and Respiratory Health Research (CERH), University of Oulu, Oulu, Finland, [19]Department of Environmental Health, School of Public Health, Fudan University, Shanghai, China, [20]Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea, [21]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Canada, [22]Air Health Science Division, Health Canada, Ottawa, Canada, [23]Department of Epidemiology, Lazio

Regional Health Service, Rome, Italy, [24]Department of Environmental Health, University of São Paulo, São Paulo, Brazil, [25]Santé Publique France, Department of Environmental Health, French National Public Health Agency, Saint Maurice, France, [26]Swiss Tropical and Public Health Institute, Basel, Switzerland, [27]University of Basel, Basel, Switzerland, [28]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [29]Department of Environmental Engineering, Graduate School of Engineering, Kyoto University, Kyoto, Japan, [30]Shanghai Children's Medical Centre, Shanghai Jiao-Tong University, Shanghai, China, [31]School of Public Health and Institute of Environment and Human Health, Anhui Medical University, Hefei, China and [32]School of Public Health and Social Work, Queensland University of Technology, Brisbane, Australia

*Corresponding author. Department of Social and Environmental Health Research, London School of Hygiene & Tropical Medicine, 15–17 Tavistock Place, London, WC1H 9SH, UK. E-mail: francesco.sera@lshtm.ac.uk

Editorial decision 22 November 2018; Accepted 25 February 2019

Abstract

Background: The health burden associated with temperature is expected to increase due to a warming climate. Populations living in cities are likely to be particularly at risk, but the role of urban characteristics in modifying the direct effects of temperature on health is still unclear. In this contribution, we used a multi-country dataset to study effect modification of temperature–mortality relationships by a range of city-specific indicators.

Methods: We collected ambient temperature and mortality daily time-series data for 340 cities in 22 countries, in periods between 1985 and 2014. Standardized measures of demographic, socio-economic, infrastructural and environmental indicators were derived from the Organisation for Economic Co-operation and Development (OECD) Regional and Metropolitan Database. We used distributed lag non-linear and multivariate meta-regression models to estimate fractions of mortality attributable to heat and cold (AF%) in each city, and to evaluate the effect modification of each indicator across cities.

Results: Heat- and cold-related deaths amounted to 0.54% (95% confidence interval: 0.49 to 0.58%) and 6.05% (5.59 to 6.36%) of total deaths, respectively. Several city indicators modify the effect of heat, with a higher mortality impact associated with increases in population density, fine particles ($PM_{2.5}$), gross domestic product (GDP) and Gini index (a measure of income inequality), whereas higher levels of green spaces were linked with a decreased effect of heat.

Conclusions: This represents the largest study to date assessing the effect modification of temperature–mortality relationships. Evidence from this study can inform public-health interventions and urban planning under various climate-change and urban-development scenarios.

Key words: Temperature, heat, mortality, epidemiology, cities, climate

Key Messages

- Urban populations may experience higher risks due to exposure to non-optimal temperature, particularly in a changing climate, but the role of urban characteristics in modifying such direct health effects is still unclear.
- This represents the largest study to date assessing the effect modification of temperature–mortality relationships, performed by comparing different cities across the world and using standardized indicators.
- The effects of heat on mortality are higher in cities characterized by a higher level of inequalities, higher air-pollution exposure, fewer green spaces and lower availability of health services.
- Evidence from this study can inform public-health interventions and urban planning under various climate-change and urban-development scenarios.

Introduction

Several studies have evaluated the relationship between ambient temperature and mortality, consistently reporting increased risks at high and low temperatures.[1–3] These risks are associated with a substantial health burden across populations living in different parts of the world, indicating that exposure to non-optimal temperature represents an important global contributor to excess mortality.[1–3] The situation is not likely to improve in the context of climate change, as the health burden associated with non-optimal temperature is projected to increase in a warming planet.[4] In addition, scenarios of socio-economic pathways suggest that future susceptibility is likely to increase with ageing populations, rapid urbanization and growing inequalities.[5]

Populations living in cities are particularly vulnerable to non-optimal temperature. The structure of urban areas could enhance temperature-related health risks through a combination of higher exposures (e.g. urban heat island effect) and higher vulnerability (e.g. population density and socio-economic differentials).[6,7] Evidence of this excess health burden, particularly during extreme events as in Chicago in 1995, Paris in 2003 and Moscow in 2010, has motivated the development of public-health measures to reduce preventable mortality and morbidity (e.g. Heat Health Watch Warning System). Several heat health watch warning systems (also called 'heat warning systems' or 'heat health warning systems') have been implemented in several countries (e.g. the USA, Italy, Germany, France, Spain, Portugal, the UK, Australia, Canada, South Korea and China), some of which attempt to target potentially vulnerable groups in urban communities.[8,9] In this context, identifying aspects that modify the susceptibility to the impacts of non-optimal temperatures can help improve health-protection programmes and contribute to the development of city-level mitigation and adaptation strategies, including urban planning and design.

A number of studies have contributed to this topic, investigating potential effect modifiers of temperature–mortality associations. In particular, some studies have adopted ecological study designs to assess community-level factors, such as urbanization, number of green areas or vegetative covering.[10–19] However, most of the published studies included homogeneous populations; only a few compared regions with different geographic and climatic conditions, and populations with highly variable socio-economic and demographic characteristics.

In this study, we used data from the Multi-City Multi-Country (MCC) collaborative network (http://mccstudy.lshtm.ac.uk/) to evaluate the role of cities' characteristics in modifying susceptibility to high and low temperatures. The MCC database includes time-series data for hundreds of cities in 22 countries and provides a unique opportunity to compare health effects across highly heterogeneous populations. Specifically, we linked the MCC data with standardized measures of contextual factors at the city level and analysed their effect modification for mortality risks associated with both heat and cold.

Methods

MCC data

The analysis is restricted to 340 cities or metropolitan areas (from now on generally referred to as cities) available in the MCC dataset, distributed across 22 countries. For each location, the dataset comprises time series of daily mean temperature and mortality counts for all causes or non-external causes only (International Classification of Diseases—ICD-9: 0–799; ICD-10: A00–R99) in largely overlapping periods ranging from 1 January 1985 to 31 December 2014. The full list of cities, together with additional information, can be found in Supplementary Material A and B, available as Supplementary data at *IJE* online.

Indicators

OECD Regional and Metropolitan Database

We collected data on several city-specific socio-economic indicators and urban development from the Organisation for Economic Co-operation and Development (OECD) Regional and Metropolitan Database.[20,21] The OECD Regional Database provides a set of comparable statistics and indicators on about 2000 regions and 281 OECD metropolitan areas in 34 OECD member countries and other economies (http://stats.oecd.org/Index.aspx). They include yearly time series, from 2000 to 2014, for around 40 indicators of demography, economic, labour-market, social, environmental and innovation themes. Details on the regional and metropolitan OECD Database can be found in a specific OECD publication.[22] OECD follows a strict Quality Framework for Statistical Activities.[23] The OECD quality framework defines two dimensions: the quality of national statistics that OECD receives and the quality of OECD internal processes for collection, processing, analysis and dissemination of data and metadata. OECD statistics have a high reputation for quality and integrity throughout the world and we are confident that the data we used have a high level of accuracy.

First, 136 cities in the MCC database were linked with the OECD Metropolitan Database at the metropolitan area (MA) level. In addition, all 340 MCC cities were

Table 1. OECD Regional and Metropolitan Database indicators included in the analysis: definition, years and geographical level of observation

Indicator	Definition	Years	MA	SR	LR
Demographic					
% population ≥65 years	% old population (65 years or more)	2000	136	147	37
Life expectancy (years)	Life expectancy at birth (years)	2005–06; 2010–11			288
Socio-economic					
GDP (US$)	GDP per capita (US$) (current prices, current PPP)	2001; 2010	136	59	130
Labour productivity (US$)	Labour productivity (Gross Value Added (GVA) per worker) (current prices, current PPP)	2005; 2009–10			280
Educational level (%)	Share of labour force with at least secondary-level education	2000			265
Unemployment rate (%)	Unemployment rate (%)	2001; 2010	136	130	41
Gini index	Gini (disposable income, after taxes and transfers); high index means high inequality	2009–14			280
Poverty gap	Poverty rate after taxes and transfers; the poverty line reflects 60% of the national median income	2009–14			280
Health system					
Hospital-bed rates	Hospital-bed rates (hospital beds per 10 000 population)	2008–10			279
Urban characteristics					
Type of surrounding region (rural/urban)	The OECD regional typology is based on the following criteria: population density, degree of rurality and size of the urban centres located within a region: Predominantly urban = 1 Intermediate = 2 Predominantly rural = 3 Predominantly rural close to a city = 4 Predominantly rural remote = 5	2000		272	
Urbanized area share (%)	Urbanized area share (%): share of the urbanized area over total land of a metropolitan area	2000–01; 2006	136		
Green area (square metres per million persons)	Land in the MA covered by vegetation, forest and parks in 2000 (source: MODIS MCD12Q1), divided by the population of the MA and then multiplied by 1 million	2000	136		
Concentration of population in the core (%)	Share of population living in the core areas over the total metropolitan population	2000	136		
Sprawl	The sprawl index measures the growth [over the periods 2000–06 and 2000–12, except Japan (1997–2006) and the USA (2001–06 and 2001–11)] in built-up areas adjusted for the growth in the city population	2006	100		

MA, city/metropolitan area; SR, small region; LR, large region.

linked with the OECD Regional Database both at small-region and large-region geographical levels. The former represents provinces or prefectures, the latter administrative regions or small states.[20] In total, a set of 14 indicators were selected from OECD Regional and Metropolitan Databases. These indicators encompass demographic, socio-economic, health-system and urban characteristics (Table 1). For each indicator, we used the data collected at the smallest geographical level available, using the value measured in a single year or averaged across multiple years in order to minimize the amount of missing data. The definition of each OECD indicator considered in this analysis is provided in Table 1.

The set of indicators related to urbanization (e.g. urbanized area, green area, concentration of population in the core, Sprawl index) is available for 136 MCC cities that are in the OECD MA Database. A subset of socio-economic indicators (e.g. Gini index, educational level) is available for OECD country members, but not for OECD country partners (e.g. China, Brazil, Colombia, Iran, Moldova, Philippines, Viet Nam). Other indicators (e.g. GDP, % population ≥65 years, unemployment rate) were available also among some OECD country partners (Brazil, China, Colombia). For each indicator, the list of countries with available information is reported in Supplementary Table 1, available as Supplementary data at *IJE* online.

Air-pollution indicators

To characterize long-term air-pollution exposures in each city, we used global estimates of annual fine particulate ($PM_{2.5}$) levels of the Data Integration Model for Air Quality available for the year 2014[24] and global annual mean ground-level nitrogen dioxide (NO_2) concentrations (3 years running mean for year 2001), developed by Geddes *et al.*[25] Both global estimates were calculated for grid cells with a spatial resolution of 0.1° for latitude and longitude.

We linked the 340 MCC cities with the databases containing the $PM_{2.5}$ and the NO_2 global estimates. Specifically, for each city, we assigned the $PM_{2.5}$ and NO_2 level of the grid cell [spatial resolution ($0.1° \times 0.1°$), which is approximately 11×11 km at the equator] including the coordinates of the city as defined by the World Cities database (https://simplemaps.com/data/world-cities).[26]

Population and density data

The World Cities database was used to retrieve population and density indicators for the year 2015. The former is an estimate of the city's population, whereas the latter is defined as population per square kilometre.

Weather variables

For each city, we calculated the average daily mean temperature and daily mean temperature range from the observed daily temperature distribution in the MCC dataset, in the city-specific observation period (between 1985 and 2014). These were used as basic indicators to avoid confounding by weather/climatological conditions.

Statistical methods

Description of the indicators

We summarize the distribution of indicators by country with the median, standard deviation and interquartile range (IQR). The relationships between indicators were examined first through the correlation matrix among all pairs of indicators. To remove the between-countries effects from the correlation, for all cities of a given country, the original indicator value was scaled by the country average indicator value. The country-adjusted correlation matrix was used as the input of a principal component analysis (PCA). The PCA is a statistical method that identifies factors (principal components) that best explain the covariability of the data. The principal components show groups of indicators that co-vary similarly in most cities, as can be illustrated in a score plot.

Association between the indicators and temperature–mortality impacts

We adopted a three-step approach to evaluate the association between the indicators and temperature–mortality impacts. Briefly, in the first stage, we calculated the city-specific temperature–mortality associations, followed by the estimation of the corresponding heat- and cold-attributable fractions and, in the last step, we fitted meta-regression models to evaluate the association between each indicator and heat and cold AF%. The three steps are described in more detail below.

First-stage time-series analysis

We estimated the city-specific temperature–mortality associations through quasi-Poisson regression[27] and distributed lags non-linear models (DLNMs).[28] We modelled the cross-basis function of daily mean temperature with a natural cubic spline function for the temperature dimension with three internal knots at the 10th, 75th and 90th percentiles of the city area-specific temperature distributions, and natural cubic spline with an intercept and two internal knots placed in equally spaced values in the log scale for the lag dimension. We extended the lag period to 21 days to capture the long delay in cold–mortality associations. We included a natural cubic B-spline function with 8 degrees of freedom (df) per year to control for long-term trends and seasonality, along with an indicator for day of the week. The model selection was based on previous work using a similar dataset.[3] We tested these modelling choices in a sensitivity analysis.

Estimation of city-specific heat- and cold-attributable fraction

To estimate the city-specific temperature at which mortality was minimal (called minimum mortality temperature, MMT) with greater precision, we applied a shrinkage procedure that borrows information across cities in the same country with similar climate. Details of this method are given in previous work.[3] We estimated attributable fractions (AF%, in percentage) using the first-stage (unshrunk) cumulative exposure–response associations, following a procedure described elsewhere.[29] In summary, we computed mortality attributable to cold and heat by summing the temperature-related deaths occurring in days with temperatures lower or higher than the MMT, and then dividing by the total number of deaths. We calculated empirical standard error using Monte Carlo simulations,[29] assuming a multivariate normal distribution of the first-stage reduced coefficients.

Table 2. MCC dataset: number of cities, deaths, period of observation and mean daily average temperature by country

Country	Cities	Level of development[a]	Deaths	Period	Daily average temperature (degrees Celsius)—mean (range)
Australia	3	Advanced economy	1 177 950	1988–2009	18.1 [5.6; 35]
Brazil	18	Developing economy	3 401 136	1997–2011	24.6 [3.6; 33.5]
Canada	26	Advanced economy	2 989 901	1986–2011	6.8 [–39.7; 32.1]
Chile	4	Developing economy	325 462	2004–14	13.7 [–1.7; 27.5]
China	15	Developing economy	950 130	1996–2008	15.1 [–23.7; 36.4]
Colombia	5	Developing economy	956 539	1998–2013	23.4 [10.5; 31.1]
Finland	1	Advanced economy	130 325	1994–2011	6.2 [–22.9; 25.5]
France	18	Advanced economy	1 197 555	2000–10	12.6 [–11.6; 32.4]
Iran	1	Developing economy	121 585	2004–13	16.0 [–14.7; 33.3]
Italy	16	Advanced economy	645 420	2001–10	15.7 [–10.7; 39.5]
Japan	7	Advanced economy	3 123 487	1985–2009	15.0 [–12.0; 33.1]
Mexico	10	Developing economy	2 980 086	1998–2014	18.8 [0.4; 35.3]
Moldova	4	Developing economy	59 906	2001–10	10.7 [–25.0; 32.6]
Philippines	4	Developing economy	274 516	2006–10	28.2 [21.8; 33.3]
South Korea	7	Advanced economy	1 726 938	1992–2010	13.7 [–15.7; 33.0]
Spain	51	Advanced economy	3 479 881	1990–2010	15.5 [–10.9; 36.8]
Sweden	1	Advanced economy	201 197	1990–2010	7.2 [–21.5; 26.8]
Switzerland	8	Advanced economy	243 638	1995–2013	10.4 [–14.9; 29.0]
Taiwan	3	Advanced economy	765 893	1994–2007	24.0 [8.1; 33.0]
UK	1	Advanced economy	1 325 902	1990–2012	11.6 [–5.5; 29.1]
USA	135	Advanced economy	22 953 896	1985–2006	14.9 [–31.4; 41.4]
Vietnam	2	Developing economy	108 173	2009–13	27.1 [14.4; 33.9]

[a]International Monetary Fund Advanced and Developing Economies List. World Economic Outlook, April 2016, p. 148; World Economic Outlook, April 2015, pp. 150–53, retrieved 26 June 2015; World Economic Outlook, Database—WEO Groups and Aggregates Information, April 2015, retrieved 26 June 2015.

Association between the indicators and heat- and cold-attributable fraction

We estimated whether the city-specific estimated temperature–mortality associations differed by city characteristics. For each indicator, we used the set of cities with available information and two separate meta-regression models were used to evaluate the association between the indicator and heat and cold AF% including indicators for countries, and average and range of daily mean temperature as meta-predictors. We tested and reported residual heterogeneity using the Cochran Q test and I^2 statistic, respectively.[30]

Results

Description of the sample

Descriptive statistics of mortality and temperature data are reported in Table 2. Almost 50 million deaths were observed in the study period. The 340 cities are located in 22 countries, 13 of which, according to the International Monetary Fund, are developed countries whereas 9 are developing countries (Table 2). Figure 1 shows the geographical distribution of the 340 cities and their average daily mean temperature, illustrating how this study covers various regions and climatic areas across the world.

Descriptive statistics of the 18 indicators considered in the analysis are shown in Table 3. Cities considered in this analysis show highly variable socio-economic, demographic, urban characteristics and air-pollution levels.

Weather variables, country and attributable mortality

Overall, we estimated that 0.54% (95% confidence interval: 0.49 to 0.58%) and 6.05% (5.59 to 6.36%) of mortality in the 340 MCC cities were attributable to heat and cold, respectively (Supplementary Table 2, available as Supplementary data at *IJE* online). Larger between-city heterogeneity was observed for heat AF% ($I^2 = 85.4\%$) than for cold AF% (64.2%). Country explained 15.7 and 10.9% of heterogeneity for heat AF% and cold AF%, respectively. In total, weather variables explained a further 22% of cold AF% heterogeneity, whereas heat AF% heterogeneity decreased by only 2.3%.

Demographic, socio-economic, environmental and urban indicators and attributable mortality

Associations between the indicators and heat- and cold-related AF% are reported in Figure 2 and Supplementary Table 3, available as Supplementary data at *IJE* online. Results are expressed as AF% variation for a standard

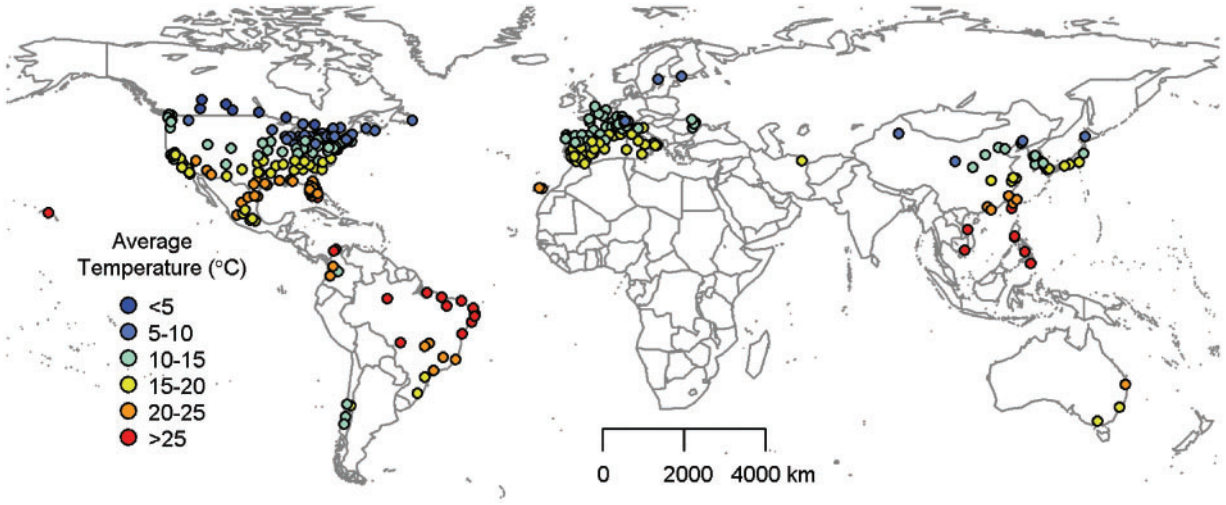


Figure 1. Average daily mean temperature in 340 MCC cities.

Table 3. Descriptive statistics of the 18 city-specific indicators considered in the analysis

Indicator	Number of cities	Median	IQR range	Range	SD
Demographic					
Population	340	418 800	[174 184; 1 416 981]	[7678; 26 174 599]	3 068 757.2
Density (population/km^2)	339	2771.0	[1282.6; 5638.6]	[9.3; 49 045.1]	7289.3
% population ≥65 years	320	12.8%	[10.4%; 15.1%]	[3.1%;27.2%]	4.7%
Life expectancy (years)	288	80.3	[78.5; 81.6]	[70.6; 85.0]	2.3
Socio-economic					
GDP (US$)	325	37 660	[27 096; 47 585]	[3168; 78 444]	15 838.5
Labour productivity (US$)	280	70 450	[64 019; 79 388]	[14 647; 366 027]	29 071.5
Educational level (%)	265	21.5%	[19.8%; 25.6%]	[9.0%; 39.3%]	5.3%
Unemployment rate (%)	307	6.5%	[4.4%; 9.4%]	[2.5%; 29.7%]	5.2%
Gini index	280	0.355	[0.315; 0.398]	[0.253; 0.484]	0.047
Poverty gap	280	22.1%	[18.2%; 26.3%]	[9.2%;40.0%]	6.0%
Health system					
Hospital-bed rates	279	29.0	[23.8; 35.3]	[1.6; 192.0]	23.0
Urban characteristics					
Type of surrounding region (rural/urban)	272	Predominantly urban = 125 Intermediate = 84 Predominantly rural = 63			
Urbanized area (%)	136	13.8%	[8.9%; 24.1%]	[0.2%; 68.7%]	13.4%
Green area (m^2 per million persons)	136	196.6	[37.6; 824.6]	[0.01; 6660.6]	1042.9
Concentration of population in the core (%)	136	83.5%	[72.8%; 93.4%]	[22.6%;100.0%]	16.0%
Sprawl	100	–0.99	[–2.71; 1.79]	[–12.13;10.97]	4.0
Air pollution					
$PM_{2.5}$ (μg/m^3)	340	9.6	[8.2; 13.9]	[4.7; 103.1]	13.2
NO_2 (ppb)	339	2.37	[1.08; 4.47]	[0.04; 23.3]	3.16

deviation (SD) increase of the indicator (provided in Table 3). No indicator is associated with cold-related AF. For heat, among demographic indicators, high life expectancy and high population density predicted high AF. Regarding the socio-economic indicators, GDP and educational level were positively associated with heat-related AF%. An inverse association was observed between number of hospital beds pro-capita and heat-related AF%. Cities with more inequalities (higher Gini index) had a larger mortality impact attributable to heat. Among the urban and environmental indicators considered, cities surrounded by a predominantly rural region and those with a larger green surface showed lower heat-related AF%, whereas $PM_{2.5}$ was positively associated with heat AF%.

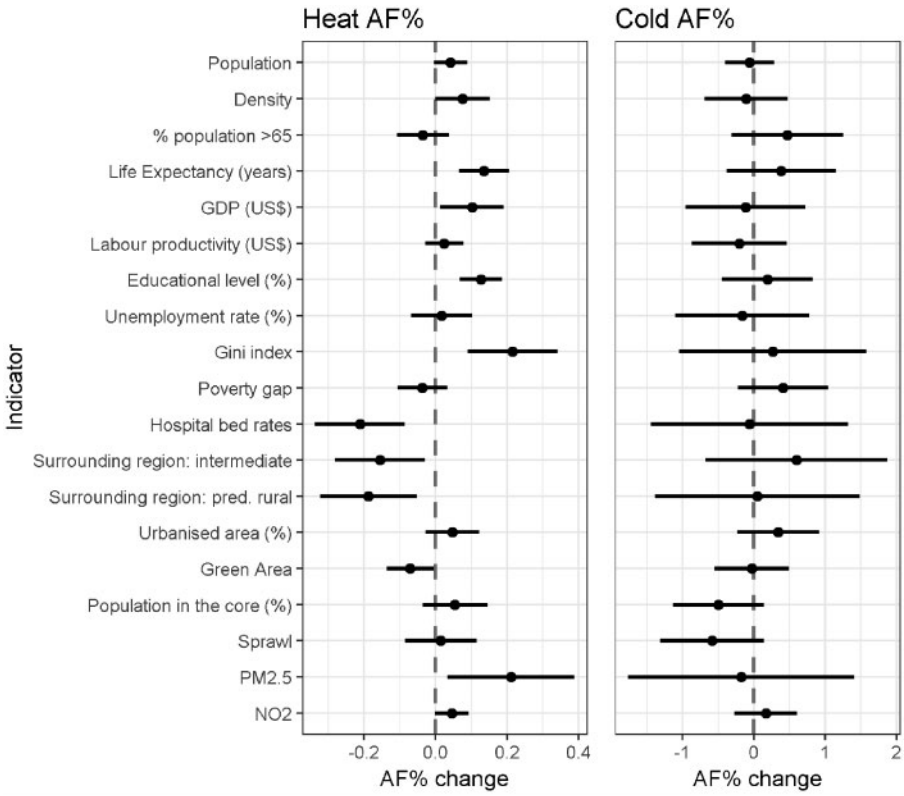


Figure 2. Associations between the indicators and heat and cold AF%. Coefficients and 95% confidence intervals calculated from a meta-regression model adjusted by country and weather variables. Results are expressed as AF% change for standard deviation increase in the indicators. The estimates of the coefficients and 95% confidence intervals are reported in Supplementary Table 3, available as Supplementary data at *IJE* online.

To give some insight into the inter-relationship between indicators and their association with attributable mortality, we performed a PCA. Supplementary Figures S1 and S2, available as Supplementary data at *IJE* online, show the correlation matrix and the results of the analysis. The first two principal components explained 44.4% of the total inertia. The first component seems characterized by the economic development of the MCC cities: high positive loading scores (represented by arrows) were observed for GDP, educational level and life expectancy. All these three variables showed a positive association with heat-related AF%. The second component characterized cities with higher levels of air pollution ($PM_{2.5}$ and NO_2), unemployment rates, inequalities (Gini index), poverty gaps, population and density. $PM_{2.5}$, Gini index and density were all positively associated with heat-related AF%.

Discussion

This study is based on the largest dataset ever collected to assess city-level modifiers of the temperature–health associations, which include more than 50 million deaths in 22 countries. The analysis allows investigating the heterogeneity of temperature-attributable mortality across 340 cities with a wide range of demographic, socio-economic and urban characteristics. Strengths of the study are the use of a standardized set of indicators, as well as the application of flexible statistical methods. Our findings suggest that more developed cities are perhaps surprisingly characterized by higher mortality attributable to heat, as indicated by the significant association with GDP, life expectancy and educational level. Furthermore, a second pattern emerged, with higher impact of heat on mortality in cities characterized by high population density, inequalities, pollution levels and fewer green spaces.

Cities have been centres of innovation and growth and the engines of economic development, but they are particularly vulnerable to the effects of climate change.[6,31,32] The nature of urban infrastructure creates microclimates that affect temperature; the urban heat island effect is an example, where cities are warmer than their surrounding hinterlands due to the thermal storage capacity of the built environment.[33] In our results, urban density is associated with an increased heat effect, which is also shown in other contextual studies.[5,14,19]

We used the OECD regional typology to characterize the region surrounding the urban setting considered in the analysis. This indicator is based on population density, degree of rurality and size of the urban centres located within the region. This indicators allows the identification of 63

cities in predominately rural regions (mainly based in the USA and Spain) and 84 in intermediate (both rural and urban) regions more evenly distributed across countries. These cities show a lower heat effect—a result that could be explained by a lower urban heat island effect and that is consistent with the increased heat effect observed for urban density.

Additional factors contribute to the vulnerability of cities. Among those of particular relevance are demographic structure, low socio-economic status and social inequity. In our study, we found a positive association between the Gini index of the city's region (an indicator of inequality) and heat impact. This result is consistent with those observed in contextual[10,11,19,34] and individual[35] studies showing a higher heat effect on communities or subjects with lower socio-economic status. Poorer housing conditions, lower prevalence of air conditioning, poorer health status and limited access to health care have been suggested as factors responsible for the increased heat effect in more deprived communities.[10,11,19] The elderly are more sensitive to non-optimal temperatures due to their higher prevalence of debilitating diseases, such as heart conditions, Alzheimer's disease and dementia,[36] which are associated with an increased effect of temperature on mortality.[19] In our study, we did not observe evidence of an association between the proportion of people aged more than 65 years and heat (or cold) attributable fraction. These results could be partially explained by the limited range of variation in age distributions across areas within the same country, as shown in our study, where the IQR range of the country-centred proportion aged more than 65 years was (–1.8%; 1.1%) on average within countries. Moreover, the proportion of elderly populations is higher in less urbanized and dense cities [+2.5% (+1.5%; +3.5%)]. The limited range of the exposure and possible confounding effect of urbanization could have limited in our study the power to detect the modifier role of age on the heat effect. We also note that our data are community-level and that future work with individual-level data is more suited to investigating these issues.

Urbanization is part of the development process and is generally associated with higher income, education and productivity level[33]; this relationship is shown in our study with a positive correlation ($r = 0.33$) between GDP and city density. At the individual level, higher income and education have been associated with lower heat-related mortality[35] due to the higher quality of housing and better access to information. In our study, however, heat-related impacts are higher in cities with a higher economic development characterized by a higher GDP, productivity, educational level and life expectancy. Using GDP as an indicator, Anderson and Bell[10] observed a similar positive contextual association in 107 US urban communities, whereas Hajat and Kosatky[37] found a negative association with GDP when meta-regressing heat coefficients across studies internationally. No association was observed at a contextual level in three other studies.[13,34,38] The increased impacts of heat on cities with higher GDP, educational level and life expectancy are not necessarily due to those cities being more unequal, as the correlation of those indicators with Gini is low. One explanation might be an association between economic development with features of urbanization such as the urban heat island, but further studies, including individual-level socio-economic indicators, are needed to clarify this.

The vulnerability of cities to climate change has motivated the development of city-level adaptation measures,[6,39] among which are urban planning and design including, for instance, cooling by greening and ventilation. Several studies have evaluated the modification effect of urban-landscape characteristics on the temperature–mortality association.[40–49] They used different neighbour-level indicators related to urban land use and land cover (e.g. impervious surface, open space, vegetation abundance), with some evidence of a protective effect of vegetation to reduce the heat effect on mortality.[41,44,49] These results are consistent with the negative association between green areas and heat AF% observed in our multi-city study.

Air pollution is also a well-known public-health risk factor. Particulate matters (PM_{10}, $PM_{2.5}$), ozone, nitrogen dioxide and sulphur dioxide have been linked to increases in morbidity and mortality.[50] There has been increasing interest in the synergist effect of temperature and pollution on morbidity and mortality.[51,52] Suggested mechanisms under the synergy hypothesis are, among others, that episodes of air pollution can increase vulnerability to the effects of temperature (e.g. respiratory diseases) and that elderly populations with deficiency of thermoregulation might suffer from high pollution levels.[53] The synergistic effect of pollutants and temperature has been studied mostly using case-only or time-series studies, with some evidence of an increased effect of particulate matters at higher temperatures.[51,52] In our study, we found a tendency for a higher AF% for heat in cities with higher levels of pollution as measured by $PM_{2.5}$ and NO_2; Benmarhnia *et al.*[11] found a similar contextual association between NO_2 and heat effect in Paris. These results need to take into account possible ecological confounding, as, in our dataset, the chronic level of pollutant examined ($PM_{2.5}$.and NO_2) is correlated with the city population and density, and shares with these urban-density indicators the tendency to increase the measured heat AF%.

Few studies have evaluated the role of healthcare access to reduce the temperature-related mortality.[13,34,37] Our

finding of reduced heat AF in cities with more hospital beds provides some evidence that an increased level of health services is an important component of adaptive capacity in an urban context.

Few studies have evaluated the role of area-level indicators as modifiers of cold effects on mortality with inconsistent results.[10,12–14,19,54] In our analysis, climate variables explain 22% of the heterogeneity, suggesting for cold-related effects a greater role of acclimatization. Moreover, more complex mechanisms for cold-related effects have been described[55] that may not be well captured by our set of indicators. Further research is needed in this area, possibly increasing the number of cities or the set of indicators, or with addition data such as individual-level data.

This study has several advantages. It represents the first investigation in which modifiers to both cold and heat effects at the city level were simultaneously assessed in a wide multi-country setting through a common study design and statistical framework. Previous multi-country studies[34,35,37,38] relied on simplifications of the exposure–response function[34,37,38] or qualitatively reviewed the evidence.[35] The statistical framework used in this analysis is based on a two-stage design that incorporates DLNMs and multivariate meta-regression to flexibly characterize complex temperature–health dependencies at a local level and to investigate their variations across cities.[56] We used the OECD Regional and Metropolitan Database as a source for defining socio-demographic indicators at the city level. This choice ensures a set of indicators collected using standardized criteria. We must also acknowledge some limitations. The observational periods and data-collection procedures are not uniform across all countries. Logistical constraints hinder perfectly consistent data streams across the globe, as different countries have various protocols for data acquisition and maintenance. However, our study design is not sensitive to potential biases arising from these differences and can appropriately pool information from data obtained from different sources. Specifically, our two-stage analytical framework includes indicators for countries as meta-predictors in the second-stage meta-regression. This means that, implicitly, the comparison is based on variations across locations within the same country, as any structural difference across countries is accounted for by the fixed-effects indicators. These differences include potential variations due to non-overlapping periods. The time frame of data collection varied for some variables and the reference period used for indicators varied between 2000 and 2014. Moreover, some of the indicators were measured after the actual city-specific time period of investigation. As a consequence, there could be some measurement errors in the level of the indicator associated with each city for the observational period. Under the hypothesis of no systematic bias within a country, this measurement error should lower the association under study towards a conservative error. However, we found a high correlation between indices in different years (data not shown), consequently this conservative error should be minor. The dataset includes several regions around the world, including developed and developing countries, but entire areas of the world are not covered, and there is a lack of information from countries with a lower degree of socio-economic development. Results might therefore not be globally representative. In our analysis, we considered each indicator as an explanatory variable in a meta-regression model adjusted by country and weather variables. We did not attempt a multivariable model, as many indicators exhibited collinearity, as shown in the PCA. Although it is an interesting research area, we did not plan subgroup analyses by climate zones or geographical regions. Further work increasing the number of locations, hopefully including developing countries, is needed to address this research question.

Conclusion

This study identifies several city characteristics that modify the vulnerability of urban populations to heat. These results can be used for determining the health burden projected in the future under specific climate-change and socio-demographic scenarios, and for the implementation of urban-development plans to mitigate the risk.

Supplementary data

Supplementary data are available at *IJE* online

Funding

This work was primarily supported by the Medical Research Council—UK (MR/M022625/1). The following individual grants also supported this work: Y.G. was supported by the Career Development Fellowship of Australian National Health and Medical Research Council (APP1107107); A.T. was supported by the Ministry of Education of Spain (PRX17/00705); J.J.K.J. and N.R.I.R. were supported by the Research Council for Health, Academy of Finland (266314); Y.L.G. was supported by the National Health Research Institutes of Taiwan (NHRI-EM-106-SP03); M.L.B. was supported by a US Environmental Protection Agency Assistance Agreement awarded to Yale University (83587101).

Conflict of Interest: None declared.

References

1. Basu R. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environ Health* 2009; 8:40.

2. Basu R, Samet JM. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev* 2002;**24**:190–202.
3. Gasparrini A, Guo Y, Hashizume M. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet* 2015;**386**:369–75.
4. Gasparrini A, Guo Y, Sera F *et al*. Projections of temperature-related excess mortality under climate change scenarios. *Lancet Planet Health* 2017;**1**:e360–67.
5. Hajat S, O'Connor M, Kosatsky T. Health effects of hot weather: from awareness of risk factors to effective health protection. *Lancet* 2010;**375**:856–63.
6. Carter JG, Cavan G, Connelly A, Guy S, Handley J, Kazmierczak A. Climate change and the city: building capacity for urban adaptation. *Prog Plann* 2015;**95**:1–66.
7. van den Bosch M, Sang ÅO. Urban natural environments as nature-based solutions for improved public health—a systematic review of reviews. *Environ Res* 2017;**158**:373–84.
8. Kalkstein LS, Sheridan SC, Kalkstein AJ. Heat/health warning systems: development, implementation, and intervention activities. In: Burton I, Ebi KL, McGregor G (eds.). *Biometeorology for Adaptation to Climate Variability and Change*. Dordrecht: Springer, 2009, pp. 33–48.
9. Toloo G, FitzGerald G, Aitken P, Verrall K, Tong S. Evaluating the effectiveness of heat warning systems: systematic review of epidemiological evidence. *Int J Public Health* 2013;**58**:667–81.
10. Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology* 2009;**20**:205.
11. Benmarhnia T, Oulhote Y, Petit C *et al*. Chronic air pollution and social deprivation as modifiers of the association between high temperature and daily mortality. *Environ Health* 2014;**13**:53.
12. Curriero FC, Heiner KS, Samet JM, Zeger SL, Strug L, Patz JA. Temperature and mortality in 11 cities of the eastern United States. *Am J Epidemiol* 2002;**155**:80–87.
13. Huang Z, Lin H, Liu Y *et al*. Individual-level and community-level effect modifiers of the temperature-mortality relationship in 66 Chinese communities. *BMJ Open* 2015;**5**:e009172.
14. Medina-Ramon M, Schwartz J. Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occup Environ Med* 2007;**64**:827–33.
15. Medina-Ramon M, Zanobetti A, Cavanagh DP, Schwartz J. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environ Health Perspect* 2006;**114**:1331–36.
16. O'Neill MS, Zanobetti A, Schwartz J. Modifiers of the temperature and mortality association in seven US cities. *Am J Epidemiol* 2003;**157**:1074–82.
17. Stafoggia M, Forastiere F, Agostini D *et al*. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology* 2006;**17**:315–23.
18. Yu W, Vaneckova P, Mengersen K, Pan X, Tong S. Is the association between temperature and mortality modified by age, gender and socio-economic status? *Sci Total Environ* 2010;**408**:3513–18.
19. Zanobetti A, O'Neill MS, Gronlund CJ, Schwartz JD. Susceptibility to mortality in weather extremes: effect modification by personal and small-area characteristics. *Epidemiology* 2013;**24**:809–19.
20. Maraut S, Dernis H, Webb C, Spiezia V, Guellec D. *The OECD REGPAT Database*. Paris, France: OECD Working Paper, 2008.
21. Brezzi M, Piacentini M, Rosina K, Sanchez-Serra D. *Redefining Urban Areas in OECD Countries. Redefining "Urban": A New Way to Measure Metropolitan Areas*. Paris, France: OECD, 2012, pp. 19–58.
22. OECD. *Regions at a Glance 2016*. Paris: OECD Publishing, 2016.
23. OECD. *Quality Framework and Guidelines for OECD Statistical Activities*. Paris, France: OECD, 2011.
24. Shaddick G, Thomas ML, Green A *et al*. Data integration model for air quality: a hierarchical approach to the global estimation of exposures to ambient air pollution. *J R Stat Soc C* 2018;**67**:231–53.
25. Geddes JA, Martin RV, Boys BL, van Donkelaar A. Long-term trends worldwide in ambient NO_2 concentrations inferred from satellite observations. *Environ Health Perspect* 2016;**124**:281.
26. Simplemaps. *SimpleMaps Geographic Data Products, World Cities Database*, 2016. http://simplemaps.com/data/world-cities (1 May 2017, date last accessed).
27. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. *Int J Epidemiol* 2013;**42**:1187–95.
28. Gasparrini A. Modeling exposure–lag–response associations with distributed lag non-linear models. *Stat Med* 2014;**33**:881–99.
29. Gasparrini A, Leone M. Attributable risk from distributed lag models. *BMC Med Res Methodol* 2014;**14**:55.
30. Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;**21**:1539–58.
31. Habitat U. *Cities and Climate Change: Global Report on Human Settlements 2011*. London: Earthscan, 2011.
32. Schauser I, Otto S, Schneiderbauer S *et al*. *Urban Regions: Vulnerabilities, Vulnerability Assessments by Indicators and Adaptation Options for Climate Change Impacts: Scoping Study*. (ETC/ACC Technical Paper 2010/12, December 2010). The Netherlands: The European Topic Centre on Air and Climate Change (ETC/ACC), 2010. Retrieved from http://acm.eionet.europa.eu/reports/docs/ETCACC_TP_2010_12_Urban_CC_Vuln_Adapt.pdf.
33. Kamal-Chaoui L, Robert A. *Competitive Cities and Climate Change*. Paris, France: OECD Regional Development Working Papers 2009/1, 2009.
34. Leone M, D'Ippoliti D, De Sario M *et al*. A time series study on the effects of heat on mortality and evaluation of heterogeneity into European and Eastern-Southern Mediterranean cities: results of EU CIRCE project. *Environ Health* 2013;**12**:55.
35. Romero-Lankao P, Qin H, Dickinson K. Urban vulnerability to temperature-related hazards: a meta-analysis and meta-knowledge approach. *Global Environ Change* 2012;**22**:670–83.
36. Feigin V. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1459–544.

37. Hajat S, Kosatky T. Heat-related mortality: a review and exploration of heterogeneity. *J Epidemiol Commun Health* 2010;**64**: 753–60.
38. McMichael AJ, Wilkinson P, Kovats RS *et al*. International study of temperature, heat and urban mortality: the 'ISOTHURM' project. *Int J Epidemiol* 2008;**37**:1121–31.
39. Georgi B, Swart R, Marinova N, Van Hove B, Jacobs C, Klostermann J. *Urban Adaptation to Climate Change in Europe: Challenges and Opportunities for Cities Together with Supportive National and European Policies*. Report No.: 929213308X. Copenhagen, Denmark: EEA, 2012.
40. Eisenman DP, Wilhalme H, Tseng CH *et al*. Heat death associations with the built environment, social vulnerability and their interactions with rising temperature. *Health Place* 2016;**41**: 89–99.
41. Harlan SL, Brazel AJ, Prashad L, Stefanov WL, Larsen L. Neighborhood microclimates and vulnerability to heat stress. *Soc Sci Med* 2006;**63**:2847–63.
42. Harlan SL, Declet-Barreto JH, Stefanov WL, Petitti DB. Neighborhood effects on heat deaths: social and environmental predictors of vulnerability in Maricopa County, Arizona. *Environ Health Perspect* 2013;**121**:197–204.
43. Klein Rosenthal J, Kinney PL, Metzger KB. Intra-urban vulnerability to heat-related mortality in New York City, 1997–2006. *Health Place* 2014;**30**:45–60.
44. Madrigano J, Ito K, Johnson S, Kinney PL, Matte T. A case-only study of vulnerability to heat wave-related mortality in New York City (2000–2011). *Environ Health Perspect* 2015;**123**: 672–78.
45. Rey G, Fouillet A, Bessemoulin P *et al*. Heat exposure and socioeconomic vulnerability as synergistic factors in heat-wave-related mortality. *Eur J Epidemiol* 2009;**24**:495–502.
46. Smoyer KE, Rainham DG, Hewko JN. Heat-stress-related mortality in five cities in Southern Ontario: 1980–1996. *Int J Biometeorol* 2000;**44**:190–97.
47. Tan J, Zheng Y, Tang X *et al*. The urban heat island and its impact on heat waves and human health in Shanghai. *Int J Biometeorol* 2010;**54**:75–84.
48. Uejio CK, Wilhelmi OV, Golden JS, Mills DM, Gulino SP, Samenow JP. Intra-urban societal vulnerability to extreme heat: the role of heat exposure and the built environment, socioeconomics, and neighborhood stability. *Health Place* 2011;**17**:498–507.
49. Xu Y, Dadvand P, Barrera-Gomez J *et al*. Differences on the effect of heat waves on mortality by sociodemographic and urban landscape characteristics. *J Epidemiol Community Health* 2013; **67**:519–25.
50. WHO. *Review of Evidence on Health Aspects of Air pollution—REVIHAAP Project*. Technical Report. Copenhagen, 2013. Available from: http://www.euro.who.int/__data/assets/pdf_file/0004/193108/REVIHAAP-Final-technical-report-final-version.pdf
51. Analitis A, De' Donato F, Scortichini M *et al*. Synergistic effects of ambient temperature and air pollution on health in Europe: results from the PHASE project. *Int J Environ Res Public Health* 2018;**15**. doi: 10.3390/ijerph15091856.
52. Chen F, Fan Z, Qiao Z *et al*. Does temperature modify the effect of PM10 on mortality? A systematic review and meta-analysis. *Environ Pollut* 2017;**224**:326–35.
53. Lam CKC. Air pollution, heat and mortality in urban populations. *Reinvention Int J Undergrad Res* 2014;**7**. Available from: http://www2.warwick.ac.uk/fac/cross_fac/iatl/reinvention/issues/volume7issue1/lam/
54. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? *Occup Environ Med* 2007;**64**:93–100.
55. Kinney PL, Schwartz J, Pascal M *et al*. Winter season mortality: will climate warming bring benefits? *Environ Res Lett* 2015;**10**: 064016.
56. Gasparrini A, Armstrong B, Kenward M. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat Med* 2012;**31**:3821–39.

Supplementary Material

Table S1. List of countries with available information by indicator.

Indicator	N cities	Countries
Demographic		
Population	340	Australia; Brazil; Canada; Chile; China; Colombia; Finland; France; Iran; Italy; Japan; Mexico; Moldova; Philippines; South Korea; Spain; Sweden; Switzerland; Taiwan; UK; USA; Viet Nam
Density	339	Australia; Brazil; Canada; Chile; China; Colombia; Finland; France; Iran; Italy; Japan; Mexico; Moldova; Philippines; South Korea; Spain; Sweden; Switzerland; Taiwan; UK; USA; Viet Nam
Old population (%)	320	Australia; Brazil; Canada; Chile; China; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Life Expectancy (years)	288	Australia; Canada; Chile; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Socioeconomic		
GDP ($)	325	Australia; Brazil; Canada; Chile; China; Colombia; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Labour productivity ($)	280	Australia; Canada; Chile; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; UK; USA
Educational level (%)	265	Canada; Chile; France; Japan; Mexico; South Korea; Spain; Sweden; UK; USA
Unemployment rate (%)	307	Australia; Canada; Chile; China; Colombia; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Gini index	280	Australia; Canada; Chile; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; UK; USA
Poverty gap	280	Australia; Canada; Chile; Finland; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; UK; USA
Health system		
Hospital bed rates	279	Australia; Canada; Chile; France; Italy; Japan; Mexico; Spain; Sweden; Switzerland; USA
Urban characteristics		
Type of surrounding region (rural/urban)	272	Australia; Canada; China; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; USA
Urbanised area	136	OECD Metropolitan area database Australia; Canada; Chile; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Green Area	136	OECD Metropolitan area database Australia; Canada; Chile; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Concentration of population in the core (%)	136	OECD Metropolitan area database Australia; Canada; Chile; France; Italy; Japan; Mexico; South Korea; Spain; Sweden; Switzerland; UK; USA
Sprawl	100	OECD Metropolitan area database

		Chile; France; Italy; Japan; Mexico; Spain; Sweden; Switzerland; UK; USA
Environmental		
PM2.5 ($\mu g/m^{-3}$)	340	Australia; Brazil; Canada; Chile; China; Colombia; Finland; France; Iran; Italy; Japan; Mexico; Moldova; Philippines; South Korea; Spain; Sweden; Switzerland; Taiwan; UK; USA; Viet Nam
NO_2 (ppb)	339	Australia; Brazil; Canada; Chile; China; Colombia; Finland; France; Iran; Italy; Japan; Mexico; Moldova; Philippines; South Korea; Spain; Sweden; Switzerland; Taiwan; UK; USA; Viet Nam

Table S2. AF% for heat and cold calculated in the all sample, and AF% change due a one degree increase of each climate indicator

		Cold				Heat		
	I^2	**AF%**	**95% CI**		I^2	**AF%**	**95% CI**	
Covariate	64.2%	6.05	(5.59; 6.36)		85.4%	0.54	(0.49; 0.58)	
Country	52.3%				69.7%			
		AF% change	**95% CI**	**p value**		**AF% change**	**95% CI**	**p value**
Climate	30.3%				67.4%			
Average mean temperature		-0.227	(-0.383; -0.072)	0.004		0.007	(-0.010; 0.023)	0.422
Range mean temperature		0.095	(0.010; 0.181)	0.029		0.022	(0.013; 0.031)	<0.001

Table S3. Associations between the indicators and heat and cold AF%. Coefficients and 95%CI calculated form a meta-regression model adjusted by country and weather variables. Results are expressed as AF% change for SD increase of the indicators.

		Cold			Heat		
Indicator	**n**	**AF% change**	**95% CI**	**p value**	**AF% change**	**95% CI**	**p value**
Demographic							
Population	340	-0.025	(-0.164; 0.115)	0.729	0.042	(-0.005; 0.089)	0.079
Density (population/km2)	339	-0.065	(-0.415; 0.285)	0.717	0.076	(-0.001; 0.152)	0.052
% population ≥65 years	320	0.468	(-0.313; 1.250)	0.240	-0.035	(-0.109; 0.038)	0.342
Life Expectancy (years)	288	0.518	(-0.502; 1.556)	0.328	0.136	(0.066; 0.206)	<0.001
Socioeconomic							
GDP (US$)	325	-0.118	(-0.958; 0.722)	0.783	0.102	(0.013; 0.191)	0.025
Labour productivity (US$)	280	-0.205	(-0.872; 0.463)	0.548	0.024	(-0.029; 0.078)	0.369
Educational level (%)	265	0.188	(-0.451; 0.827)	0.564	0.127	(0.068; 0.185)	<0.001
Unemployment rate (%)	307	-0.161	(-1.102; 0.779)	0.737	0.017	(-0.069; 0.103)	0.695
Gini index	280	0.268	(-1.047; 1.583)	0.690	0.216	(0.090; 0.341)	0.001
Poverty gap	280	0.410	(-0.223; 1.044)	0.204	-0.037	(-0.106; 0.033)	0.301
Health system							
Hospital bed rates	279	-0.065	(-1.448; 1.318)	0.927	-0.212	(-0.337; -0.087)	0.001
Urban characteristics							
Type of surrounding region: intermediate	272	0.597	(-0.679; 1.873)	0.359	-0.156	(-0.281; -0.031)	0.015
Type of surrounding region: predominately rural	272	0.049	(-1.388; 1.486)	0.946	-0.188	(-0.323; -0.053)	0.006
Urbanised area (%)	136	0.341	(-0.235; 0.917)	0.246	0.047	(-0.027; 0.122)	0.214
Green Area (m^2per million person)	136	-0.031	(-0.554; 0.493)	0.909	-0.071	(-0.137; -0.005)	0.036
Concentration of population in the core (%)	136	-0.498	(-1.14; 0.143)	0.128	0.054	(-0.036; 0.145)	0.241
Sprawl	100	-0.587	(-1.317; 0.143)	0.115	0.015	(-0.085; 0.115)	0.769
Air pollution							
$PM_{2.5}$ ($\mu g/m^{-3}$)	340	-0.178	(-1.766; 1.411)	0.826	0.211	(0.033; 0.389)	0.020
NO_2 (ppb)	339	0.168	(-0.273; 0.609)	0.456	0.045	(-0.002; 0.093)	0.059

[2]Meta-regression model adjusted by Country and climate variables

Table S4: Sensitivity analysis. Attributable fraction (AF) % (total, heat, and cold), using varying modelling choices.

Modelling choices (340 cities)	Total (%)	Cold AF %	Heat AF %
Main model	6.59	6.05	0.54
Knots for exposure-response: 10th, 50th, and 90th	6.62	5.97	0.65
Knots for exposure-response: 10th, 25th, 75th and 90th	6.66	6.08	0.58
Quadratic B-spline for exposure-response	7.27	6.66	0.62
Df for lag-response: 6	6.56	6.03	0.54
Lag period: 14 days	5.24	4.64	0.60
Lag period: 28 days	7.73	7.14	0.59
Df/year for seasonal control: 4	5.73	5.10	0.63
Df/year for seasonal control: 6	6.43	5.96	0.46
Df/year for seasonal control: 10	5.63	5.05	0.58

Figure S1. Correlation matrix between the country-centered indicators. Analysis made on 255 MCC cities with no missing data on the indicators.

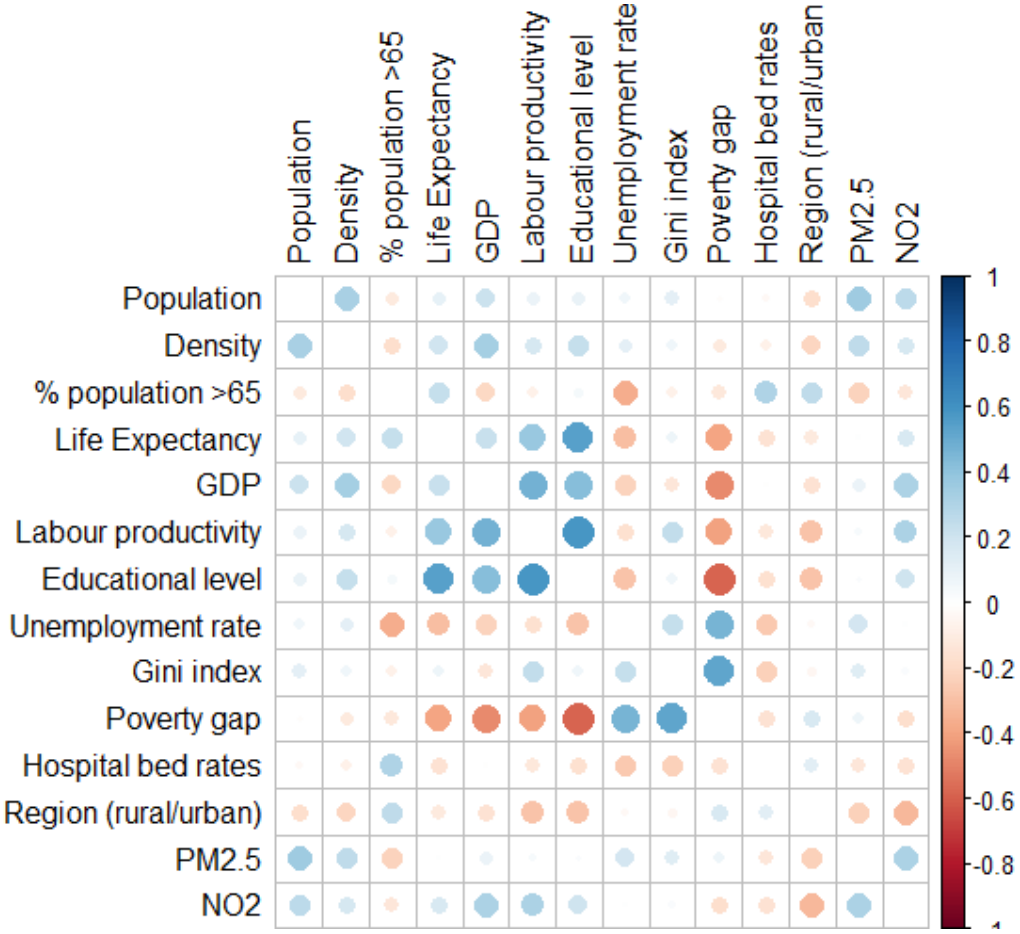


Figure S2. Score plot (for cities) and loading plot (for the variables) of the principal component analysis performed on 248 MCC locations.

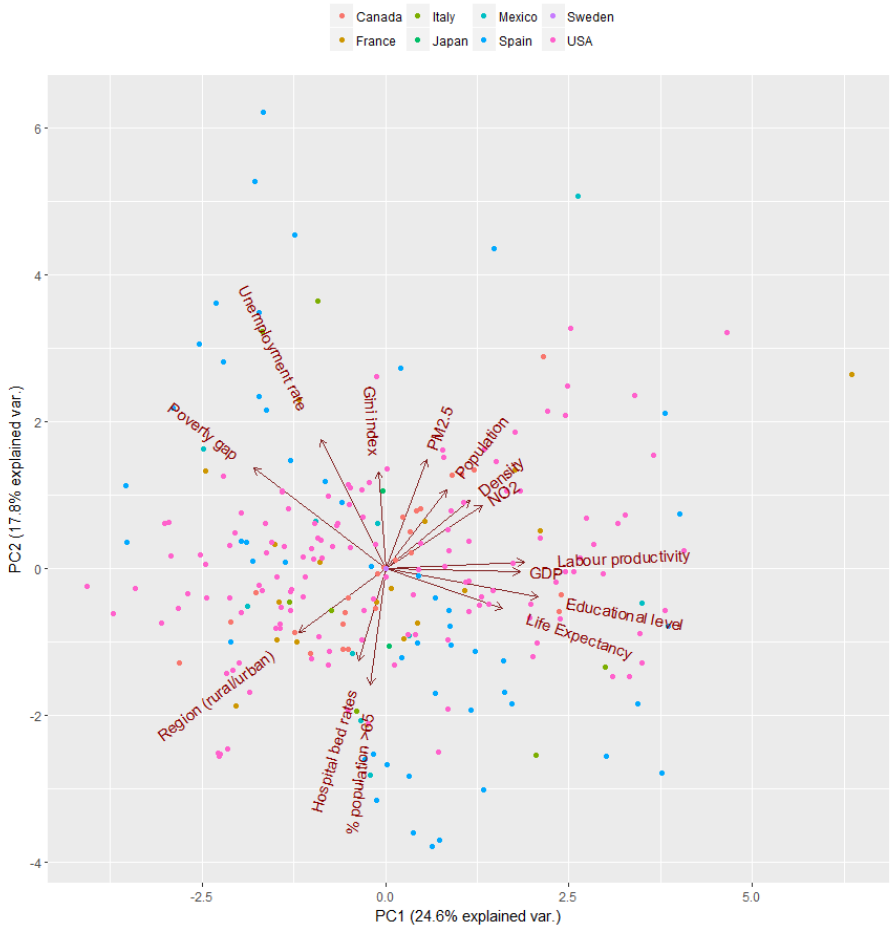
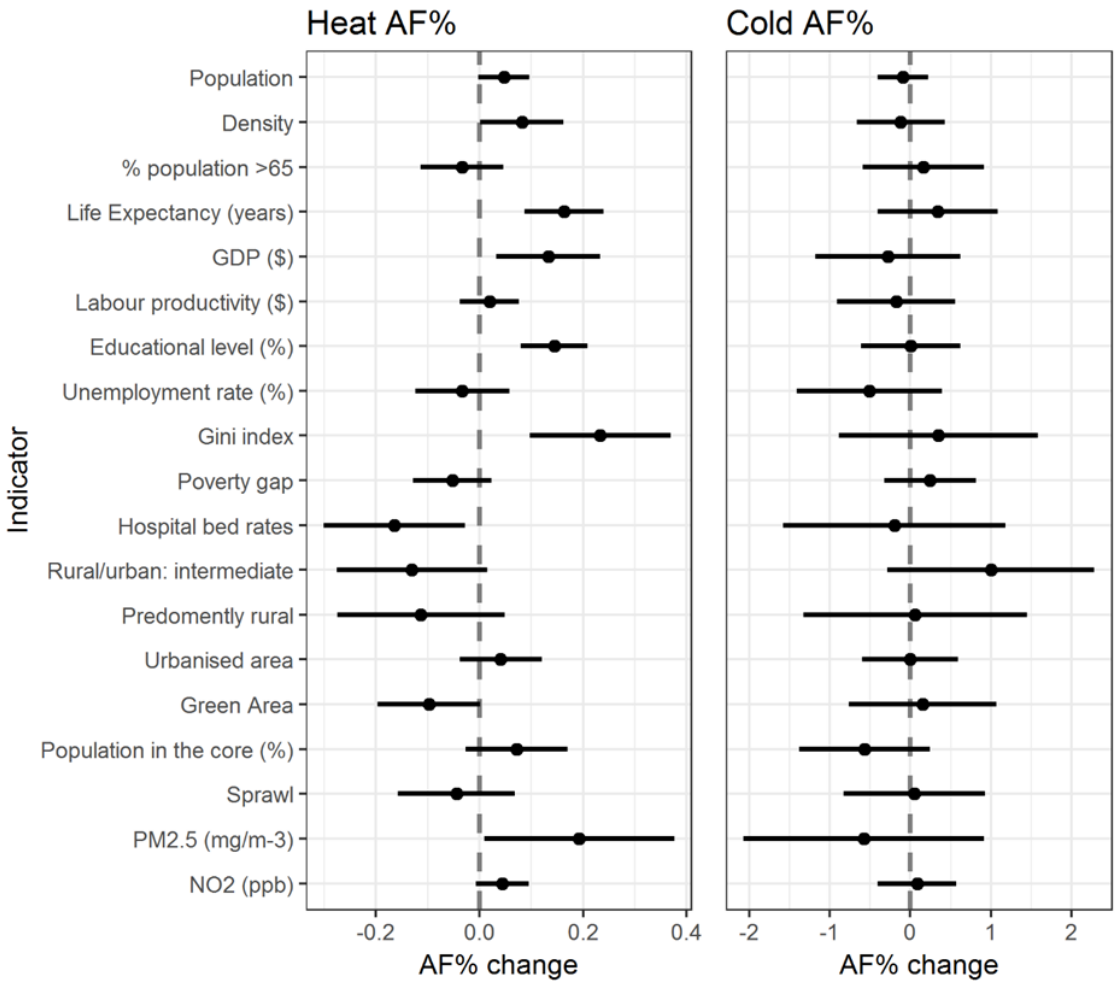


Figure S3. Associations between the indicators and heat and cold attributable fraction (AF) % between 1995 and 2010. Coefficients and 95%CI calculated form a meta-regression model adjusted by country and weather variables. Results are expressed as AF% change for SD increase of the indicators.



SUPPLEMENTARY INFORMATION ON THE MCC DATASET

The dataset has been collected through the Multi-City Multi-Country (MCC) network, an international collaboration of research teams working on a program aiming to produce epidemiological evidence on associations between weather and health (http://mccstudy.lshtm.ac.uk/). The dataset collects time series data on environmental and health variables, together with metadata, from several locations (either cities, metropolitan areas, provinces, prefectures, or regions) in multiple countries. The current analysis includes data from 340 locations from 22 countries. The datasets from 8 countries (Australia, Brazil, China, South Korea, Spain, Taiwan, and USA) were described in the online appendix of a previous publication[1]. Here we provide details on 15 countries (Canada, Chile, Colombia, Finland, France, Iran, Italy, Japan, Mexico, Moldova, Philippines, Sweden, Switzerland, UK, and Viet Nam).

Canada. We collected data from 25 census metropolitan areas (CMA) and the city of Hamilton between the 1st of January 1986 and 31st of December 2011. A list of 20 CMA is reported in the online appendix of a previous publication.[1] The other five CMA are Niagara, Oakville, Oshawa, Sarnia, and Sault Ste Marie. Daily mortality, obtained from Statistics Canada through access to the Canadian Mortality Database, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C) and relative humidity (in %), computed as the 24-hour average based on hourly measurements, were obtained from Environment Canada. A single weather station was selected for each city using the airport monitoring station located closest to the CMA centre. Measures of ozone (O_3, in ppb), nitrogen dioxide (NO_2, in ppb) and particles ($PM_{2·5}$, in ppb) were available in the same period from the National Air Pollution Surveillance (NAPS) network of Environment Canada. Daily level of pollutants was computed as the 24-hour mean based on hourly measurements in different stations, and then averaged across stations with no missing data, with an average of 4 stations per city. In total, missing data amount for 2·00% and 0·97% of the mortality and temperature series, respectively. These data were used and described in previous publications.[2,3]

Chile. We collected data from the city of Chillan (1st of January 2008 to 31st of December 2014), Santiago de Chile (1st of January 2008 to 31st of December 2014), Temuco (31st of March 2004 to 31st of July 2014), and Valparaiso between 7th of August 2004 and 27th of June 2014. Daily mortality, obtained from Departamento de Estadísticas e Información de Salud (Ministerio de Salud), is represented by counts of deaths for all causes. Mean daily temperature (in ˚C), computed as 24-hour average based on hourly measurements, were obtained from Sistema de Información Nacional de Calidad del Aire (SINCA), Ministerio del Medio Ambiente. For Santiago de Chile a single weather station "Parque Ohiggins" was selected. In total, missing data amount for 0.00% and 10.09% of the mortality and temperature series, respectively.

Colombia. We collected data from Bogota, Barranquilla, Cali, Cartagena and Medellin between 1st of January 1998 and 31st of December 2013. Daily mortality, obtained Administrativo Nacional de Estadistica (DANE), is represented by counts of deaths for all causes and for non-external causes only (ICD-9: 0-799; ICD-10: A00-R99). Mean daily temperature (in ˚C) and relative humidity (in %), computed as 24-hour average based on hourly measurement, were obtained from Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia (IDEAM). A single weather station was selected for each city; in particular, for Bogota, Barranquilla, Cartagena and

Medellin we used the airport monitoring station located closest to the city centre. In total, missing data amount for 0.01% and 2.81% of the mortality and temperature series, respectively.

Finland. We collected data from the city of Helsinki between the 1st of January 1994 and 31st of December 2011. Daily mortality, obtained from Statistics Finland, is represented by counts of deaths for all causes. A dataset containing minimum, mean, and maximum daily temperatures was obtained from the Finnish Meteorological Institute. In this dataset, point measurements from the weather measuring stations around the country have been interpolated onto a 10×10 km grid covering the whole of Finland, using a Kriging model. The temperature variables in the Helsinki Temperature Time-series have been extracted from the GIS-database for KKJ-coordinates 6675470:2552920 (KKJ, Finnish National Coordinate System based on ED50). These are the coordinates for weather measuring station Kallion urheilukenttä of Helsinki Region Environmental Services Authority HSY. In total, missing data amount for 0.00% and 0.00% of the mortality and temperature series, respectively.

France. We collected data from 18 cities (see full list in Table B1) between 1st of January 2000 and 31st of December 2010. Daily mortality, obtained from French National Institute of Health and Medical Research (CepiDC), is represented by counts of deaths for all causes. Mean daily temperature (in ˚C), computed as the mean of the minimal and maximal temperature, were obtained from the Meteo France. A single weather station was selected for each city; for 15 out 18 locations the weather station was located at the nearest airport. Two cities (Lille and Lens) have the same temperature series from the same weather station. In total, missing data amount for 0.00% and 0.06% of the mortality and temperature series, respectively.

Iran. We collected data from the city of Mashhad between 1st of January 2004 and 31st of December 2013. Daily mortality, obtained from Database of Mashhad Municipality, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C) and relative humidity (in %), computed as the 24-hour average based on hourly measurements, were obtained from the Iran Meteorological Organization. A single weather station was selected. In total, missing data amount for 2.11% and 0.03% of the mortality and temperature series, respectively.

Italy. We collected data from 16 cities (see full list in Table B1) between 1st of January 2001 and 31st of December 2010. Daily mortality, obtained from local mortality registries and from the rapid mortality surveillance system operational since 2004, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C), computed as the 24-h average based on 20-minutes measurements, were obtained from the Meteorological Service of the Italian Air Force. A single weather station was selected for each city, using the airport monitoring station located closest to the city centre. In total, missing data amount for 1.61% and 2.58% of the mortality and temperature series, respectively.

Japan. We collected data from 7 cities (see full list in Table B1) between 1st of January 1985 and 31st of December 2009. Daily mortality, obtained from computerized death certificate data from the Ministry of Health, Labour and Welfare, Japan, is represented by counts of deaths for all causes and for non-external causes only (ICD-9: 0-799; ICD-10: A00-R99). Mean daily temperature (in ˚C) and relative humidity (in %) were computed as the 24-hour average based on hourly measurements. In total, missing

data amount for 0.00% and 0.00% of the mortality and temperature series, respectively.

Mexico. We collected data from 10 metropolitan areas (see full list in Table B1) between 1st of January 1998 and 31st of December 2014. Daily mortality, obtained from the National Institute of Statistics, Geography and Informatics is represented by counts of deaths for all causes. Temperature and relative humidity data, were obtained from Servicio Meteorológico Nacional (SMN) and the Instituto Nacional de Ecología y Cambio Climático (INECC). Daily mean temperature as well as the maximum or minimum daily temperature was calculated using hourly data, with a minimum of adequacy of information of 50%. In the case of the data from airports, these were daily averages of temperature and relative humidity. We obtained the maximum and minimum temperatures as well as the average relative humidity across all stations that met the sufficiency criteria of having at least 75% data for the day. Measures of ozone (O_3, in ppb), particles (PM_{10}, in $\mu g/m^3$), and fine particles measures ($PM_{2.5}$, in $\mu g/m^3$) were available in the same period. Daily level of pollutants was computed as the 24-hour mean based on hourly measurements. In total, missing data amount for 0.00% and 26.67% of the mortality and temperature series, respectively.

Moldova. We collected data from the city of Anenii Noi (1st of January 2003 to 31st of December 2010), Cahul (1st of January 2003 to 31st of December 2010), Chisinau (1st of January 2001 to 31st of December 2010), and Falesti (1st of January 2003 to 31st of December 2010). Daily mortality, obtained from National Centre for Health Management, Moldova, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C) computed as the average between daily minimum and maximum, were obtained from State Hydrometeorological Service, Moldova. A single weather station was selected for each city. In total, missing data amount for 0.00% and 0.00% of the mortality and temperature series, respectively.

Philippines. We collected data from four cities (Cebu, Davao, Manila, and Quezon) between the 1st of January 2006 and 31st of December 2010. Daily mortality, obtained from Philippine Statistics Authority-National Statistics Office, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C), and relative humidity (in %), computed as the 24-hour average based on hourly measurements, were obtained from Philippine Atmospheric Geophysical and Astronomical Services Administration, and National Oceanic and Atmospheric Administration. A single weather station was selected for each city. In particular, for Davao and Cebu we used the airport monitoring station closest to the city centre, while the weather station was located in Manila Port for Manila, and in the Science Garden for Quezon. In total, missing data amount for 0.14% and 0.01% of the mortality and temperature series, respectively.

Sweden. We collected data from the city of Stockholm between 1st of January 1990 and 31st of December 2010. Daily mortality, obtained from the Swedish Cause of Death Register at the Swedish National Board of Health and Welfare, is represented by counts of non-external causes only (ICD-9: 0-799; ICD-10: A00-R99). Mean daily temperature (in ˚C) computed as the 24-hour average based on hourly measurements, were obtained from the Swedish Meteorological and Hydrological Institute. A measuring station, located at Bromma Airport, was selected. In total, missing data amount for 0.00% and 0.25% of the mortality and temperature series, respectively.

Switzerland. We collected data from 8 cities (Basel, Bern, Zurich, Geneva, Lausanne, Lucerne, Lugano, St. Gallen) between 1st January 1995 to 31st December 2013.

Lugano also includes the small municipalities around the main city of Lugano with similar altitude. Daily mortality, provided by the Federal Office of Statistics (Switzerland), is represented by counts of non-external deaths with accidents included (International Classification of Diseases, 10th revision (ICD-10) codes A00-R99, V01-V99, W00-X59). Daily data on several meteorological indicators were collected from the IDAWEB database (a service provided by MeteoSwiss, the Swiss Federal Office of Meteorology and Climatology). A single weather station located within or near the urban area was selected for each city. The meteorological indicators were: mean daily temperature (in ˚C) and relative humidity (in %), computed as the 24-hour average based on hourly measurements. Daily measurements of nitrogen dioxide (NO_2, in ppb), particles (PM_{10}, in ppb), and ozone (O_3, in ppb), were provided by the Immissionsdatenbank Luft (IDB, Federal Office of the Environment, Bern, Switzerland) and computed as 24-h average for the former two and as 1h-maximum for the latter. In total, missing data amount for 0.00% and 0.00% of the mortality and temperature series, respectively.

UK. We collected data for the region of London between 1st of January 1990 and 31st of August 2012. Daily mortality, obtained from the Office of National Statistics, is represented by counts of deaths for all causes. Mean daily temperature (in ˚C) computed from the 24-h average of hourly measurements) were obtained from the British Atmospheric Data Centre. Data from seven meteorological stations were used to derive the mean daily temperature time serie. In total, missing data amount for 0.00% and 0.00% of the mortality and temperature series, respectively

Vietnam. We collected data from the cities of Ho Chi Minh City (1st of January 2010 to 31st of December 2013), and Hue (1st of January 2009 to 31st of December 2013). Since 1992, a mortality data-collecting system based on the commune health center has been introduced in an official book named A6 [Vietnam Ministry of Health. Decision No 822/BYT.QD to issue mortality reporting book A6/YTCS. 1992]. Data from the A6 are collected at the commune health center level and then forwarded to the provincial and central levels. In this study, daily mortality data from the A6 mortality reporting system, is represented by counts of deaths for all causes and for non-external causes only (ICD-9: 0-799; ICD-10: A00-R99). Mean daily temperature (in ˚C), and relative humidity (in %) computed as computed from the 24-h average of hourly measurements, were obtained from National Oceanic and Atmospheric Administration's (NOAA) National Climate Data Center (NCDC). A single weather station was selected for each city. In total, missing data amount for 0.00% and 0·45% of the mortality and temperature series, respectively.

REFERENCES

1. Gasparrini, A. et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 386, 369–375 (2015).
2. Kaplan, G. G. et al. Ambient ozone concentrations and the risk of perforated and nonperforated appendicitis: a multicity case-crossover study. Environ. Health Perspect. 121, 939–943 (2013).
3. Martin, S. L., Cakmak, S., Hebbern, C. A., Avramescu, M.-L. & Tremblay, N. Climate change and future temperature-related mortality in 15 Canadian cities. Int J Biometeorol 56, 605–619 (2012).
4. Armstrong, B. G. et al. Association of mortality with high temperatures in a temperate climate: England and Wales. J Epidemiol Community Health 65, 340–345 (2011).
5. Gasparrini, A., Armstrong, B., Kovats, S. & Wilkinson, P. The effect of high temperatures on cause-specific mortality in England and Wales. Occup Environ Med 69, 56–61 (2012).

Table B1. Descriptive statistics and indicators for the 340 MCC cities included in the analysis.

City	Country	Deaths	Start year	End year	T mean	Min Temp	Max Temp	Population	Density	GDP	Gini Index	$PM_{2.5}$
Brisbane	Australia	191996	1988	2009	20.3	7.4	31.5	1856353	2023.5	37520	0.332	6.567
Melbourne	Australia	449092	1988	2009	15.7	5.6	35.0	2178625	3224.6	37745	0.319	6.446
Sydney	Australia	536862	1988	2009	18.3	7.9	31.8	5956422	3824.3	39965	0.348	6.040
Belem	Brazil	132910	1997	2011	26.9	23.4	29.4	2491237	22674.8	5040	NA	13.918
Belo Horizonte	Brazil	464778	1997	2011	22.0	11.9	29.5	5160724	7379.2	8799	NA	9.626
Brasilia	Brazil	114801	1997	2011	21.3	14.1	29.1	4421461	846.1	30793	NA	8.324
Cuiaba	Brazil	51336	1997	2011	26.4	11.4	33.4	803287	2578.5	11413	NA	18.757
Curitiba	Brazil	155276	1997	2011	17.7	3.6	27.4	3040860	3232.2	10775	NA	9.973
Fortaleza	Brazil	202340	1997	2011	27.0	23.2	29.7	4116276	16704.3	4592	NA	8.278
Goiania	Brazil	134263	1997	2011	24.5	15.1	32.0	2182195	2005.6	8438	NA	10.189
Joao Pessoa	Brazil	53970	1997	2011	27.0	22.2	30.3	1128883	5400.3	4377	NA	8.360
Maceio	Brazil	92525	1997	2011	25.2	20.8	29.1	1407431	14050.4	4050	NA	8.233
Manaus	Brazil	98204	1997	2011	27.2	21.3	32.8	2396744	12595.7	8769	NA	13.901
Natal	Brazil	80753	1997	2011	26.5	21.8	29.8	1307043	9768.8	5215	NA	7.582
Porto Alegre	Brazil	215734	1997	2011	19.7	5.2	33.5	3331241	14503.9	9203	NA	11.688
Recife	Brazil	263492	1997	2011	26.0	21.8	29.4	3303598	6974.5	5356	NA	8.363
Salvador	Brazil	223545	1997	2011	25.6	20.3	29.2	4420845	15829.1	5582	NA	7.312
Sao Luis	Brazil	81852	1997	2011	26.9	23.0	29.9	530385	4602.1	3610	NA	9.351
Sao Paulo	Brazil	916233	1997	2011	20.3	7.8	28.6	19443795	20285.3	15222	NA	15.969
Teresina	Brazil	67174	1997	2011	27.4	22.4	32.5	1040221	8304	3502	NA	10.849
Vitoria	Brazil	51950	1997	2011	24.8	16.2	31.4	1164656	9805.6	11436	NA	9.557
Abbotsford	Canada	27098	1986	2011	10.7	-12.9	29.0	227525	3509.8	35381	0.323	6.666
Calgary	Canada	132883	1986	2011	4.5	-33.9	24.3	1470322	1852.2	57370	0.320	7.012
Edmonton	Canada	154899	1986	2011	4.1	-39.7	29.4	1241391	4129.2	57370	0.320	6.397

Halifax	Canada	71264	1986	2011	6.7	-23.5	26.6	402430	7059.3	30070	0.303	5.963
Hamilton	Canada	113876	1986	2011	8.1	-24.1	29.6	880476	8206.9	41180	0.331	9.126
Kingston	Canada	39019	1986	2011	7.6	-26.0	29.5	159498	4767.1	37343	0.331	7.182
Kitchener-Waterloo	Canada	71387	1986	2011	7.0	-25.2	29.8	617613	2363.8	37343	0.331	7.827
London Ontario	Canada	96799	1986	2011	8.2	-24.7	30.7	508418	1902	37343	0.331	7.584
Montreal	Canada	255272	1986	2009	6.9	-27.3	29.3	4195556	20207.3	34202	0.290	9.195
Ottawa	Canada	136955	1986	2011	6.6	-28.4	29.7	1384629	11165.9	39505	0.331	7.236
Regina	Canada	49530	1986	2011	3.1	-36.5	31.4	264275	5731.9	43880	0.308	9.135
Saint John NB	Canada	44302	1986	2011	5.3	-25.0	25.7	98378	9.3	30251	0.285	5.802
Saskatoon	Canada	56891	1986	2011	2.6	-38.9	32.1	288672	2089.5	43880	0.308	7.047
St. John's NFL	Canada	54328	1986	2011	5.2	-21.3	24.3	164917	244	38856	0.306	5.201
Sudbury	Canada	40907	1986	2011	4.3	-33.8	28.8	159436	946.6	37343	0.331	5.513
Thunder Bay	Canada	35663	1986	2011	3.1	-33.6	27.3	108359	855	37343	0.331	6.946
Toronto	Canada	673074	1986	2011	8.5	-24.7	31.5	6540921	8383.8	41180	0.331	9.109
Vancouver	Canada	329577	1986	2011	10.5	-10.7	28.4	1760166	2466	35381	0.323	6.801
Victoria	Canada	84747	1986	2011	10.2	-10.4	26.7	396171	1162.6	35381	0.323	7.051
Windsor	Canada	65259	1986	2011	10.1	-25.0	31.5	370514	3684.6	37343	0.331	10.088
Winnipeg	Canada	168512	1986	2011	3.2	-38.6	30.9	891345	3834.3	34573	0.298	6.748
Niagara	Canada	98788	1986	2011	9.3	-21.0	30.5	50193	3249.3	37343	0.331	8.594
Oakville	Canada	58991	1986	2011	8.6	-21.9	31.1	193832	1314	37343	0.331	8.478
Oshawa	Canada	72386	1986	2011	7.6	-26.1	30.2	473471	3185	37343	0.331	7.366
Sarnia	Canada	28656	1986	2011	8.8	-25.0	30.7	155084	1725.2	37343	0.331	8.344
Sault Ste. Marie	Canada	28838	1986	2011	5.1	-29.6	27.7	73368	328	37343	0.331	6.380
Chillan	Chile	8619	2008	2014	13.2	0.2	25.4	224550	391.4	9237	0.412	18.262
Santiago	Chile	254344	2008	2014	15.4	2.1	27.5	2906611	6478.2	18919	0.484	27.784
Temuco	Chile	16709	2004	2013	11.5	-1.7	25.9	371080	747	5388	0.449	24.427
Valparaiso	Chile	45790	2004	2013	14.7	6.0	25.7	442938	857.3	14004	0.436	20.076
Anshan	China	30076	2004	2006	11.4	-19.0	36.4	2018775	11072.7	8174	NA	57.669
Beijing	China	74786	2007	2008	14.8	-6.8	30.7	13033601	35522.2	15756	NA	88.334

Fuzhou	China	17142	2004	2006	20.7	3.7	32.2	2482720	29313	7885	NA	33.981
Guangzhou	China	57721	2007	2008	22.8	5.4	33.5	7567325	14240.7	9155	NA	55.218
Hangzhou	China	21743	2002	2004	17.9	-1.4	36.4	3381629	15459.9	10421	NA	69.003
Hong Kong	China	213860	1996	2002	23.7	6.9	33.8	8154579	32289.7	NA	NA	33.637
Lanzhou	China	33877	2004	2008	7.4	-20.0	33.0	3284718	8900.5	3168	NA	68.355
Shanghai	China	172940	2001	2004	17.7	-2.4	34.0	22102012	14681.7	16548	NA	58.743
Shenyang	China	96588	2005	2008	8.2	-22.0	28.0	5905692	46512.8	8174	NA	70.279
Suzhu	China	49633	2005	2008	17.2	-2.8	33.8	2168091	6244.4	10254	NA	71.035
Taiyuan	China	43771	2004	2008	11.2	-13.6	30.5	4178975	44550.9	5098	NA	71.338
Tianjin	China	15857	2005	2008	13.3	-10.5	31.3	7356207	15598.7	14322	NA	103.132
Wuhan	China	62440	2003	2005	17.9	-1.5	35.8	7805706	44001.9	5302	NA	76.964
Wulumqi	China	12281	2006	2007	8.6	-23.7	32.6	2583725	42517.9	4932	NA	65.762
Xian	China	47415	2004	2008	13.4	-8.0	32.0	8705600	870	5050	NA	80.344
Barranquilla	Colombia	91624	1998	2013	27.7	23.0	31.0	2143491	9167.4	6642	NA	22.596
Bogota	Colombia	426297	1998	2013	13.9	10.5	17.4	10219661	5553.1	13728	NA	19.221
Cali	Colombia	188960	1998	2013	24.5	19.3	29.6	3305836	4511.1	9000	NA	13.443
Cartagena	Colombia	53115	1998	2013	28.0	24.3	31.1	1330500	1895.5	7754	NA	20.534
Medellin	Colombia	196543	1998	2013	22.8	18.3	27.5	3648479	7478.2	8672	NA	20.494
Helsinki	Finland	130325	1994	2011	6.2	-22.9	25.5	1115957	3123.6	50028	NA	8.193
Bordeaux	France	53219	2000	2010	13.7	-4.9	31.3	633344	5541	35675	0.281	13.049
Clermont-Ferrand	France	19040	2000	2010	12.0	-8.7	30.3	255206	3235.3	31288	0.308	11.063
Dijon	France	18786	2000	2010	11.3	-8.4	31.1	234755	3673.1	34627	0.311	11.345
Grenoble	France	32738	2000	2010	12.1	-10.1	29.0	352839	8562.7	34516	0.288	14.305
Le Havre	France	24323	2000	2010	11.5	-4.6	29.3	307034	3309.5	31700	0.253	11.816
Lens-Douai	France	36686	2000	2010	11.0	-7.0	28.0	31398	2700	23699	0.296	13.366
Lille	France	90900	2000	2010	11.0	-7.0	28.0	750328	6614.1	29825	0.296	14.682
Lyon	France	77106	2000	2010	12.9	-6.7	31.9	1183817	8235.3	43138	0.288	16.047
Marseille	France	94792	2000	2010	15.6	-2.0	31.6	1494811	24351.3	35479	0.304	11.564
Montpellier	France	26978	2000	2010	15.3	-2.5	30.9	411879	4773.5	30537	0.302	12.454

Nancy	France	28945	2000	2010	10.9	-10.5	29.1	239822	7128.2	30038	0.269	13.239
Nantes	France	43547	2000	2010	12.5	-4.6	31.3	496538	4594.6	34899	0.272	12.326
Nice	France	51959	2000	2010	16.2	1.3	31.4	802120	4528.5	35929	0.304	13.287
Paris	France	455460	2000	2010	12.5	-5.4	32.4	4963177	9705.2	57112	0.343	17.468
Rennes	France	16600	2000	2010	12.2	-4.2	31.8	303972	4053.5	34178	0.270	10.201
Rouen	France	41927	2000	2010	10.6	-6.7	29.3	376556	5440.2	32403	0.253	13.181
Strasbourg	France	34874	2000	2010	11.1	-11.6	28.9	494831	3692.5	35640	0.317	18.919
Toulouse	France	49675	2000	2010	13.8	-4.4	31.8	856555	4031.5	38663	0.272	12.458
Mashhad	Iran	121585	2004	2013	16.0	-14.7	33.3	3401753	348	NA	NA	33.336
Bari	Italy	7234	2005	2007	15.8	-0.8	36.0	566821	2776.3	26504	0.307	15.704
Bologna	Italy	39553	2001	2010	14.6	-8.9	32.3	615303	2613.8	48783	0.302	22.739
Brindisi	Italy	2405	2006	2009	17.6	0.9	36.5	141300	272.4	19035	0.307	16.297
Cagliari	Italy	12716	2001	2010	17.3	3.0	31.9	310005	2832.5	31668	0.311	11.754
Florence	Italy	31078	2001	2009	15.6	-3.0	32.6	1121517	3477.2	45866	0.280	17.218
Genoa	Italy	29546	2003	2006	15.6	-0.7	31.5	925700	2357.4	42574	0.322	16.343
Milan	Italy	102836	2001	2010	14.2	-4.7	31.8	2779161	6806.5	56647	0.304	28.394
Naples	Italy	33443	2006	2009	16.9	-2.1	31.2	2113972	7814.7	22881	0.353	17.015
Padua	Italy	4303	2006	2007	13.5	-3.4	29.1	214125	2300	37619	0.269	24.901
Palermo	Italy	48411	2002	2010	19.1	4.1	39.5	1103675	4066.6	23115	0.369	16.566
Pisa	Italy	7699	2001	2009	15.1	-2.4	29.7	191362	454	35368	0.280	17.619
Rome	Italy	208407	2001	2010	15.9	-0.4	31.0	1704952	2071.2	48656	0.347	17.764
Taranto	Italy	12764	2001	2009	17.4	1.6	32.3	196598	878.8	21654	0.307	13.189
Trieste	Italy	12825	2006	2010	15.7	-2.1	30.4	319202	1983.1	41508	0.261	15.596
Turin	Italy	72835	2001	2010	12.8	-10.7	29.6	1691263	6828.2	40263	0.286	18.978
Venice	Italy	19365	2001	2009	14.0	-6.5	30.1	406224	1383	39532	0.269	20.015
Fukuoka	Japan	185959	1985	2009	17.1	-0.9	32.4	2788289	4305.1	29834	0.314	18.192
Kitakyushu	Japan	206425	1985	2009	17.1	-0.9	32.4	983037	2019	29849	0.314	14.759
Nagoya	Japan	367480	1985	2009	16.0	-2.0	32.7	3806279	7107.9	35725	0.275	13.384
Osaka	Japan	550458	1985	2009	17.0	-0.2	32.9	8849000	6706	32991	0.302	14.829

Sapporo	Japan	256792	1985	2009	9.0	-12.0	30.1	3133786	1675.5	29149	0.327	9.976
Sendai	Japan	126279	1985	2009	12.5	-4.8	31.2	2162562	1325.3	29847	0.292	9.187
Tokyo	Japan	1430094	1985	2009	16.4	0.5	33.1	26174599	4381.5	40165	0.300	14.121
Guadalajara	Mexico	351426	1998	2014	21.1	11.9	28.0	3739589	5445	14193	0.438	17.540
Leon	Mexico	110095	1998	2014	19.5	8.0	33.0	1858626	16759.2	11478	0.412	21.586
Ciudad Juarez	Mexico	91721	1998	2014	20.2	0.4	35.3	2014500	4197.7	13442	0.441	14.996
Comarca Lagunera	Mexico	99640	1998	2014	23.3	3.2	33.8	1488613	33	16709	0.441	14.367
Monterrey	Mexico	286416	1998	2014	22.1	1.2	33.3	2978874	2440.4	25484	0.423	25.058
Puebla-Tlaxcala	Mexico	206429	1998	2014	16.6	5.9	24.0	2489599	7950.1	9023	0.454	21.703
San Luis Petosi	Mexico	77816	1998	2014	17.6	2.4	27.8	685934	1781.6	12304	0.461	17.337
Tijuana	Mexico	96965	1998	2014	17.5	7.0	32.0	2152957	2729.4	15305	0.417	9.517
Toluca de Lerdo	Mexico	129783	1998	2014	13.9	5.7	20.3	819561	1841.7	10234	0.437	26.011
Valley of Mexico	Mexico	1529795	1998	2014	16.4	6.2	24.8	21163226	9700	20216	0.437	27.075
Anenii Noi	Moldova	799	2003	2010	10.5	-22.0	31.6	7678	NA	NA	NA	16.058
Cahul	Moldova	2812	2003	2010	11.3	-20.2	32.3	66196	104.8	NA	NA	18.799
Chisinau	Moldova	54832	2001	2010	10.8	-21.4	32.6	980061	1219.7	NA	NA	17.157
Falesti	Moldova	1463	2003	2010	10.2	-25.0	31.2	17800	3544	NA	NA	18.265
Cebu	Philippines	43855	2006	2010	28.1	23.4	32.0	1206134	20411.5	NA	NA	18.064
Davao	Philippines	44467	2006	2010	28.1	23.1	30.8	1913504	21692.6	NA	NA	17.980
Manila	Philippines	94009	2006	2010	28.8	23.5	33.3	8627575	49045.1	NA	NA	33.341
Quezon	Philippines	92185	2006	2010	28.0	21.8	32.9	4142580	21379	NA	NA	30.103
Busan	South Korea	340551	1992	2010	14.9	-7.1	30.2	5220000	12132.3	19293	NA	23.973
Daegu	South Korea	207086	1992	2010	14.4	-8.5	32.9	3690000	5042.3	17797	NA	25.636
Daejeon	South Korea	105049	1992	2010	13.0	-12.6	31.8	2182330	1785.7	19943	NA	25.069
Gwangju	South Korea	108222	1992	2010	14.1	-9.0	31.3	2136938	4704.4	19193	NA	30.777
Incheon	South Korea	193478	1992	2010	12.5	-14.7	31.5	3825000	16775.5	27111	NA	30.025
Seoul	South Korea	716638	1992	2010	12.8	-15.7	33.0	14694000	4981.2	27111	NA	27.757
Ulsan	South Korea	55914	1992	2010	14.5	-7.4	30.8	1493365	7468.3	58480	NA	25.438
Vitoria	Spain	38581	1990	2010	11.8	-6.4	30.6	243918	880	46938	0.302	6.026

A Coruna	Spain	75572	1990	2010	15.0	2.4	29.0	427924	6299	31029	0.302	9.057
Albacete	Spain	34158	1990	2010	14.4	-10.7	31.6	176147	158.6	26104	0.333	9.372
Alicante	Spain	51253	1990	2010	18.4	4.0	32.2	434759	1792.3	26017	0.334	10.058
Almeria	Spain	41618	1990	2010	19.1	5.4	36.2	214363	704.6	26505	0.344	10.368
Avila	Spain	21147	1990	2010	11.2	-6.3	28.6	59258	255	26310	0.310	8.721
Badajoz	Spain	37579	1990	2010	17.2	1.6	33.9	161211	105.1	22688	0.315	9.000
Barcelona	Spain	365724	1990	2010	16.3	1.4	30.9	4041595	17073	37751	0.325	14.953
Bilbao	Spain	82030	1990	2010	14.8	-0.4	32.2	790963	8601.4	37254	0.302	10.219
Burgos	Spain	41193	1990	2010	10.9	-9.8	29.8	215411	1715.6	36592	0.310	9.069
Caceres	Spain	22017	1990	2010	16.4	0.1	34.1	95855	55	22058	0.315	7.873
Cadiz	Spain	40682	1990	2010	18.6	3.9	32.9	123948	10000	23898	0.344	12.844
Castellon	Spain	41632	1990	2010	17.8	3.4	32.0	171669	1600	30826	0.334	9.897
Ceuta	Spain	9793	1990	2010	18.7	4.1	32.8	82376	4500	29406	0.414	14.107
Ciudad Real	Spain	22650	1990	2010	15.8	-3.2	33.7	74921	260	25023	0.333	10.090
Cordoba	Spain	71562	1990	2010	18.3	1.0	36.3	440336	268	23301	0.344	10.138
Cuenca	Spain	16498	1990	2010	13.3	-4.4	29.8	57032	63	25420	0.333	7.581
Girona	Spain	30810	1990	2010	14.8	-2.3	30.4	97586	2502	37578	0.325	9.626
Granada	Spain	77716	1990	2010	15.7	-3.1	32.7	432393	2861.7	23273	0.344	9.422
Guadalajara	Spain	20115	1990	2010	13.3	-3.8	29.0	67388	417.9	28568	0.333	9.222
Huelva	Spain	43169	1990	2010	18.2	3.8	36.2	167377	1213.6	24038	0.344	12.437
Huesca	Spain	15670	1990	2010	14.1	-5.5	32.0	52347	330	35584	0.315	6.774
Jaen	Spain	36427	1990	2010	17.0	-3.1	35.3	127618	278	22459	0.344	9.156
Leon	Spain	44880	1990	2010	11.1	-8.5	27.9	201915	3376.7	29365	0.310	7.400
Lleida	Spain	37200	1990	2010	15.2	-7.6	30.6	139176	660	38142	0.325	9.348
Logrono	Spain	33058	1990	2010	14.0	-4.9	31.5	175988	2044.4	35454	0.317	10.460
Lugo	Spain	35048	1990	2010	12.1	-4.1	29.1	98560	300	26122	0.302	8.505
Madrid	Spain	576566	1990	2010	15.2	-1.8	32.4	2833937	5676	42081	0.339	10.190
Malaga	Spain	116461	1990	2010	18.7	4.4	34.2	803734	1548.3	23644	0.344	10.246
Melilla	Spain	8727	1990	2010	19.0	3.8	36.1	78476	6380	27198	0.380	16.485

Murcia	Spain	77678	1990	2010	19.0	3.8	36.1	533242	534.3	27096	0.309	12.485
Ourense	Spain	38757	1990	2010	15.1	-1.3	31.5	167137	1265.2	25444	0.302	8.015
Oviedo	Spain	71913	1990	2010	13.3	-1.6	28.4	329224	1210.3	29861	0.305	12.492
Palma Mallorca	Spain	83128	1990	2010	16.7	1.8	32.1	401270	1900	35667	0.339	9.606
Palmas G. Canaria	Spain	85973	1990	2010	21.3	13.7	33.4	383308	3800	28124	0.331	10.397
Pamplona	Spain	56897	1990	2010	13.1	-5.2	31.6	330439	8210.6	40740	0.287	9.630
Pontevedra	Spain	31206	1990	2010	14.8	2.4	30.4	82549	701.04	28515	0.302	8.523
Salamanca	Spain	45440	1990	2010	12.3	-5.1	29.1	239737	3871.9	28186	0.310	8.607
San Sebastian	Spain	66047	1990	2010	13.7	-1.7	30.3	186095	3686	42175	0.302	8.911
Santander	Spain	58362	1990	2010	14.6	1.1	27.8	178465	5100	31229	0.296	10.658
Segovia	Spain	17095	1990	2010	12.4	-6.3	31.4	56660	350	30067	0.310	8.258
Sevilla	Spain	177514	1990	2010	19.5	2.7	36.8	1309044	5283.4	25265	0.344	13.059
Soria	Spain	12594	1990	2010	11.2	-7.2	28.1	39838	150	30407	0.310	7.892
Tarragona	Spain	26417	1990	2010	17.9	-0.1	32.4	152770	2613.6	36984	0.325	10.074
Tenerife	Spain	52474	1990	2010	21.6	13.4	34.3	538000	1400	26993	0.331	8.526
Teruel	Spain	11995	1990	2010	12.2	-10.9	27.8	35396	80	34544	0.315	8.127
Toledo	Spain	31042	1990	2010	15.9	-2.0	34.0	70441	401.6	26710	0.333	8.829
Valencia	Spain	214073	1990	2010	18.5	3.4	33.8	1209304	7356.3	30226	0.334	11.789
Valladolid	Spain	67795	1990	2010	12.9	-4.8	30.9	437595	1618.5	33590	0.310	9.671
Zamora	Spain	20858	1990	2010	13.3	-4.3	30.8	66293	440	25343	0.310	8.495
Zaragoza	Spain	143087	1990	2010	15.7	-4.8	32.9	773209	738.9	34226	0.315	9.948
Stockholm	Sweden	201197	1990	2010	7.2	-21.5	26.8	1885309	8512.5	54811	0.314	5.630
Basel	Switzerland	37607	1995	2013	10.8	-12.4	29.0	585635	7608.6	52566	0.285	12.578
Bern	Switzerland	28193	1995	2013	9.4	-14.2	26.4	259296	1028.8	42997	0.268	13.937
Geneve	Switzerland	26306	1995	2013	11.0	-9.2	28.8	812385	12833.2	54530	0.319	13.434
Lausanne	Switzerland	20810	1995	2013	11.3	-10.3	28.9	249602	3214.4	43253	0.319	13.191
Lugano	Switzerland	28567	1995	2013	12.9	-5.3	28.2	79059	1982.8	42326	0.256	14.045
Luzern	Switzerland	15073	1995	2013	9.9	-11.6	27.1	202491	1380.8	37793	0.280	13.448
St. Gallen	Switzerland	13543	1995	2013	8.6	-14.9	27.0	105858	1915.3	40905	0.256	11.213

Zürich	Switzerland	73539	1995	2013	9.7	-13.5	27.7	895730	4479	61496	0.314	14.085
Kaohsiung	Taiwan	212330	1994	2007	25.2	10.5	32.0	2904247	11691.3	NA	NA	34.994
Taichung	Taiwan	162814	1994	2007	23.6	8.1	32.0	2355387	21559.9	NA	NA	37.864
Taipei	Taiwan	390749	1994	2007	23.2	8.1	33.0	2704974	10000	NA	NA	29.485
London	UK	1325902	1990	2012	11.6	-5.5	29.1	11704709	1353.2	50197	0.386	14.937
Akron, OH	USA	107392	1985	2006	10.1	-26.7	30.8	199110	1368.2	42511	0.367	10.074
Albuquerque, NM	USA	73279	1985	2006	14.2	-14.7	32.2	545852	2212.2	45276	0.410	6.664
Allentown-Bethlehem, PA	USA	61366	1985	2006	11.0	-20.6	31.4	118032	3630.9	52115	0.387	9.260
Atlanta, GA	USA	310249	1985	2006	17.2	-15.0	32.5	420003	1080.6	62350	0.390	8.671
Atlantic City, NJ	USA	49410	1985	2006	12.2	-18.1	32.2	39558	2814.5	52115	0.387	8.420
Austin, TX	USA	69427	1985	2006	20.8	-8.6	35.0	790390	1628.5	50801	0.398	8.922
Bakersfield, CA	USA	88852	1985	2006	18.3	-1.9	36.7	347483	1402.2	47585	0.409	9.072
Baltimore, MD	USA	319591	1985	2006	13.2	-17.8	32.5	620961	2256.3	54045	0.375	10.706
Barnstable-Yarmouth, MA	USA	51337	1985	2006	10.2	-17.2	30.6	68986	317.3	53721	0.402	7.496
Baton Rouge, LA	USA	62561	1985	2006	20.0	-8.6	32.8	229493	1777.7	47800	0.422	8.570
Bergen-Passaic, NJ	USA	239023	1985	2006	13.1	-18.6	34.7	1406342	1300	52115	0.387	9.453
Birmingham, AL	USA	171109	1985	2006	17.2	-13.3	32.2	212237	2306.5	54122	0.394	8.837
Boston, MA	USA	533170	1985	2009	10.9	-16.9	32.2	617594	10917.7	77140	0.402	6.712
Brownsville, TX	USA	36059	1985	2006	23.6	-3.9	33.1	175023	2778.5	44975	0.398	8.032
Buffalo, NY	USA	212201	1985	2006	9.3	-21.1	30.3	261310	3082	39390	0.415	9.847
Canton-Massillon, OH	USA	77288	1985	2006	10.1	-26.7	30.8	105156	921.94	38972	0.367	9.221
Charleston, WV	USA	49105	1985	2006	13.2	-20.6	31.7	51400	1989.3	40344	0.381	7.581
Charlotte, NC	USA	82255	1985	2006	16.2	-12.5	32.5	731424	4985.7	71057	0.390	8.291
Chattanooga, TN	USA	60219	1985	2006	16.1	-15.8	32.8	176588	471.8	40476	0.390	9.007
Chicago, IL	USA	1115158	1985	2006	10.0	-26.7	33.6	2695598	7364.9	55966	0.392	10.498
Cincinnati, OH	USA	171958	1985	2006	12.9	-23.6	33.1	296943	3656.5	48995	0.352	9.944
Cleveland, OH	USA	404057	1985	2006	10.5	-24.4	30.8	396815	475.6	59358	0.367	11.502

Columbia, SC	USA	75994	1985	2006	17.8	-11.1	32.8	129272	526	44301	0.383	8.870
Columbus, OH	USA	159353	1985	2006	11.8	-24.4	31.7	787033	1389.3	53034	0.367	10.129
Dallas, TX	USA	260718	1985	2006	19.1	-13.1	35.8	1197816	2247.5	66880	0.398	9.566
Dayton, OH	USA	108776	1985	2006	11.3	-27.5	32.2	141527	981.9	45561	0.367	10.186
Daytona Beach, FL	USA	107272	1985	2006	21.8	-3.3	31.7	61005	260	37657	0.408	7.996
Denver, CO	USA	182600	1985	2006	10.5	-25.6	30.3	600158	2722.2	60559	0.378	7.762
Des Moines, IA	USA	54488	1985	2006	10.4	-27.5	32.8	203433	587.6	60867	0.343	8.311
Detroit, MI	USA	729077	1985	2006	10.2	-24.4	31.4	713777	1463.8	51958	0.373	10.088
Dutchess County, NY	USA	43055	1985	2006	9.8	-20.3	31.7	297488	374	53680	0.415	7.987
El Paso, TX	USA	73269	1985	2006	18.1	-8.6	36.7	649121	724.8	31581	0.398	6.442
Erie, PA	USA	54723	1985	2006	10.1	-23.9	30.0	101786	4499.5	41271	0.376	10.136
Flint, MI	USA	75484	1985	2006	8.8	-24.4	30.8	102434	1579.9	36264	0.373	8.940
Fort Myers-Cape Coral, FL	USA	88850	1985	2006	24.1	1.9	31.4	154305	671.8	37657	0.408	7.639
Fort Pierce-Port St. Lucie, FL	USA	67004	1985	2006	23.0	-1.4	30.8	417637	1786.3	37657	0.408	7.888
Fort Worth-Arlington, TX	USA	172892	1985	2006	19.1	-7.5	35.6	1198757	1055.6	46770	0.398	9.241
Fresno, CA	USA	104033	1985	2006	18.0	-2.8	38.6	494665	5135.6	30998	0.409	13.428
Ft. Lauderdale, FL	USA	308032	1985	2006	23.0	0.0	30.8	165521	1423	37657	0.408	6.626
Galveston, TX	USA	40680	1985	2006	20.3	-3.6	32.8	47743	764.2	44975	0.398	7.290
Gary, IN	USA	90669	1985	2006	10.0	-27.8	31.7	80294	1325.3	39330	0.352	11.654
Grand Rapids, MI	USA	78804	1985	2006	9.1	-24.4	31.7	188040	2877.7	51211	0.373	9.455
Greensboro, NC	USA	65906	1985	2006	14.9	-14.2	30.6	269666	2000.6	42208	0.390	8.062
Greenville, SC	USA	58344	1985	2006	16.0	-11.7	32.8	58409	792.4	33801	0.383	8.866
Hamilton, OH	USA	49618	1985	2006	12.9	-23.6	33.1	62477	1866.8	38972	0.367	10.041
Harrisburg-Carlisle, PA	USA	49992	1985	2006	12.1	-21.9	32.5	49528	4625.9	50990	0.376	10.056
Hartford, CT	USA	159050	1985	2006	10.3	-18.3	31.1	124775	5311.3	58867	0.376	7.987
Honolulu, HI	USA	75775	1985	2006	25.5	18.1	30.3	337256	2236.1	46159	0.339	4.706
Houston, TX	USA	366340	1985	2006	20.9	-8.1	33.3	2099451	2770.7	64390	0.398	7.828

Indianapolis, IN	USA	149459	1985	2006	11.8	-26.9	32.2	820445	2004.3	62643	0.352	11.700
Jacksonville, FL	USA	124017	1985	2006	20.4	-7.5	32.5	821784	5465.2	42803	0.408	7.950
Jersey City, NJ	USA	103084	1985	2006	13.1	-18.6	34.7	247597	6928.4	52115	0.387	9.934
Kansas City, MO-KS	USA	218933	1985	2006	12.7	-26.7	33.9	459787	1816.6	53019	0.352	8.642
Knoxville, TN	USA	80418	1985	2006	15.1	-21.1	30.3	178874	1358.6	38377	0.395	7.997
Lakeland-Winter Haven, FL	USA	95395	1985	2006	23.1	-0.8	32.2	134584	600.8	37657	0.408	8.120
Lancaster, PA	USA	80724	1985	2006	11.7	-19.7	31.1	59322	6098.2	41271	0.376	10.116
Lansing, MI	USA	37393	1985	2006	8.7	-23.3	30.8	114297	765.5	36264	0.373	8.743
Las Vegas, NV-AZ	USA	182220	1985	2006	20.3	-5.6	41.1	583756	2694.6	45509	0.392	6.243
Little Rock, AR	USA	63901	1985	2006	17.1	-14.2	35.3	193524	123.1	50491	0.372	9.515
Los Angeles, CA	USA	1239036	1985	2006	17.4	6.1	31.1	3792621	2772.5	49493	0.409	11.678
Louisville, KY	USA	139347	1985	2006	14.4	-22.8	34.2	597337	4627	47035	0.352	10.251
Lubbock, TX	USA	35407	1985	2006	16.1	-14.4	35.6	229573	1919.9	44975	0.398	6.511
Madison, WI	USA	48763	1985	2006	8.3	-28.9	32.8	233209	8345.3	61390	0.347	8.577
McAllen-Edinburg-Mission, TX	USA	49998	1985	2006	24.3	-3.6	36.7	291358	1046.5	17996	0.398	8.113
Melbourne-Titusville-Palm Bay, FL	USA	88449	1985	2006	22.6	-0.3	31.9	76068	168.6	37657	0.408	6.839
Memphis, TN	USA	152003	1985	2006	17.3	-16.7	33.9	646889	5044.5	47912	0.372	9.188
Miami, FL	USA	372130	1985	2006	25.0	3.3	31.4	399457	4291.6	44723	0.408	7.930
Middlesex, NJ	USA	110324	1985	2006	11.7	-17.5	32.5	13635	1622.7	52115	0.387	9.232
Milwaukee, WI	USA	232056	1985	2006	9.1	-27.2	33.9	594833	784	54365	0.347	9.891
Minneapolis-St. Paul, MN	USA	241475	1985	2006	7.9	-31.4	32.5	382578	5962.6	59371	0.343	8.454
Mobile, AL	USA	72746	1985	2006	19.7	-9.4	32.8	195111	406.8	34138	0.394	8.329
Monmouth-Ocean, NJ	USA	235036	1985	2006	11.9	-17.2	31.9	642081	522	52115	0.387	8.689
Myrtle Beach, SC	USA	30268	1985	2006	17.8	-8.3	33.3	27109	1524.7	42208	0.390	8.061
Naples, FL	USA	36951	1985	2006	24.0	3.6	30.8	19537	384.2	37657	0.408	7.956
Nashua, NH	USA	51115	1985	2006	8.7	-20.3	30.0	86494	2883.1	43673	0.357	7.963

Nashville, TN	USA	97358	1985	2006	15.5	-20.3	32.5	601222	292.9	56295	0.395	10.189
Nassau-Suffolk, NY	USA	460192	1985	2006	11.4	-16.4	32.2	1506776	608	53680	0.415	8.435
New Haven-Meriden, CT	USA	157415	1985	2006	10.3	-18.3	31.1	129779	4129.8	58867	0.376	8.444
New London, CT	USA	40419	1985	2006	10.7	-16.9	31.1	27620	2973.9	58867	0.376	7.965
New York, NY	USA	1367085	1985	2006	13.3	-16.4	34.4	8175133	13073.6	67702	0.415	9.361
Newark, NJ	USA	220980	1985	2006	13.1	-18.6	34.7	277140	1586.5	52115	0.387	9.934
Newburgh, NY	USA	49890	1985	2006	10.2	-19.7	30.8	28866	6850.5	53680	0.415	9.112
Oakland, CA	USA	325028	1985	2006	15.2	0.6	27.8	390724	6614.8	47585	0.409	7.641
Ocala, FL	USA	58345	1985	2006	21.7	-3.6	32.2	56315	281	37657	0.408	7.834
Oklahoma City, OK	USA	118753	1985	2006	16.0	-17.8	34.7	579999	705.2	44578	0.371	7.951
Omaha, NE	USA	71558	1985	2006	11.0	-26.7	33.3	408958	310	52624	0.343	8.653
Orange County, CA	USA	320343	1985	2006	18.7	5.3	32.8	3010232	1200	47585	0.409	10.985
Orlando, FL	USA	157019	1985	2006	22.8	-2.5	32.2	238300	1776	47691	0.408	6.775
Pensacola, FL	USA	50546	1985	2006	20.3	-9.2	34.2	51923	1260.1	37657	0.408	8.691
Philadelphia, PA-NJ	USA	911888	1985	2006	13.4	-17.5	33.3	1526006	4969.1	54347	0.376	10.998
Phoenix, AZ	USA	386802	1985	2006	24.0	2.5	41.4	1445632	623.8	47691	0.410	8.827
Pittsburgh, PA	USA	317935	1985	2006	11.0	-25.0	30.3	305704	6740.5	63152	0.376	9.750
Portland, ME	USA	46217	1985	2006	8.0	-20.3	29.7	66194	4504.9	36692	0.361	8.257
Portland, OR	USA	210301	1985	2006	12.5	-11.1	29.7	583776	4456.3	56705	0.361	7.150
Providence-Fall River, RI-MA	USA	36108	1985	2006	10.9	-16.7	31.4	178042	3814.9	44355	0.402	7.680
Punta Gorda, FL	USA	37773	1985	2006	23.5	3.6	32.2	16641	492.8	37657	0.408	7.640
Raleigh, NC	USA	58561	1985	2006	15.7	-15.6	32.2	403892	3426.5	52985	0.390	8.290
Reading, PA	USA	72337	1985	2006	15.7	-15.6	32.2	88082	5601.3	41271	0.376	10.099
Riverside-San Bernardino, CA	USA	433285	1985	2006	19.2	3.6	36.7	303871	1522.6	47585	0.409	12.012
Rochester, NY	USA	127040	1985	2006	9.1	-21.1	30.3	210565	3093.6	53680	0.415	8.850
Rockford, IL	USA	46380	1985	2006	9.3	-28.3	31.7	152871	1327.2	47422	0.392	9.201
Sacramento, CA	USA	172136	1985	2006	16.4	-2.5	34.7	466488	3690.7	44379	0.409	8.544

Saginaw, MI	USA	39515	1985	2006	8.6	-23.9	31.1	51508	422.3	36264	0.373	8.960
Salinas, CA	USA	45929	1985	2006	14.4	0.6	27.8	150441	1987.9	47585	0.409	6.568
Salt Lake City, UT	USA	89770	1985	2006	11.7	-20.0	32.8	186440	1232.8	57148	0.320	7.602
San Antonio, TX	USA	186461	1985	2006	20.9	-7.8	34.2	1327407	1662.5	37285	0.398	7.566
San Diego, CA	USA	369956	1985	2006	17.7	6.4	30.8	1307402	10912.3	54256	0.409	8.282
San Francisco, CA	USA	248607	1985	2006	14.5	0.3	30.0	805235	11866.2	74576	0.409	10.202
San Jose, CA	USA	176066	1985	2006	16.3	-1.9	32.2	945942	2744.9	47585	0.409	8.780
Sarasota-Bradenton, FL	USA	151551	1985	2006	23.2	0.8	32.8	107763	1539.6	37657	0.408	7.520
Scranton--Wilkes-Barre--Hazleton, PA	USA	150119	1985	2006	10.0	-22.2	31.9	76089	2610.7	41271	0.376	9.267
Seattle, WA	USA	225451	1985	2006	11.4	-9.7	28.1	608660	11537.1	78444	0.389	7.867
Shreveport, LA	USA	51716	1985	2006	18.9	-10.0	34.2	199311	1323.7	44401	0.422	8.273
Spokane, WA	USA	68681	1985	2006	8.8	-23.6	30.6	208916	1463.8	47485	0.389	6.960
Springfield, MA	USA	94971	1985	2006	10.3	-18.3	31.1	153060	4636.1	53721	0.402	8.069
St. Louis, MO-IL	USA	312923	1985	2006	14.0	-22.8	33.6	319294	3342.2	47422	0.392	10.753
Stamford-Norwalk, CT	USA	142216	1985	2006	11.4	-16.1	30.8	216763	1381.5	58867	0.376	9.001
Stockton-Lodi, CA	USA	82225	1985	2006	16.7	-3.3	36.9	359792	1843.2	47585	0.409	9.701
Syracuse, NY	USA	84451	1985	2006	9.2	-23.6	30.3	145170	2948.3	53680	0.415	9.042
Tacoma, WA	USA	96086	1985	2006	11.8	-9.4	26.7	198397	6320.3	47485	0.389	7.295
Tampa-St. Petersburg-Clearwater, FL	USA	158555	1985	2006	23.0	-1.1	31.7	335709	4664.6	51813	0.408	8.409
Toledo, OH	USA	92004	1985	2006	10.2	-25.0	31.7	287208	1951.3	46146	0.367	9.782
Trenton, NJ	USA	58430	1985	2006	11.7	-16.4	31.7	84913	5001.8	52115	0.387	9.395
Tucson, AZ	USA	131053	1985	2006	21.1	-1.4	37.2	520116	1242.5	33110	0.410	7.328
Tulsa, OK	USA	95475	1985	2006	16.1	-17.5	35.3	391906	1223.2	46716	0.371	8.641
Utica-Rome, NY	USA	53724	1985	2006	8.2	-25.8	28.9	95960	404	53680	0.415	8.121
Ventura County, CA	USA	87603	1985	2006	16.2	3.6	29.2	823318	140	47585	0.409	10.378
Virginia Beach, VA	USA	187233	1985	2006	16.0	-15.0	33.1	437994	1749	42208	0.390	8.683

Washington, DC-MD-VA	USA	141028	1985	2006	14.6	-16.7	33.9	601723	109.8	72136	0.375	9.753
West Palm Beach-Boca Raton, FL	USA	233887	1985	2006	24.3	1.7	31.1	322923	1514.6	37657	0.408	6.412
Wichita, KS	USA	68542	1985	2006	14.0	-22.2	33.9	382368	3077.8	44816	0.352	8.937
Wilmington, DE	USA	76254	1985	2006	12.7	-18.9	31.9	70851	4006.6	66071	0.362	9.175
Worcester, MA	USA	135785	1985	2006	8.8	-21.1	28.9	181045	5528.9	58867	0.376	8.262
York, PA	USA	62767	1985	2006	12.2	-22.2	31.1	43718	2280	41271	0.376	9.097
Youngstown-Warren, OH	USA	86656	1985	2006	9.7	-25.6	30.6	108539	837.5	38972	0.367	8.821
Ho Chi Minh City	Vietnam	101959	2010	2013	28.5	23.0	32.1	6124331	28432.8	NA	NA	23.624
Hue	Vietnam	6214	2009	2013	25.7	14.4	33.9	815000	5066.6	NA	NA	23.500

Chapter 6

Research paper IV

Title: Air conditioning and heat-related mortality: a multi-country longitudinal study.

Author(s): Francesco Sera, Masahiro Hashizume, Yasushi Honda, Eric Lavigne, Joel Schwartz, Antonella Zanobetti, Aurelio Tobias, Carmen Iñiguez, Ana M. Vicedo-Cabrera, Marta Blangiardo, Ben Armstrong, Antonio Gasparrini

Journal/Publisher: Epidemiology.

Type of publication: Research paper.

Stage of publication: Published in Volume 31, Issue 6, November 2020, Pages 779–787.

URL: `http://https://https://journals.lww.com/epidem/Abstract/2020/11000/Air_Conditioning_and_Heat_related_Mortality__A.5.aspx`.

Academic peer-reviewed: Yes.

Candidate's role: See Section 2.3.

Senior author: (Prof. Antonio Gasparrini)

Air Conditioning and Heat-related Mortality
A Multi-country Longitudinal Study

Francesco Sera,[a,b] *Masahiro Hashizume,*[c] *Yasushi Honda,*[d] *Eric Lavigne,*[e,f] *Joel Schwartz,*[g] *Antonella Zanobetti,*[g] *Aurelio Tobias,*[h] *Carmen Iñiguez,*[i,j] *Ana M. Vicedo-Cabrera,*[a,k,l] *Marta Blangiardo,*[m] *Ben Armstrong,*[a,b] *and Antonio Gasparrini*[a,b,n]

Background: Air conditioning has been proposed as one of the key factors explaining reductions of heat-related mortality risks observed in the last decades. However, direct evidence is still limited.

Methods: We used a multi-country, multi-city, longitudinal design to quantify the independent role of air conditioning in reported attenuation in risk. We collected daily time series of mortality, mean temperature, and yearly air conditioning prevalence for 311 locations in Canada, Japan, Spain, and the USA between 1972 and 2009. For each city and sub-period, we fitted a quasi-Poisson regression combined with distributed lag non-linear models to estimate summer-only temperature–mortality associations. At the second stage, we used a novel multilevel, multivariate spatio-temporal meta-regression model to evaluate effect modification of air conditioning on heat–mortality associations. We computed relative risks and fractions of heat-attributable excess deaths under observed and fixed air conditioning prevalences.

Results: Results show an independent association between increased air conditioning prevalence and lower heat-related mortality risk. Excess deaths due to heat decreased during the study periods from 1.40% to 0.80% in Canada, 3.57% to 1.10% in Japan, 3.54% to 2.78% in Spain, and 1.70% to 0.53% in the USA. However, increased air conditioning explains only part of the observed attenuation, corresponding to 16.7% in Canada, 20.0% in Japan, 14.3% in Spain, and 16.7% in the USA.

Conclusions: Our findings are consistent with the hypothesis that air conditioning represents an effective heat adaptation strategy, but suggests that other factors have played an equal or more important role in increasing the resilience of populations.

Keywords: Adaptation; Air conditioning; Longitudinal; Meta-analysis; Multilevel; Temperature

(*Epidemiology* 2020;31: 779–787)

Submitted December 12, 2019; accepted July 21, 2020.

From the [a]Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London, United Kingdom; [b]Centre for Statistical Methodology, London School of Hygiene & Tropical Medicine, London, United Kingdom; [c]Department of Global Health Policy, School of International Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [d]Faculty of Health and Sport Sciences, University of Tsukuba, Tsukuba, Japan; [e]Air Health Science Division, Health Canada, Ottawa, Canada; [f]School of Epidemiology & Public Health, University of Ottawa, Ottawa, Canada; [g]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [h]Institute of Environmental Assessment and Water Research, Spanish Council for Scientific Research, Barcelona, Spain; [i]Department of Statistics and Computational Research, Universitat de València, Spain; [j]CIBERESP, Madrid, Spain; [k]Institute of Social and Preventive Medicine. University of Bern, Bern, Switzerland; [l]Oeschger Center for Climate Change Research. University of Bern, Bern, Switzerland; [m]Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom; and [n]Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, United Kingdom

This work was supported by the Medical Research Council-UK (grant ID: MR/R013349/1), Natural Environment Research Council UK (grant ID: NE/R009384/1), and European Union Horizon 2020 programme (grant ID: 820655).

The authors report no conflicts of interest.

The computer code used to conduct analyses for this article is available from the first author upon request. The mortality data have been obtained through a restricted data use agreement with each national institute (Statistics Canada for Canada, Ministry of Health, Labour and Welfare for Japan, Spain National Institute of Statistics for Spain, and National Center for Health Statistics [NCHS] for the USA), and are therefore not available for public dissemination.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Francesco Sera, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, United Kingdom. E-mail: francesco.sera@lshtm.ac.uk.

ISSN: 1044-3983/20/3106-0779
DOI: 10.1097/EDE.0000000000001241

Epidemiologic studies in various countries have provided evidence of a decrease in mortality risks associated to exposure to heat over the last decades.[1,2] Several mechanisms have been suggested as potential drivers of such attenuation, including physiologic (referred to as acclimatization), behavioral (e.g., clothing), infrastructural (green spaces), and technological (heat warning system).[2–4] However, evidence is still limited, and an appropriate characterization of factors responsible for the attenuation of heat-related risks is still lacking. This information is nonetheless critical for planning effective public health and climate policies.[1–3]

Air conditioning is one of the most straightforward strategies to reduce heat stress, and previous investigations have assessed its role in modifying mortality risks associated to exposure to high temperature using both individual- or aggregated-level designs, although with conflicting results.[5–13] These

studies adopted either a cross-sectional and/or longitudinal design, comparing risks at different air conditioning prevalence between individuals/locations or at different times. However, they faced a number of methodologic challenges. Analyses based on the cross-sectional comparison of subjects or cities with different air conditioning use and prevalence are prone to bias, as other characteristics, such as socioeconomic or climatic conditions, can be responsible for differences in health risks. Longitudinal designs can address this issue, but they need data consistently collected over a long period of time to allow for substantial variation in air conditioning use within each location. More importantly, these studies can be affected by temporal confounding due to concurrent changes in other modifying factors, such as infrastructural changes and access to health care. Finally, the complexity of exposure–response relationships, characterised by non-linearity and temporally delayed effects, presents additional problems in modeling temperature–mortality associations. A recent investigation by Nordio et al [10] partly addressed these issues by comparing estimates from several USA cities over 5 decades, while using flexible exposure–response functions and adjusting for underlying trends. However, that study was performed in a single country, and its estimates of the role of air conditioning can be affected by the lack of separation between spatial and temporal contrasts.

In this contribution, we extend the assessment to a multi-country setting and adopting sophisticated longitudinal designs to control for spatial and temporal confounding. Specifically, the analysis makes use of a unique dataset with time-series data from 331 locations in 4 countries (USA, Japan, Canada, and Spain) in the period 1972–2019, and applies novel 2-stage methods based on multilevel multivariate spatio-temporal meta-regression models.

METHODS

Data

We collated data on mortality, temperature, and air conditioning prevalence from multiple locations in the 4 countries (see eTable 1; http://links.lww.com/EDE/B701). For each location, the data consist of daily counts of all-cause (Canada, Japan, and Spain) or non-accidental (USA) mortality and temperature series in summer months (June to September), and air conditioning prevalence from survey data in multiple years within the study period. Table 1 lists the study locations, the observation period as well as the air conditioning variable and surveys used to derive air conditioning prevalences in the 4 countries included in this study. Across countries, air conditioning prevalence data comes from different surveys with different frequency of reporting (see eAppendix; http://links.lww.com/EDE/B701). More detailed information on the data collected in each country are reported in the eAppendix; http://links.lww.com/EDE/B701.

Statistical Methods

The analytical strategy was based on 3 steps, briefly summarized here and described in detail below. In the first step, each country-specific study interval was split into multiple periods. Then, we fitted separate regression models to obtain estimates of heat–mortality associations for each location and period. In addition, we reconstructed location-specific air conditioning trends and assigned prevalence estimates to each location or period unit. In the second step, we pooled the set of coefficients defining the associations to evaluate changes in heat-related mortality risks by calendar year and air conditioning prevalence, accounting for both within- and between-city variations. Finally, in the third and last step, we used the coefficients of the meta-regression models to derive trends in relative risk (RR) and attributable fractions predicted using observed and alternative scenarios of air conditioning prevalence trends.

Step 1: Estimating Location and Period-specific Air Conditioning Prevalence and Risks

In the first step, for each location, we divided the observation time was divided into specific time intervals. The number and the different periods for each country are reported in eTable 2; http://links.lww.com/EDE/B701. Time intervals have a length of 4 or 5 years. The length of time intervals was chosen a priori to provide enough statistical power to derive period-specific estimates, and enough time points to detect changes over time. For each country and locations, using the original air conditioning data, which was assessed intermittently, we estimated the air conditioning prevalence for each period, as described in the eAppendix; http://links.lww.com/EDE/B701. Briefly, for the USA, Canada, and Spain, we fitted a linear mixed-effects model with a B-spline parameterization of the time variable (years), and city as grouping level. We used best linear unbiased prediction estimates were used to predict yearly air conditioning prevalence in mid-summer (1st of July) in each city of the 3 countries. For Japan, we used the original yearly data and assigned it to mid-summer. To assessed if changes in reporting air conditioning prevalence over time affected the predicted trends, we performed a sensitivity analysis including an indicator that defines pre- and post-periods corresponding to implementation of the new reporting methods (see eAppendix; http://links.lww.com/EDE/B701).

We estimated the location and period-specific temperature–mortality associations through quasi-Poisson regression[14] with distributed lag non-linear models.[15] Based on previous work,[16] we specified the cross-basis function of daily mean temperature using a quadratic B-spline function for the temperature dimension, with 2 internal knots at the 50th and 90th percentiles of the location and period-specific summer temperature distributions, and unconstrained parameterization over lag 0–2. To control for long-term trends and residual seasonality, we included interaction terms between a natural cubic B-spline function with 4 degrees of freedom of the day of the year and indicators of year, along with an indicator of day of the week. We tested these modeling choices in a sensitivity analysis.

TABLE 1. Geographical Boundaries, Observation Period, and Definition of Air Conditioning Prevalence in Each Country

Country	Locations	Period	Air Conditioning Variable	Survey
Canada	20 census metropolitan areas + city of Hamilton	1991–2009	Proportion of dwellings with an air conditioning system (central or with a window or room mounted air conditioning system)	Survey of Household and Energy Use (SHEU)[a] Households and Environment Survey (HES)[b]
Japan	47 prefectures	1972–2009	Proportion of households with 2 or more occupants with air conditioning	Regional statistics database[c]
Spain	52 capital cities	1990–2009	Proportion of family homes with "refrigeration"; and from 2007 proportion of "homes with air conditioning"	Population and Housing Census[d] "Life Conditions" Survey[e]
USA	211 metropolitan areas	1973–2006	Proportion of households in each metropolitan area with central air conditioning	Census of Population[f] American Housing Survey (AHS)[g] Residential Energy Consumption Survey[h]

[a]Estimates at regional level in years 1993, 1997, and 2003.
[b]Estimates at city level in years 2006, 2007, and 2009.
[c]Asahi Newspaper Publishing (2015).
[d]Estimates at city level in years 1991 and 2001.
[e]Estimates at regional level in 2007.
[f]Estimates before 1985 at city level.
[g]AHS use a rotation sampling of cities; data available yearly from 1985.
[h]Used to estimate air conditioning prevalence in northern New England cities.

Step 2: Modeling Spatial and Temporal Variation in risk

The location and period-specific estimates obtained from the quasi-Poisson model in step 1 were then combined using multilevel multivariate spatio-temporal models that consider possible non-independence of estimates within each location.[17] For each location $i = 1,\ldots,m$ and year $t = 1,\ldots,T_i$ (defined as mid-points of periods), we obtained a $k = 4$ length column vector of spline coefficients θ_{it} representing the temperature–mortality association cumulated over lag 0–2 in location *i* and period *t*, and associated $k \times k$ estimated (co)variance matrix S_{it}. The multilevel multivariate spatio-temporal meta-regression model for the multivariate vector response θ_{it} can be written as:

$$\theta_{it} = X_{it}\beta + Z_i b_i + \varepsilon_{it} \quad (1)$$

with $b_i \sim N(0, \psi_1)$, and $\varepsilon_{it} \sim N(0, S_{it})$.

The matrix X_{it} in the meta-regression model in equation 1 included fixed-effect predictors, represented by indicators of country, calendar year, period-specific average and interquartile range of daily mean temperature, in addition to air conditioning prevalence. Temperature variables were selected following previous evidence of their role in modifying heat-related mortality risks, while a linear term for calendar year was included to control for underlying variations in risk unrelated to air conditioning use. We compared the role of different fixed-effect predictors through likelihood ratio test in models fitted with a maximum-likelihood estimator. We included random terms at city- or prefecture-level, represented by indicators Z_i with random coefficients b_i. The random coefficients have unstructured (co)variance matrices Ψ_1. The term S_{it} represents the estimation error within location/period combinations. A restricted maximum-likelihood estimator was used for the final model.

This modeling approach allows investigation of the independent effect of changes over time in air conditioning prevalence on the temperature–mortality association, while adjusting for country- and location-specific trends. Using random terms at location level allows the use of information both within and between locations.

Step 3: Quantifying Heat-related Risks and AC Contribution

The estimated fixed-effects coefficients $\hat{\beta}$ from the multilevel multivariate spatio-temporal meta-regression model (1) fitted in step 2 can be used to predict a set of spline coefficients $\hat{\theta}_{ct}$ that represent pooled heat–mortality association curves for any combination of country, year, and air conditioning prevalence. Specifically, associations were predicted longitudinally or at the end of country-specific study periods, either using observed values of meta-predictors or under specific scenarios of air conditioning prevalence. Results were first reported in terms of country-averaged RR, using country-specific temperature distributions and minimum mortality temperature as references. In addition, we also derived summaries corresponding to estimated mortality fractions (in percentage) attributed to summer heat for each country/sub-period, following a procedure described elsewhere.[18] In brief, we computed the mortality attributable to heat first by summing the temperature-related deaths occurring in days with temperatures higher than the location specific 50th percentile of the summer distribution, and then by dividing this excess by the total number of deaths. We calculated empirical standard error using Monte Carlo simulations, assuming a multivariate normal distribution of the fixed-effects coefficients estimated in step 2.[18]

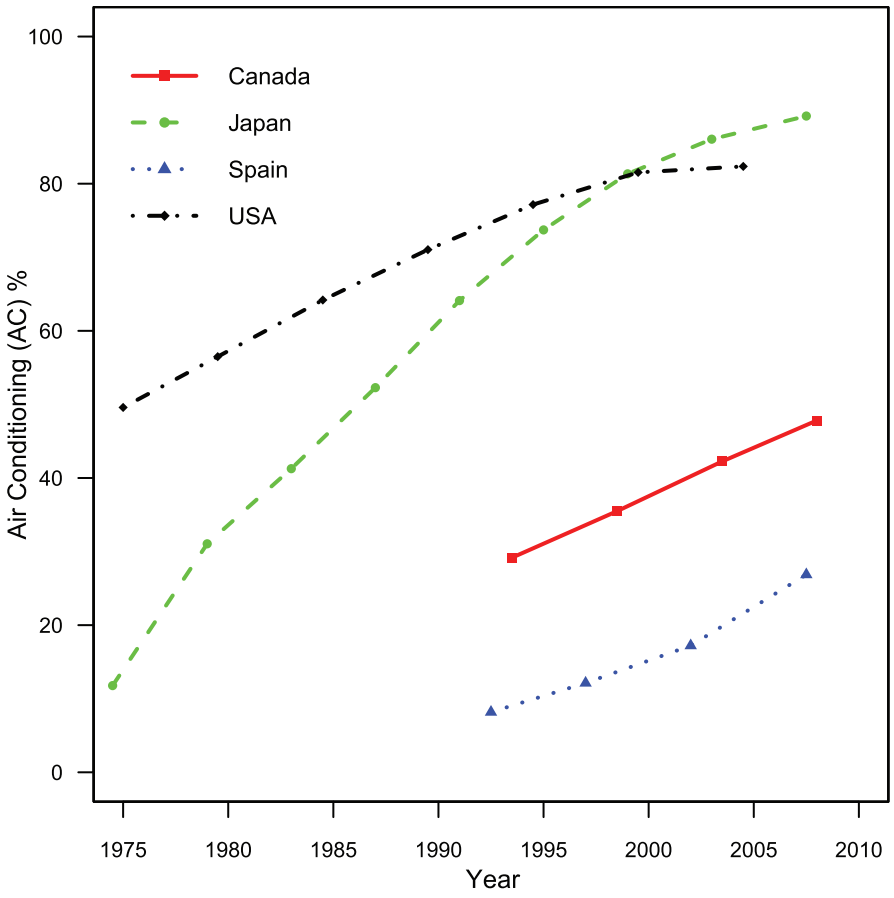


FIGURE 1. Air conditioning (AC) prevalence (%) by year in Canada, Japan, Spain, and the USA.

RESULTS

Data Description

During the study period, more than 23 million deaths were registered in the 331 locations assessed in the 4 countries. On average, air conditioning prevalence increased in all countries (Figure 1), with the highest prevalence at the end of the study period observed in Japan (89.2%), followed by the USA (82.8%), Canada (48.8%), and Spain (26.9%).

Multilevel Multivariate Spatio-temporal Meta-regression Model

The results of meta-regression models with different fixed-effects specifications are shown in eTable 3; http://links.lww.com/EDE/B701. In the final specification of the multilevel multivariate spatio-temporal meta-regression model, air conditioning prevalence shows an independent association with heat-related risks ($P = 0.011$), while accounting for country-specific trends and adjusting also by locations and period-specific average and interquartile range of mean temperature. We did not find strong evidence of a differential effect of air conditioning prevalence between countries ($P = 0.084$). Inspection of distribution of the residuals and their scatter plot versus time and air conditioning prevalence suggested a good fit of the model (see eFigure 3; http://links.lww.com/EDE/B701).

Quantification of the Heat-related Risk and its Trend

Figure 2 represents the changes in the heat–mortality association curves predicted by spatio-temporal meta-regression, at the beginning and end of the study periods in the 4 countries. Japan showed a strong attenuation in risk, with a decline of the RRs across almost all the summer temperature range. The USA and Spain also displayed a decrease in risk, although more evident at highest temperature percentiles. Canada showed little evidence of a reduction in heat-related RR over the observed period.

Table 2 presents air conditioning prevalence, estimated RR at 99th percentile of the temperature distribution versus minimum mortality temperature, and estimated excess mortality by country and calendar year. The trend is consistent with the attenuation in risk, especially in Japan where the RR declined from 1.32 to 1.08 during the period 1975–2007. In the same period, the heat-related excess deaths reduced from 3.57% to 1.10%. A reduction in RR is also evident in the USA and Spain, with a reduction of excess deaths due to heat from 0.54% to 2.78% in Spain, and 1.70% to 0.53% in the USA. In Canada, there was no evidence of reduction of the RR corresponding to the 99th temperature percentile, but we observed a decrease in mortality fraction attributable to heat, from 1.40% to 0.80%, due to an attenuation in risk at

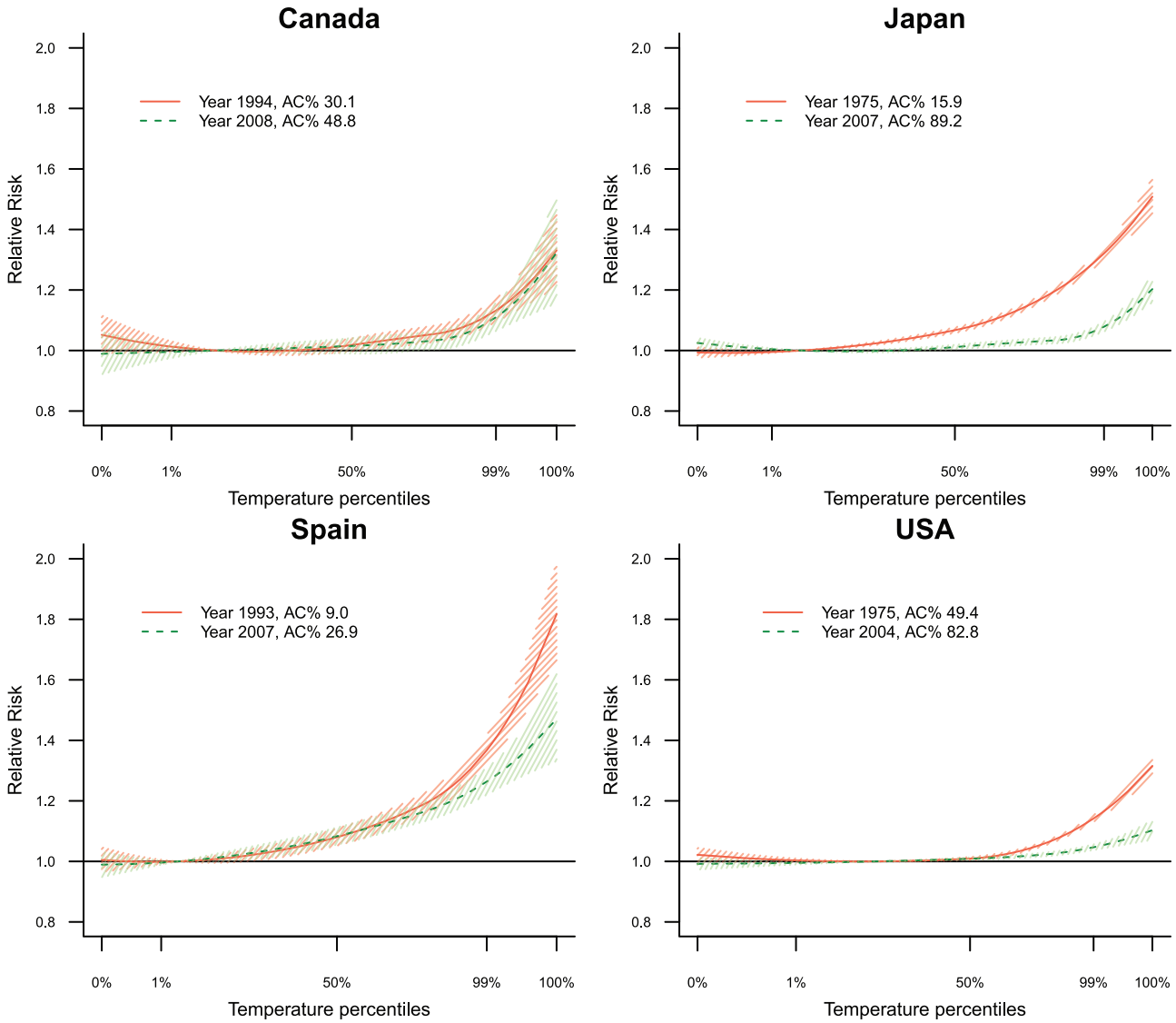


FIGURE 2. Country-average exposure–response curves (in RR) predicted at the beginning and end of the study periods in Canada, Japan, Spain, and the US. The *x*-axis represents relative temperatures in percentiles, but rescaled using the average distribution of absolute temperature across cities in each country.

lower temperature percentiles (90th and 50th), as shown in eFigure 2; http://links.lww.com/EDE/B701.

Temporal changes in temperature-related risks are generated by both variation in air conditioning prevalence and underlying trends due to other factors. To quantify the role of air conditioning, we fixed the calendar year at the end of the study period and calculated the RR at 99th temperature percentile and heat-related mortality fraction for different levels of air conditioning prevalence (Table 3). Results indicate that increasing the AC prevalence from 30% to 80% would be associated with important reduction in heat-related death: 30.2% in the USA, 24.9% in Canada, 20.3% in Japan, and 8.8% in Spain.

Finally, to separate and quantify the contribution of air conditioning prevalence from other time-varying factors in attenuating heat-related risks, we compared the excess mortality under scenarios of observed increase or no change in air conditioning prevalence (Figure 3). The dark and light blue bars represent the excess mortality fraction calculated at the beginning and at the end of the study periods, using the actual air conditioning prevalences, with figures reported in Table 2. The middle blue bar represent instead the excess mortality fraction at the end of study period assuming no change in air conditioning prevalence: the comparison indicates that an increased air conditioning prevalence is responsible for only part of the observed attenuation, corresponding approximately to 16.7% in Canada, 20.0% in Japan, 14.3% in Spain, and 16.7% in the USA. These results suggest that other adaptation factors can be equally and, in some cases, more important for explaining the decreasing trend (see eTable 4; http://links.lww.com/EDE/B701).

TABLE 2. Reconstructed Air Conditioning (AC) Prevalence, RR at 99Th Percentile of the Temperature Distribution Versus Minimum Mortality Temperature, and Attributed Mortality Fraction AF% with 95% Confidence Intervals (CI) by Country and Year

Country	Year	AC%	99th RR (95% CI)	AF% (95% CI)
Canada	1994	30.1	1.13 (1.09, 1.17)	1.40 (1.23, 1.55)
	1998	35.5	1.12 (1.08, 1.16)	1.33 (1.20, 1.44)
	2003	41.9	1.11 (1.08, 1.14)	1.22 (1.05, 1.38)
	2008	48.8	1.11 (1.07, 1.16)	0.80 (0.59, 0.98)
Japan	1975	15.9	1.32 (1.29, 1.34)	3.57 (3.53, 3.61)
	1979	31.1	1.28 (1.26, 1.30)	3.13 (3.10, 3.17)
	1983	41.3	1.24 (1.23, 1.26)	2.83 (2.79, 2.86)
	1987	52.3	1.21 (1.19, 1.22)	2.52 (2.49, 2.56)
	1991	64.1	1.18 (1.16, 1.19)	2.24 (2.20, 2.28)
	1995	73.7	1.15 (1.13, 1.16)	1.90 (1.86, 1.94)
	1999	81.3	1.12 (1.11, 1.14)	1.70 (1.66, 1.75)
	2003	86.0	1.10 (1.08, 1.11)	1.43 (1.39, 1.46)
	2007	89.2	1.08 (1.06, 1.10)	1.10 (1.05, 1.14)
Spain	1993	9.0	1.37 (1.32, 1.42)	3.54 (3.38, 3.69)
	1998	12.9	1.42 (1.37, 1.46)	3.54 (3.42, 3.65)
	2003	19.2	1.35 (1.32, 1.39)	3.51 (3.41, 3.60)
	2007	26.9	1.26 (1.22, 1.31)	2.78 (2.63, 2.92)
USA	1975	49.4	1.14 (1.13, 1.15)	1.70 (1.67, 1.73)
	1979	56.5	1.13 (1.12, 1.14)	1.56 (1.54, 1.58)
	1984	64.1	1.11 (1.10, 1.12)	1.32 (1.30, 1.33)
	1989	71.0	1.09 (1.08, 1.10)	1.09 (1.07, 1.10)
	1994	76.8	1.08 (1.07, 1.09)	0.88 (0.87, 0.90)
	1999	80.7	1.06 (1.05, 1.07)	0.67 (0.65, 0.68)
	2004	82.8	1.05 (1.04, 1.06)	0.53 (0.51, 0.55)

CI indicates confidence interval.

TABLE 3. Predicted Relative Risk (RR) at 99th Temperature Percentile, and Attributed Mortality Fraction (AF%) with 95% Confidence Intervals (CI) Calculated at the End of the Study Period for 4 Scenarios of Air Conditioning Prevalence Levels (30%, 55%, 80%, and 100%) in Canada, Japan, Spain, and the USA. AC% indicates percent with air conditioning.

Country (Year)	AC%	99th RR (95% CI)	AF% (95% CI)
Canada (2008)	30	1.12 (1.07, 1.17)	0.93 (0.75, 1.10)
	55	1.11 (1.06, 1.15)	0.82 (0.63, 1.00)
	80	1.09 (1.05, 1.14)	0.70 (0.51, 0.89)
	100	1.08 (1.03, 1.13)	0.61 (0.40, 0.80)
Japan (2007)	30	1.12 (1.09, 1.14)	1.48 (1.41, 1.54)
	55	1.10 (1.08, 1.12)	1.33 (1.28, 1.37)
	80	1.08 (1.07, 1.10)	1.18 (1.13, 1.22)
	100	1.07 (1.06, 1.09)	1.06 (1.01, 1.10)
Spain (2007)	30	1.26 (1.22, 1.31)	2.86 (2.70, 2.99)
	55	1.24 (1.20, 1.29)	2.73 (2.58, 2.87)
	80	1.23(1.18, 1.28)	2.61 (2.45, 2.77)
	100	1.21 (1.16, 1.27)	2.50 (2.32, 2.66)
USA (2004)	30	1.07 (1.05, 1.09)	0.82 (0.79, 0.84)
	55	1.06 (1.05, 1.07)	0.69 (0.67, 0.71)
	80	1.05 (1.04, 1.06)	0.57 (0.55, 0.59)
	100	1.04 (1.03, 1.05)	0.47 (0.45, 0.49)

DISCUSSION

Our results on air conditioning prevalence in Japan, the USA, Canada, and Spain are consistent with the hypothesis that air conditioning reduces heat-related mortality. This reduction occurs on top of variations in heat-related health risks possibly associated with planned and unplanned adaptation processes other than air conditioning use. These independent adaptation pathways were quantified and compared using alternative scenarios of air conditioning prevalence and underlying temporal trends. These scenarios indicate that while the increase in air conditioning use is associated with a reduction in heat-related mortality, this only explains a part of the decline in risk experienced in some countries, and other adaptation pathways have had a more important role in reducing the health burden.

Our results are consistent with published epidemiological investigations that have reported a substantial attenuation of heat-related health risk.[1,2,19] In particular, similar declining trends were observed in the USA,[6,7,10,16,20–24] Japan,[8,9,25] Spain,[26] and Canada.[16] Similar declining trends were also observed in Sweden,[27] Austria,[28] the United Kingdom,[29,30] Netherlands,[31] 9 European cities,[32] and Korea,[33,34] but not in China.[35]

Previous studies have evaluated the protective effect of air conditioning on heat-related risks. Some assessments used cohort[12] and case-control study designs,[13] and suggested a role of AC in reducing the heat-related mortality risks in the USA. These studies were followed by 2-stage studies in which the first-stage estimates obtained through case-only[11] or time-series analyses[5] in multiple cities were combined using meta-regression models with air conditioning prevalence as a contextual variable. These studies confirmed the protective effect of air conditioning in the USA, but were prone to ecologic confounding as the selected cities can differ by other unmeasured characteristics (e.g., demographic, socioeconomic, and infrastructural) related to health risk. More recent studies in the USA and Japan used a longitudinal design to disentangle the effect of air conditioning as behavioral adaptive measure. In the USA, 2 studies found an independent protective effect of air conditioning,[5,6] but Bobb et al[7] observed no evidence of protective effect. The longitudinal study of Nordio et al[10] reported independent protective effects of air conditioning while controlling for region, time trend, and mean summer temperature, using spline models in individual cities and a meta-regression approach. The 2 longitudinal studies conducted in Japan did not find evidence consistent with an independent protective effect of air conditioning over the declining heat-related risk trend.[8,9] Differences on previous studies results can be partly explained by low statistical power, as these investigations were conducted in a single country and/or the temperature–mortality curve was summarized using simplified indices. Moreover, these studies did not jointly

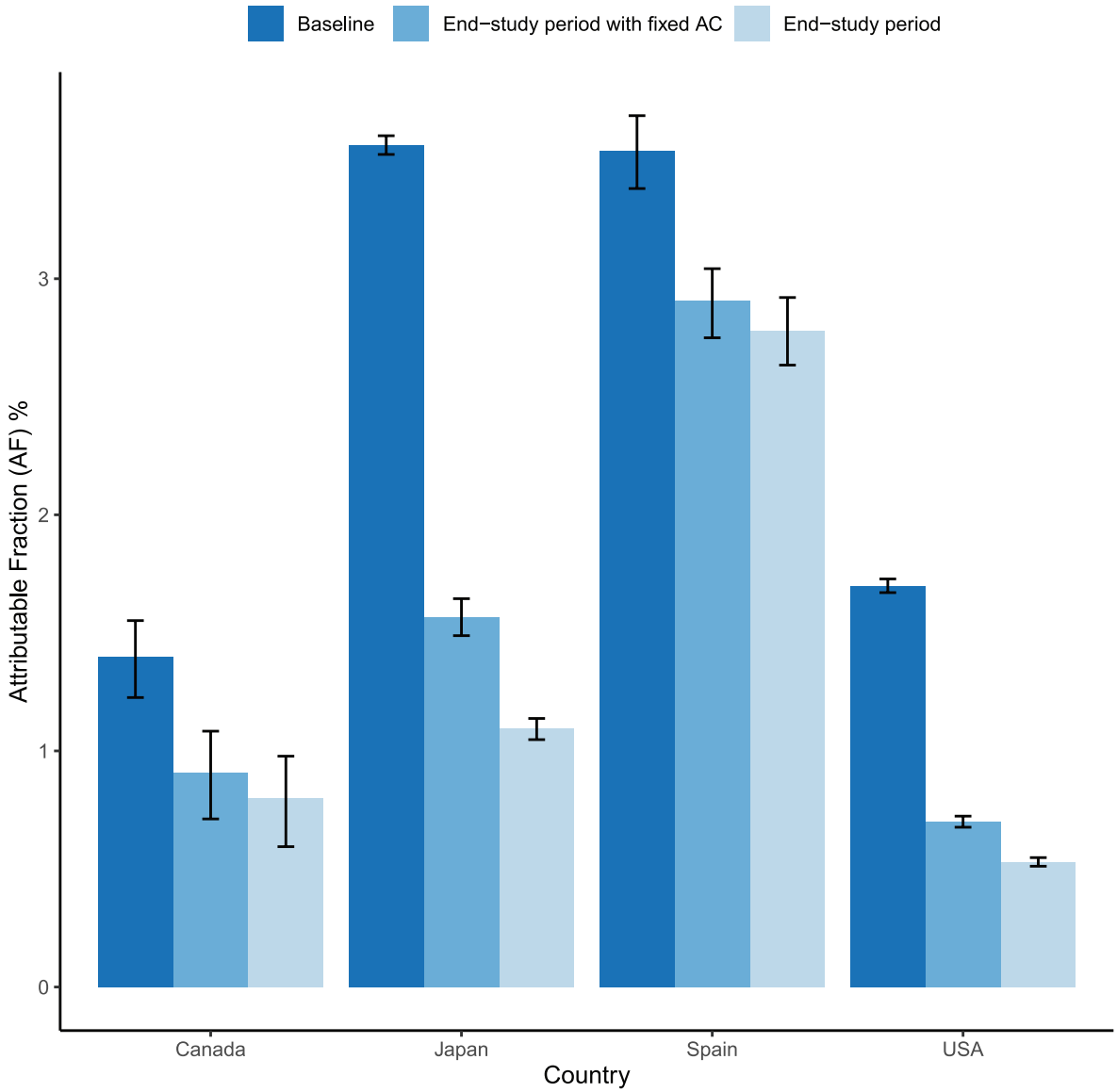


FIGURE 3. Excess mortality associated to heat reported as attributable fraction (AF%) estimated at the beginning (baseline, dark) and end of the study period assuming no change (end-study period with fixed air conditioning, medium) or with the observed change (end-study period, light) in air conditioning (AC) prevalence.

consider the longitudinal and spatial structure of the data, and the non-independence of the observations within locations.

Our study has several strengths. First, we used distributed lag non-linear modeling techniques to estimate the heat–mortality association. This modeling framework helps avoid biases due to simplification of the exposure–response association and considers possible lagged effects of heat on mortality.[15] Second, we were able to collect mortality, temperature and air conditioning data for 331 locations in 4 countries for a period of 4 decades. This provided large variability in air conditioning prevalence both within and across locations, offering sufficient statistical power to isolate the impact on modifying heat–mortality relationships. Third, we used a study design based on both spatial and longitudinal comparison, reducing the chance of ecologic bias and temporal confounding due to concurrent changes in other modifying factors, such as socioeconomic conditions and access to health care. The spatial component provides increased variability in response and exposure, while the longitudinal design compares variations in risk within a location. Finally, we used novel multilevel multivariate spatio-temporal meta-regression models that allow disentangling of the reduction in heat-related risk associated to the increase in air conditioning prevalence from underlying trends due to other adaptation pathways, while at the same time correctly accounting for correlations between repeated measures taken within the same location.[17]

We must acknowledge some limitations. First, the results of our study refer to developed countries with predominantly

temperate or continental climates. Caution should be used when extrapolating results to low-income countries, which are characterized by different climatic, sociodemographic, and development conditions, and where technology-based adaptation measures, such as increasing air conditioning use, may be problematic as many low-income countries already experience chronic shortages of power.[2] Second, we reconstructed air conditioning prevalence along the past decades by applying smoothing techniques to irregular survey data from multiple sources. However, additional analyses described in the eAppendix; http://links.lww.com/EDE/B701 show that results are robust to this filling-up procedure. The results of the sensitivity analysis suggest that the smoothing process could have introduced some error, although it is unlikely that this is correlated with the estimated period-specific risk, and therefore can probably be assumed as random. Third, our air conditioning variable is defined as presence of air conditioning units or central air conditioning at home, but does not capture its actual use. Moreover, this measure is not informative about air conditioning use in other environments, such as on public transport, stores, workplaces, and public areas. This may induce some additional problems in the interpretation of the results.

The analysis of factors related to changes in susceptibility to temperature-related mortality is critical to inform health and climate policies. Air conditioning is a solution to regulate ambient indoor temperatures and lower the heat stress imposed on the human thermoregulatory function,[36] and it represents one of the most cited behavioral adaptation strategy to climate change.[37] The results of our analysis confirm that air conditioning is an effective adaptive measure and have contributed to reduce the burden of heat-related mortality. According to our estimates in the USA and Japan, nearly 0.09% and 0.32% of deaths during summer months were delayed by increasing the air conditioning prevalence level to more than 80%, respectively. In these countries, the air conditioning market seems to have reached a plateau, but the heat-related mortality is still substantial. However, the quantitative comparison of the contribution of increase in air conditioning prevalence, and the independent attenuation of the risk reported in Figure 3, suggest that other adaptation pathways can be equally or even more effective in reducing the health burden. In Spain and Canada, the delayed deaths during summer months were both 0.05%, suggesting a further margin on reduction of heat-related mortality, especially in Spain where the reported air conditioning prevalence reaches only 30% in 2009. In addition, increasing air conditioning use has also important negative consequences, including capital and energy cost, carbon and pollution-generating energy demand, and contribution to the heat-island effect.[2] However, the current rapid transition of electricity generation to carbon zero sources is likely to ameliorate the pollution impact in the next few decades. A quantitative assessment of health and economic impacts of this and other adaptive changes is critical for generating plausible scenarios of potential mitigation and adaptation benefit and costs.

In conclusion, in this study, we found a reduction over time of the heat-related health risk in Japan, the USA, and Spain. Air conditioning prevalence was factor that independently explained part of the decrease in heat-related deaths, although we estimated that other adaptive strategies accounted for a larger proportion of the attenuation. These results can be used to inform policy measures based at individual, community, and international level, and to improve and extend projections of future heat impacts on human health.

REFERENCES

1. Arbuthnott K, Hajat S, Heaviside C, Vardoulakis S. Changes in population susceptibility to heat and cold over time: assessing adaptation to climate change. *Environ Health*. 2016;15(suppl 1):33.
2. Kinney PL. Temporal trends in heat-related mortality: implications for future projections. *Atmosphere*. 2018;9:409.
3. Hondula DM, Balling RC, Vanos JK, Georgescu M. Rising temperatures, human health, and the role of adaptation. *Curr Climate Change Rep*. 2015;1:144–154.
4. IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability*. IPCC Working Group II, Cambridge: Cambridge University Press and New York, NY; 2014.
5. Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*. 2009;20:205–213.
6. Barreca A, Clay K, Deschenes O, Greenstone M, Shapiro JS. Adapting to climate change: the remarkable decline in the US temperature–mortality relationship over the twentieth century. *J Polit Econ*. 2016;124:105–159.
7. Bobb JF, Peng RD, Bell ML, Dominici F. Heat-related mortality and adaptation to heat in the United States. *Environ Health Perspect*. 2014;122:811–816.
8. Chung Y, Yang D, Gasparrini A, et al. Changing susceptibility to non-optimum temperatures in Japan, 1972-2012: the role of climate, demographic, and socioeconomic factors. *Environ Health Perspect*. 2018;126:057002.
9. Ng CFS, Boeckmann M, Ueda K, et al. Heat-related mortality: effect modification and adaptation in Japan from 1972 to 2010. *Glob Environ Change*. 2016;39:234–243.
10. Nordio F, Zanobetti A, Colicino E, Kloog I, Schwartz J. Changing patterns of the temperature-mortality association by time and location in the US, and implications for climate change. *Environ Int*. 2015;81:80–86.
11. Ostro B, Rauch S, Green R, Malig B, Basu R. The effects of temperature and use of air conditioning on hospitalizations. *Am J Epidemiol*. 2010;172:1053–1061.
12. Rogot E, Sorlie PD, Backlund E. Air-conditioning and mortality in hot weather. *Am J Epidemiol*. 1992;136:106–116.
13. Semenza JC, Rubin CH, Falter KH, et al. Heat-related deaths during the July 1995 heat wave in Chicago. *N Engl J Med*. 1996;335:84–90.
14. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. *Int J Epidemiol*. 2013;42:1187–1195.
15. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med*. 2014;33:881–899.
16. Gasparrini A, Guo Y, Hashizume M, et al. Temporal variation in heat-mortality associations: a multicountry study. *Environ Health Perspect*. 2015;123:1200–1207.
17. Sera F, Armstrong B, Blangiardo M, Gasparrini A. An extended mixed-effects framework for meta-analysis. *Stat Med*. 2019;38:5429–5444.
18. Gasparrini A, Leone M. Attributable risk from distributed lag models. *BMC Med Res Methodol*. 2014;14:55.
19. Sheridan SC, Allen MJ. Temporal trends in human vulnerability to excessive heat. *Environ Res Lett*. 2018;13:043001.
20. Barnett AG. Temperature and cardiovascular deaths in the US elderly: changes over time. *Epidemiology*. 2007;18:369–372.
21. Davis RE, Knappenberger PC, Michaels PJ, Novicoff WM. Changing heat-related mortality in the United States. *Environ Health Perspect*. 2003;111:1712–1718.
22. Donaldson GC, Keatinge WR, Näyhä S. Changes in summer temperature and heat-related mortality since 1971 in North Carolina, South Finland, and Southeast England. *Environ Res*. 2003;91:1–7.

23. Guo Y, Barnett AG, Tong S. High temperatures-related elderly mortality varied greatly from year to year: important information for heat-warning systems. *Sci Rep*. 2012;2:830.
24. Petkova EP, Gasparrini A, Kinney PL. Heat and mortality in New York City since the beginning of the 20th century. *Epidemiology*. 2014;25: 554–560.
25. Onozuka D, Hagihara A. Variation in vulnerability to extreme-temperature-related mortality in Japan: a 40-year time-series analysis. *Environ Res*. 2015;140:177–184.
26. Martínez-Solanas È, Basagaña X. Temporal changes in temperature-related mortality in Spain and effect of the implementation of a Heat Health Prevention Plan. *Environ Res*. 2019;169:102–113.
27. Åström DO, Forsberg B, Edvinsson S, Rocklöv J. Acute fatal effects of short-lasting extreme temperatures in Stockholm, Sweden: evidence across a century of change. *Epidemiology*. 2013;24:820–829.
28. Matzarakis A, Muthers S, Koch E. Human biometeorological evaluation of heat-related mortality in Vienna. *Theor Appl Climatol*. 2011;105:1–10.
29. Carson C, Hajat S, Armstrong B, Wilkinson P. Declining vulnerability to temperature-related mortality in London over the 20th century. *Am J Epidemiol*. 2006;164:77–84.
30. Christidis N, Donaldson GC, Stott PA. Causes for the recent changes in cold-and heat-related mortality in England and Wales. *Clim Change*. 2010;102:539–553.
31. Ekamper P, Van Poppel F, Van Duin C, Garssen J. 150 years of temperature-related excess mortality in the Netherlands. *Demogr Res*. 2009;21:385–426.
32. de' Donato FK, Leone M, Scortichini M, et al. Changes in the effect of heat on mortality in the last 20 years in nine European cities. Results from the PHASE Project. *Int J Environ Res Public Health*. 2015;12:15567–15583.
33. Ha J, Kim H. Changes in the association between summer temperature and mortality in Seoul, South Korea. *Int J Biometeorol*. 2013;57:535–544.
34. Lee W, Choi HM, Kim D, Honda Y, Guo Y-LL, Kim H. Temporal changes in mortality attributed to heat extremes for 57 cities in Northeast Asia. *Sci Total Environ*. 2018;616:703–709.
35. Yang C, Meng X, Chen R, et al. Long-term variations in the association between ambient temperature and daily cardiovascular mortality in Shanghai, China. *Sci Total Environ*. 2015;538:524–530.
36. Deschenes O. Temperature, human health, and adaptation: a review of the empirical literature. *Energy Econ*. 2014;46:606–619.
37. O'Neill MS. Air conditioning and heat-related health effects. *Appl Environ Sci Public Health*. 2003;1:9–12.

Additional information on data collection

Japan

We collected data for each of the 47 prefectures in Japan in the period 1972-2009. (1) Daily counts of deaths from all causes were extracted from a computerised death certificate database maintained by the Ministry of Health, Labour and Welfare of Japan. We derived daily mean temperature by averaging hourly measurements provided by the Japan Meteorological Agency for a single weather station in the capital city of each prefecture. We obtained prefecture-specific prevalence data of AC for households with two or more occupants in each year from a regional statistics database. (1)

USA

We collected data for 211 metropolitan areas in the USA with a nationwide geographic distribution in the period 1973-2006. (2) Metropolitan areas were composed of single or multiple counties. All cause daily mortality excluding any death from accidental causes (ICD-code 10^{th} revision: V01-Y98, ICD-code 9^{th} revision: 001-799) were calculated from individual mortality data obtained from the National Center for Health Statistics (NCHS). Daily mean temperature was obtained from the airport weather station nearest to each city (National Oceanic and Atmospheric Administration [NOAA]). We estimated percentage of households in each city with central air conditioning (AC) by combining county-level or metropolitan area-level data. For years in 1970's and 1980's, county-level AC data were gathered from the USA Census of Population. For later years, we used metropolitan area data from the American Housing Survey (AHS). As the AC prevalence shows a strong (north to south) geographical pattern in the USA, for cities not included in the AHS we used the nearest metropolitan area with available data. For northern New England cities, we used regional level data from the "US Energy Information Administration, Office of Energy Consumption Residential Energy Consumption Survey".

Canada

We collected data from 20 census metropolitan areas (CMA) and the city of Hamilton in the period 1986-2009. All-cause daily mortality was obtained from Statistics Canada through access to the Canadian Mortality Database. Mean daily temperature, computed as the 24-hour average based on hourly measurements, were obtained from Environment Canada. A single weather station was selected for each city using the airport monitoring station located closest to the CMA centre. Proportion of dwellings with an air conditioning system (central or with a window or room mounted air conditioning system) was available for years 1993, 1997, 2003, 2006, 2007, 2009. The information is available at regional level until 2003 (Survey of household & energy use (SHEU)), and from 2006 at city level (Households and environment survey (HES)).

Spain

We collected data from the 52 capital cities in the period 1990-2014. All-cause daily mortality was obtained from Spain National Institute of Statistics. Mean daily temperature, computed as the 24-hour average based on hourly measurements, was obtained from Spain National Meteorology Agency. A single weather station, located within the urban area or at the near airport, was selected for each city. Single-day missing values were imputed as the average of the days before and after. For periods longer than two days, no imputation was done. AC

prevalence data were available for three years, in 1991, 2001 and 2007. Data for 1991 and 2001 available at city level come from the National Population and housing census and refers to number of family homes with “refrigeration". Data for year 2007 available at regional level (17 Regions) comes from “Life conditions" survey and refers to "homes with air conditioning".

Derivation of AC trends

For each country and location, using the original AC data, we estimated the AC prevalence for each sub-period. Briefly, for the USA, Canada and Spain we fitted a linear mixed-effects model with a B-spline parametrisation of the time variable (years), and city as grouping level. (3) The B-spline variables were used as fixed and random effects, borrowing information across locations, and allowing the random terms to model city-specific deviations in the trend. Best linear unbiased prediction (BLUP) estimates were used to predict yearly AC prevalence in mid-summer (1st of July) in each city of the three countries. For Japan, we used the original yearly data, and assigned it to mid-summer.

The original prevalence data for each country, location and sub-period for all the four countries, together with the estimated smoothed trends, are reported in eFigures 1 (a)-(d).

Sensitivity Analyses

Across countries AC prevalence data comes from different surveys with different frequency of reporting. To assess if changes in how AC prevalence was collected and reported affect our results we performed a sensitivity analysis in the linear mixed-effects models fitted for deriving trends in US and Canada. In particular we added an indicator that defines pre/post periods corresponding to implementation of the new reporting methods, using as threshold the year 1980 for US (transition from census (counties) to AHS survey (metropolitan areas), and the year 2003 for Canada (transition from regional to city level data). The parameters for these indicators are not significant at 95% (p=0.11 and p=0.10, and indeed their inclusion results in negligible changes in predicted AC prevalence).

AC data from cities in the USA come from different sources (USA Census of Population, American Housing Survey (AHS) and Residential Energy Consumption Survey), which were collected with different designs and frequency. We performed a sensitivity analysis to assess if the effect of AC in USA was different in cities with (n = 105) and without (n = 106) AHS data. Briefly, we applied multilevel multivariate meta-analytic model with calendar year, AC prevalence, average and range of mean temperature as fixed effects and city as random term. An indicator variable was introduced to represent cities with and without AHS data with an interaction term with AC prevalence to assess the AC effect is modified by the two group of cities. The results of this analysis show that the AC effect is not modified (p=0.529) by the group of cities.

1. Chung Y, Yang D, Gasparrini A, et al. Changing Susceptibility to Non-Optimum Temperatures in Japan, 1972-2012: The Role of Climate, Demographic, and Socioeconomic Factors. *Environ Health Perspect* 2018;126(5):057002.
2. Nordio F, Zanobetti A, Colicino E, et al. Changing patterns of the temperature–mortality association by time and location in the US, and implications for climate change. *Environment international* 2015;81:80-6.
3. Ruppert D, Wand MP, Carroll RJ. *Semiparametric regression*. Cambridge university press; 2003.

Additional tables

eTable 1(a). Total number of deaths during summer months, daily mean temperature (Celsius degree) and average AC prevalence by 21 study locations in Canada during the study period 1986-2009.

City	Deaths	Daily Mean Temperature	Average AC prevalence
Abbotsford	7838	17.0	16.8
Calgary	38533	14.3	13.4
Edmonton	45066	15.4	11.3
Halifax	20661	16.9	7.8
Hamilton	33352	18.8	67.0
Kingston	11469	18.8	61.8
Kitchener-Waterloo	20230	17.9	64.0
London Ontario	28166	18.8	65.8
Montreal	80028	18.9	32.0
Ottawa	39664	18.7	64.3
Regina	14581	16.3	31.1
Saint John NB	12648	15.3	31.7
Saskatoon	16794	15.8	55.5
St. John's NFL	15741	13.9	7.4
Sudbury	12019	16.7	7.5
Thunder Bay	10529	15.3	53.0
Toronto	198640	19.4	66.5
Vancouver	94778	16.8	10.5
Victoria	24457	15.8	11.3
Windsor	18810	21.0	69.8
Winnipeg	49069	17.1	34.3

eTable 1(b). Total number of deaths during summer months, daily mean temperature (Celsius degree) and average AC prevalence by 47 study locations in Japan during the study period 1972-2009.

Prefecture	Deaths	Daily Mean Temperature	Average AC prevalence
Aichi	452427	25.0	74.7
Akita	124440	21.7	30.3
Aomori	138564	20.1	18.5
Chiba	353653	24.0	63.7
Ehime	145620	25.2	62.1
Fukui	73010	24.2	68.1
Fukuoka	390851	25.5	70.5
Fukushima	189597	22.5	33.6
Gifu	167354	25.2	62.1
Gunma	163532	23.7	59.4
Hiroshima	238543	25.1	69.8
Hokkaido	467270	19.2	6.0
Hyogo	429740	25.2	73.8
Ibaraki	226688	22.4	53.4
Ishikawa	99811	24.0	63.4
Iwate	131879	20.4	20.0
Kagawa	95519	25.3	74.8
Kagoshima	185235	26.4	53.6
Kanagawa	479908	24.0	66.1
Kochi	90113	25.3	58.2
Kumamoto	168999	25.8	61.3
Kyoto	210622	25.4	79.0
Mie	156597	24.8	68.8
Miyagi	170173	21.3	33.3
Miyazaki	104213	25.6	55.0
Nagano	197618	22.3	27.7
Nagasaki	146701	25.5	59.6
Nara	104561	24.2	76.0
Niigata	228737	23.4	58.2
Oita	119665	24.9	56.4
Okayama	174873	25.4	71.5
Okinawa	78148	27.8	57.9
Osaka	625918	26.0	83.4
Saga	83179	25.5	66.6
Saitama	382546	24.0	73.7
Shiga	94724	24.2	67.8
Shimane	83086	23.9	56.5
Shizuoka	279169	24.6	58.5
Tochigi	158398	22.9	55.0

Tokushima	83947	25.2	66.3
Tokyo	839158	24.7	74.2
Tottori	60834	24.1	59.6
Toyama	104314	23.5	63.4
Wakayama	108623	25.5	70.5
Yamagata	124152	21.9	38.8
Yamaguchi	155872	24.6	62.0
Yamanashi	75953	24.0	42.1

eTable 1(c). Total number of deaths during summer months, daily mean temperature (Celsius degree) and average AC prevalence by 52 study locations in Spain during the study period 1990-2014.

City	Deaths	Daily Mean Temperature	Average AC prevalence
A Coruna	16435	18.9	4.6
Albacete	7657	23.1	15.0
Alicante	17524	24.8	23.2
Almeria	9622	25.2	26.2
Avila	3293	19.0	5.7
Badajoz	7356	24.7	28.2
Bilbao	25981	19.8	6.1
Barcelona	119966	23.1	19.4
Burgos	10884	18.1	3.0
Cadiz	9221	23.8	18.0
Caceres	4585	24.5	29.3
Ciudad Real	4078	24.7	21.8
Ceuta	3668	23.3	8.4
Cordoba	18015	26.4	39.8
Castellon	8906	24.4	20.3
Cuenca	3405	21.7	12.2
Guadalajara	3780	21.6	17.5
Girona	4579	22.1	19.3
Granada	15302	23.7	26.4
Huelva	8310	24.6	19.7
Huesca	3506	22.3	16.6
Jaen	6148	25.2	35.7
Leon	9530	18.2	3.0
Logrono	8150	21.4	7.2
Lleida	7641	23.5	22.0
Lugo	6118	17.5	3.2
Malaga	32155	24.9	21.5
Madrid	194623	23.7	21.6
Melilla	3100	24.6	11.7

Murcia	19671	26.1	35.1
Ourense	7223	21.5	4.1
Oviedo	14887	18.1	5.0
Palmas G. Canaria	20947	23.9	3.9
Palma Mallorca	20727	23.7	22.6
Palencia	5697	19.5	3.9
Pamplona	11776	20.1	7.6
Pontevedra	4520	19.6	3.7
Segovia	3706	20.3	4.1
Salamanca	10890	19.8	4.7
San Sebastian	12657	18.5	6.5
Santander	13103	19.3	6.3
Soria	2366	18.7	3.8
Sevilla	42071	26.9	42.5
Teruel	2328	20.4	15.1
Tenerife	11999	24.7	6.3
Toledo	3927	24.8	27.5
Tarragona	6777	25.1	19.0
Vitoria	11886	18.1	6.0
Valladolid	18921	20.7	6.0
Valencia	51853	24.8	28.1
Zamora	4517	21.1	2.8
Zaragoza	42089	23.8	21.2

eTable 1(d). Total number of deaths during summer months, daily mean temperature (Celsius degree) and average AC prevalence by 211 study locations in USA during the study period 1973-2006.

City	Deaths	Daily Mean Temperature	Average AC prevalence
AUGUSTA (GA)	16328	25.4	85.6
AKRON (OH)	50880	20.3	54.6
ALBANY (NY)	28663	19.7	59.2
ALBUQUERQUE (NM)	29823	23.6	61.6
ALLENTOWN (PA)	27587	21.3	73.2
ANCHORAGE (AK)	5904	13.5	1.0
ANAHEIM (CA)	137811	22.6	47.7
ANN ARBOR (MI)	14962	20.3	64.7
ANNANDALE (VA)	24150	23.7	92.9
AUSTIN (TX)	29496	27.6	94.8
ATLANTIC CITY (NJ)	23639	21.9	61.1
ATLANTA (GA)	133722	24.4	84.5
AZTEC (NM)	3051	21.7	76.7
BATH (NY)	7696	19.2	42.5
BUFFALO (NY)	102555	19.7	32.4
BAKERSFIELD (CA)	37912	27.1	73.7
BOULDER (CO)	10504	21.4	40.9
BALTIMORE (MD)	151409	23.3	79.6
BANGOR (ME)	12045	17.9	33.0
BOISE CITY (ID)	10125	20.9	50.2
PATERSON (NJ)	112797	22.4	81.0
BURLINGTON (VT)	7828	19.1	36.4
BIRMINGHAM (AL)	80149	25.1	84.2
BARNSTABLE (MA)	22275	20.1	50.7
BROWNSVILLE (TX)	15246	28.2	78.1
BOSTON (MA)	230062	20.8	58.0
BATON ROUGE (LA)	27480	26.4	92.9
CEDAR RAPIDS (IA)	12886	20.8	81.0
CHICAGO (IL)	543251	22.3	76.3
CHARLOTTE (NC)	34665	24.2	83.5
CHARLESTON (SC)	21786	26.0	86.5
CHATTANOOGA (TN)	27278	24.3	90.1
CHARLESTON (WV)	23102	21.8	77.9
COLUMBUS (OH)	73424	21.7	75.4
COLORADO SPRINGS (CO)	21173	19.0	39.0
CLEVELAND (OH)	192411	21.8	58.5
CINCINNATI (OH)	83233	22.5	78.9
CANTON (OH)	35823	20.2	58.8

COLUMBIA (SC)	32946	25.4	91.0
CARLISLE (PA)	16529	22.3	62.9
CORPUS CHRISTI (TX)	20657	27.9	84.9
LAYTON (UT)	7228	21.4	50.1
DALLAS (TX)	116462	28.3	95.4
DENVER (CO)	81168	20.4	43.0
BEAVER DAM (WI)	5773	19.7	60.5
DOVER (DE)	8362	22.8	75.9
DURHAM (NC)	14200	23.8	83.3
DES MOINES (IA)	25279	22.1	85.6
DETROIT (MI)	348759	21.6	63.1
DAVENPORT (IA)	25669	21.5	84.1
DAYTONA BEACH (FL)	44885	26.4	91.3
DAYTON (OH)	50614	21.8	78.1
EL CENTRO (CA)	6978	32.2	54.3
ELKHART (IN)	11791	22.3	73.7
EL PASO (TX)	30456	26.7	72.6
ELIZABETH (NJ)	46629	23.1	77.6
ERIE (PA)	25514	20.0	37.5
ESSEX (MA)	62360	20.5	58.2
EUGENE (OR)	22396	17.7	27.7
EVANSVILLE (IN)	17643	23.7	85.9
EVERETT (WA)	28599	17.0	6.0
FARGO (ND)	6372	19.1	48.7
FLINT (MI)	34774	19.7	51.1
FRESNO (CA)	44191	26.1	83.0
FORT LAUDERDALE (FL)	133746	28.2	93.6
FORT MYERS (FL)	34326	27.3	94.1
FORT PIERCE (FL)	26163	27.1	85.6
FORT WORTH (TX)	74381	27.9	96.0
FORT WAYNE (IN)	23452	21.0	76.6
FAYETTEVILLE (NC)	14727	25.2	85.7
GARY (IN)	42247	22.4	76.0
GREEN BAY (WI)	13173	18.9	60.8
GREENSBURG (PA)	39408	22.3	54.9
GRAND HAVEN (MI)	11578	19.4	56.3
GRAND JUNCTION (CO)	6151	23.1	49.6
GRAND RAPIDS (MI)	36477	19.9	55.3
GREENSBORO (NC)	29000	23.3	84.8
GREENVILLE (SC)	24980	24.8	82.5
GAINESVILLE (FL)	11380	25.9	87.1
GETTYSBURG (PA)	5058	22.7	56.9
HICKORY (NC)	9600	23.2	75.4
HOLLAND (MI)	5255	19.4	56.3

HONOLULU (HI)	36742	26.6	30.9
HARRISBURG (PA)	23678	22.3	61.7
HARTFORD (CT)	71541	21.4	60.4
HOUSTON (TX)	161273	27.3	94.4
INDIANAPOLIS (IN)	70216	22.2	81.8
IOWA CITY (IA)	3434	21.0	85.0
JACKSONVILLE (FL)	55432	26.9	89.1
JERSEY CITY (NJ)	52656	19.5	66.6
KLAMATH FALLS (OR)	4132	17.1	27.3
KALAMAZOO (MI)	15947	21.3	61.6
KENOSHA (WI)	10620	20.2	64.2
KANSAS CITY (KS)	100016	24.8	87.2
KNOXVILLE (TN)	36093	23.6	87.8
LAFAYETTE (IN)	8821	21.9	79.8
LAFAYETTE (LA)	10644	26.6	89.5
LAKE CHARLES (LA)	14017	27.4	88.6
LAKELAND (FL)	39449	27.8	79.7
LANCASTER (PA)	35226	22.5	78.9
LANSING (MI)	17190	19.6	53.4
LOGAN (UT)	2466	20.0	39.6
LOUISVILLE (KY)	65088	23.8	83.0
LA PORTE (IN)	9585	20.9	74.0
LOS ANGELES (CA)	585151	21.5	49.3
LAS VEGAS (NV)	60738	30.8	94.5
LITTLE ROCK (AR)	29271	25.9	92.4
MACON (GA)	15179	25.6	83.3
MCALLEN (TX)	20083	28.9	77.7
MIDDLESEX (NJ)	48927	22.8	86.1
MIDDLETOWN (OH)	21954	22.3	78.7
MEDFORD (OR)	13963	20.6	37.7
MADISON (IL)	21823	24.7	80.0
MODESTO (CA)	26730	25.8	49.9
MADISON (WI)	21529	19.8	62.4
MIAMI (FL)	173549	27.9	87.7
MELBOURNE (FL)	34939	27.3	89.3
MILWAUKEE (WI)	109839	20.2	64.7
MEMPHIS (TN)	70069	26.3	94.1
TOMS RIVER (NJ)	103364	23.0	78.4
MINNEAPOLIS (MN)	113123	20.6	75.9
MONTGOMERY (AL)	18371	27.3	87.2
MOBILE (AL)	32427	27.0	91.4
MONROE (LA)	11976	26.3	86.9
MERCER (PA)	12730	19.8	52.6
UPPER MARLBORO (MD)	33827	23.0	90.6

MUSKEGON (MI)	14426	19.4	55.2
MUNCIE (IN)	10741	22.4	74.6
MYRTLE BEACH (SC)	12073	25.6	83.4
NAMPA (ID)	4082	20.6	47.1
NASHUA (NH)	22925	21.5	47.9
MELVILLE (NY)	217220	21.4	73.4
NILES (MI)	14168	20.8	62.2
NORFOLK (VA)	69980	24.5	87.9
NASHVILLE (TN)	44063	24.5	94.8
NEWBURGH (NY)	23313	20.6	58.5
NEW HAVEN (CT)	72842	21.7	59.5
NEW LONDON (CT)	18931	20.6	54.1
NEW ORLEANS (LA)	88199	27.6	89.0
NEWARK (NJ)	107048	23.1	71.1
NEW YORK (NY)	691188	19.5	61.4
OCALA (FL)	21980	26.0	81.9
OKLAHOMA CITY (OK)	52741	25.7	93.7
OAKLAND (CA)	145642	16.9	31.0
OMAHA (NE)	33423	22.4	92.7
ORLANDO (FL)	65320	26.8	91.0
OTTAWA (IL)	11733	21.3	74.2
PHILADELPHIA (PA)	427954	22.6	78.0
PHOENIX (AZ)	152406	33.2	88.4
PALM BEACH (FL)	94124	27.3	89.5
PLYMOUTH (MA)	34916	20.2	56.5
PENSACOLA (FL)	21640	26.8	90.4
PORTLAND (OR)	94919	18.7	28.9
PROVO (UT)	11373	21.8	46.1
PORT ARTHUR (TX)	23927	26.9	90.1
PORTAGE (IN)	8380	22.4	78.8
PORTLAND (ME)	21078	18.3	36.1
PROVIDENCE (RI)	118928	20.7	50.4
PITTSBURGH (PA)	154655	21.1	57.4
RICHMOND (VA)	40673	23.8	86.4
ROCHESTER (NY)	60756	19.6	48.3
ROCKVILLE (MD)	38423	24.5	92.0
READING (PA)	32927	22.2	71.2
RENO (NV)	18059	20.6	71.0
RALEIGH (NC)	24517	23.9	89.2
RIVERSIDE (CA)	177334	23.3	79.6
SACRAMENTO (CA)	72377	22.1	89.2
SCRANTON (PA)	71109	19.9	47.2
SAN DIEGO (CA)	158466	21.3	34.0
SAN FRANCISCO (CA)	118777	16.9	7.2

SALT LAKE CITY (UT)	39245	22.8	50.9
SAN JOSE (CA)	79032	21.8	32.2
SANTA BARBARA (CA)	25352	18.2	38.1
SAN ANTONIO (TX)	81165	28.1	86.3
SPOKANE (WA)	31111	18.8	45.4
SPRINGFIELD (MA)	44171	21.1	61.9
SPRINGFIELD (MO)	18837	23.3	81.1
SPARTANBURG (SC)	19782	24.2	77.3
SARASOTA (FL)	62363	27.7	93.3
STEUBENVILLE (OH)	11219	21.5	58.6
ST. CHARLES (MO)	11185	24.4	88.9
STOCKTON (CA)	35179	23.5	84.1
EAST ST. LOUIS (IL)	23205	24.5	85.0
SOUTH BEND (IN)	22463	20.8	70.5
ST. LOUIS (MO)	131259	24.7	88.4
STAMFORD (CT)	66789	20.6	72.8
ST. PETERSBURG (FL)	68483	28.5	90.8
STATE COLLEGE (PA)	7171	19.7	55.5
SEATTLE (WA)	102243	16.0	7.3
SIOUX CITY (IA)	7325	21.5	83.5
TACOMA (WA)	41570	17.1	8.2
TAMPA (FL)	68483	27.3	89.9
TUCSON (AZ)	52297	29.1	60.8
TALLAHASSEE (FL)	10497	26.1	88.9
TOLEDO (OH)	44939	21.3	66.0
TOPEKA (KS)	14340	23.7	89.3
TRENTON (NJ)	27141	22.5	86.2
TERRE HAUTE (IN)	11482	22.2	80.6
TULSA (OK)	42061	26.1	92.6
VISALIA (CA)	22014	25.4	44.5
VANCOUVER (WA)	15616	18.6	8.6
VENTURA (CA)	37298	18.8	43.4
WICHITA (KS)	30425	24.9	92.4
OGDEN (UT)	10926	23.5	47.6
WILMINGTON (DE)	34130	22.7	81.6
WINSTON-SALEM (NC)	22695	24.2	80.4
WORCESTER (MA)	62512	18.8	46.6
WASHINGTON (DC)	67541	24.3	87.3
WASHINGTON (PA)	22831	20.7	54.5
YOUNGSTOWN (OH)	41226	19.6	56.9
YORK (PA)	28005	22.0	71.6

eTable 2. Country specific sub-periods, and period specific averagre daily mean temperature (Celsisus degree).

Country	Sub-period	Average daily mean temperature
Canada	[1991; 1995]	16.7
Canada	[1996; 2000]	17.1
Canada	[2001; 2005]	17.4
Canada	[2006; 2009]	17.0
Japan	[1972; 1976]	23.5
Japan	[1977; 1980]	24.0
Japan	[1981; 1984]	23.6
Japan	[1985; 1988]	23.9
Japan	[1989; 1992]	24.3
Japan	[1993; 1996]	24.1
Japan	[1997; 2000]	24.7
Japan	[2001; 2004]	24.6
Japan	[2005; 2009]	24.7
Spain	[1990; 1994]	22.0
Spain	[1995; 1998]	21.8
Spain	[1999; 2004]	22.5
Spain	[2005; 2009]	22.4
USA	[1973; 1976]	22.4
USA	[1977; 1981]	23.0
USA	[1982; 1986]	22.7
USA	[1987; 1991]	23.2
USA	[1992; 1996]	23.0
USA	[1997; 2001]	22.9
USA	[2002; 2006]	23.0

eTable 3. Multivariate multilevel meta-regression models with different fixed-effects specification and related p-values of Wald tests.

	Model 1	Model 2	Model 3	Model 4
Country*year interaction		<0.0001	<0.0001	<0.0001
Air Conditioning (%)			<0.0001	0.011
Average summer mean temperature ^{o}C				0.740
Interquartile range of mean temperature ^{o}C				<0.0001
I^2	35.0%	22.5%	22.1%	20.5%

Model 1: Intercept
Model 2: Intercept, country*year interaction
Model 3: Intercept, country*year interaction, AC
Model 4: Intercept, country*year interaction, AC, average mean temperature, interquartile range of mean temperature

eTable 4. Attributable fractions (AF%), Attributable deaths by country and sub periods calculated under observed air conditioning prevalence (Scenario 1) and under Scenario 2 on which, in each country, air conditioning prevalence is set at the level observed at the beginning of the observational period. Delayed deaths were calculated as difference between attributable deaths calculated between scenario 2 and scenario 1.

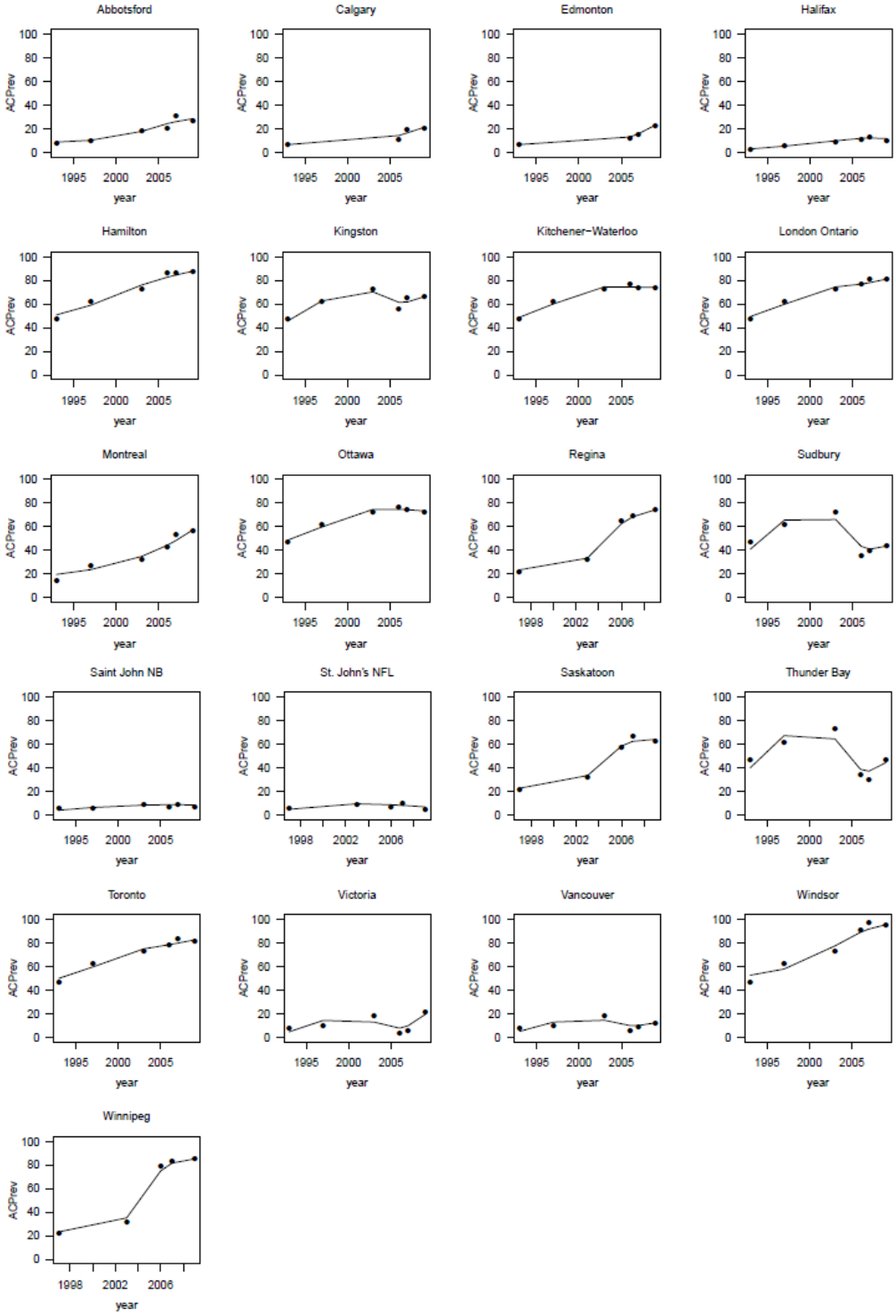
		Scenario 1. Observed air conditioning prevalence				**Scenario 2: Air conditioning prevalence set at the level observed at the beginning of the observational period**				
		AF%		**Attributable deaths**		**AF%**		**Attributable deaths**		**Delayed deaths**
Country	Period	**Point estimate**	**95%CI**	**Point estimate**	**95%CI**	**Point estimate**	**95%CI**	**Point estimate**	**95%CI**	
Canada	[1991; 1995]	1.4	(1.2; 1.6)	2366.4	(2070.9; 2642)	1.4	(1.2; 1.6)	2381.2	(2108.8; 2647)	14.8
	[1996; 2000]	1.3	(1.2; 1.5)	2284.4	(2047.4; 2506)	1.4	(1.2; 1.5)	2345.4	(2116.8; 2571.2)	61.0
	[2001; 2005]	1.2	(1.1; 1.4)	1928.1	(1663.4; 2191.7)	1.3	(1.2; 1.5)	2095.6	(1835.9; 2336.4)	167.5
	[2006; 2009]	0.8	(0.6; 1)	1002.3	(758.8; 1230.7)	0.9	(0.7; 1.1)	1136.8	(903.6; 1357.2)	134.5
									Delayed deaths	377.8
									Total deaths	793073
									Delayed AF%	0.05

Japan	[1972; 1976]	3.6	(3.5; 3.6)	37131.7	(36735.3; 37554.5)	3.6	(3.5; 3.6)	37293.1	(36862.9; 37678.3)	161.4
	[1977; 1980]	3.1	(3.1; 3.2)	26476.9	(26182.5; 26761.1)	3.3	(3.2; 3.3)	27486.9	(27206.4; 27779.2)	1010.0
	[1981; 1984]	2.8	(2.8; 2.9)	24687.3	(24383.1; 24972.6)	3.0	(3; 3.1)	26287.2	(25940.6; 26652.8)	1599.9
	[1985; 1988]	2.5	(2.5; 2.6)	23182.0	(22867.4; 23527.1)	2.8	(2.7; 2.8)	25453.3	(25014.1; 25890.6)	2271.3
	[1989; 1992]	2.2	(2.2; 2.3)	22402.2	(21979.9; 22798.3)	2.6	(2.5; 2.6)	25862.9	(25263.2; 26440.4)	3460.7
	[1993; 1996]	1.9	(1.9; 1.9)	20218.4	(19799.2; 20665.7)	2.3	(2.2; 2.4)	24519.9	(23831; 25203.5)	4301.5
	[1997; 2000]	1.7	(1.7; 1.7)	18937.1	(18452.2; 19409.5)	2.2	(2.1; 2.2)	24118.4	(23288.9; 24930.7)	5181.3
	[2001; 2004]	1.4	(1.4; 1.5)	17090.7	(16617; 17538)	1.9	(1.8; 2)	22989.7	(22051.7; 23992.8)	5899.0
	[2005; 2009]	1.1	(1; 1.1)	18268.8	(17422.3; 19040.2)	1.6	(1.5; 1.6)	26097.9	(24791.5; 27415.1)	7829.1
									Delayed deaths	31714.2
									Total deaths	9764534
									Delayed AF%	0.32
Spain	[1990; 1994]	3.5	(3.4; 3.7)	6055.3	(5791.6; 6306.5)	3.5	(3.4; 3.7)	6061.7	(5805.7; 6314.6)	6.4
	[1995; 1998]	3.5	(3.4; 3.7)	5005.7	(4848; 5179.8)	3.6	(3.5; 3.7)	5050.4	(4888.5; 5214.7)	44.7
	[1999; 2004]	3.5	(3.4; 3.6)	7775.2	(7545.3; 7997.1)	3.6	(3.5; 3.7)	7929.4	(7713.4; 8149.7)	154.2
	[2005; 2009]	2.8	(2.6; 2.9)	5201.3	(4919.3; 5455.6)	2.9	(2.8; 3)	5438.9	(5178.8; 5707.5)	237.6
									Delayed deaths	442.9
									Total deaths	918076

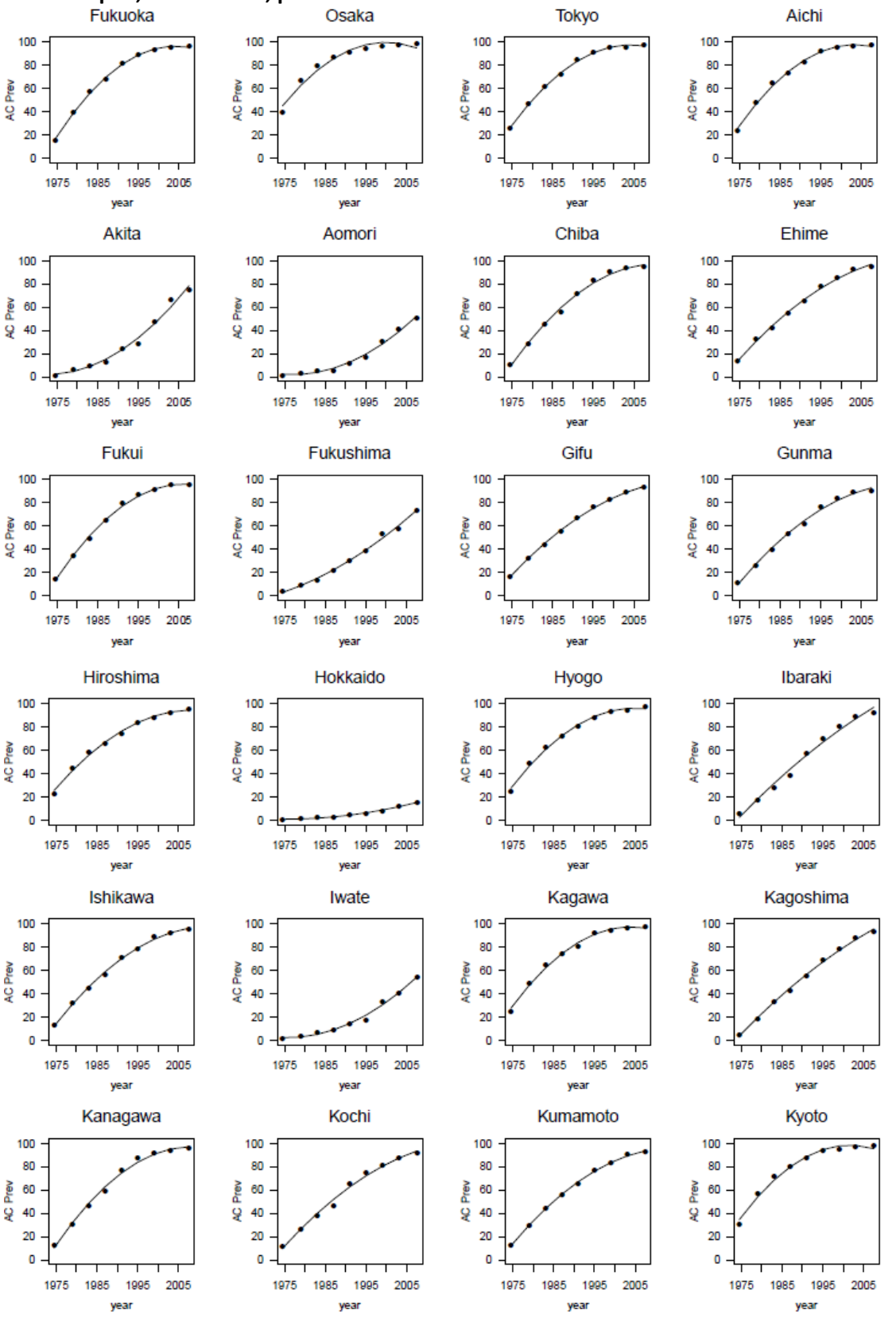
									Delayed AF%	0.05
USA	[1973; 1976]	1.7	(1.7; 1.7)	20659.3	(20327.1; 20967.7)	1.7	(1.7; 1.7)	20540.2	(20216.1; 20847.5)	-119.1
	[1977; 1981]	1.6	(1.5; 1.6)	23776.4	(23459.7; 24106.6)	1.6	(1.6; 1.6)	24229.1	(23923.3; 24518.9)	452.7
	[1982; 1986]	1.3	(1.3; 1.3)	21885.6	(21570.6; 22164.5)	1.4	(1.4; 1.4)	22920.2	(22655.3; 23188)	1034.6
	[1987; 1991]	1.1	(1.1; 1.1)	19344.4	(19079.4; 19619.4)	1.2	(1.2; 1.2)	21177.6	(20864.3; 21486.1)	1833.2
	[1992; 1996]	0.9	(0.9; 0.9)	16215.0	(15896.4; 16528.2)	1.0	(1; 1)	18368.7	(18049.8; 18680.7)	2153.7
	[1997; 2001]	0.7	(0.7; 0.7)	12353.9	(12062; 12604.9)	0.8	(0.8; 0.8)	15016.5	(14666.6; 15358.6)	2662.6
	[2002; 2006]	0.5	(0.5; 0.5)	10037.1	(9680.4; 10355.1)	0.7	(0.7; 0.7)	13255.1	(12815.9; 13693.2)	3218.0
									Delayed deaths	11235.7
									Total deaths	11839659
									Delayed AF%	0.09

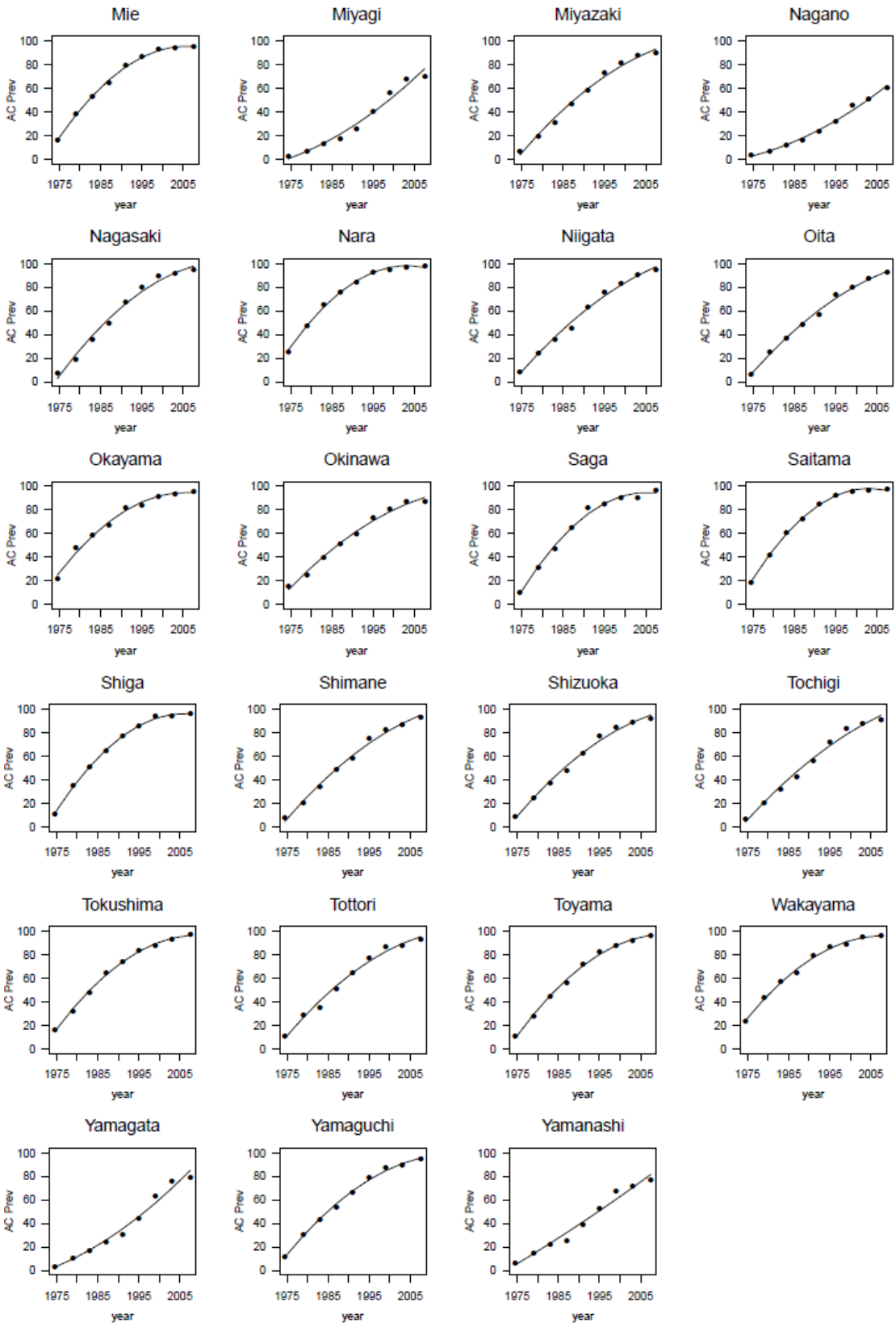
Additional figures

eFigure 1(a). Location specific air conditioning prevalence with the estimated smoothed trends. Canada, 21 locations, period 1986-2009.

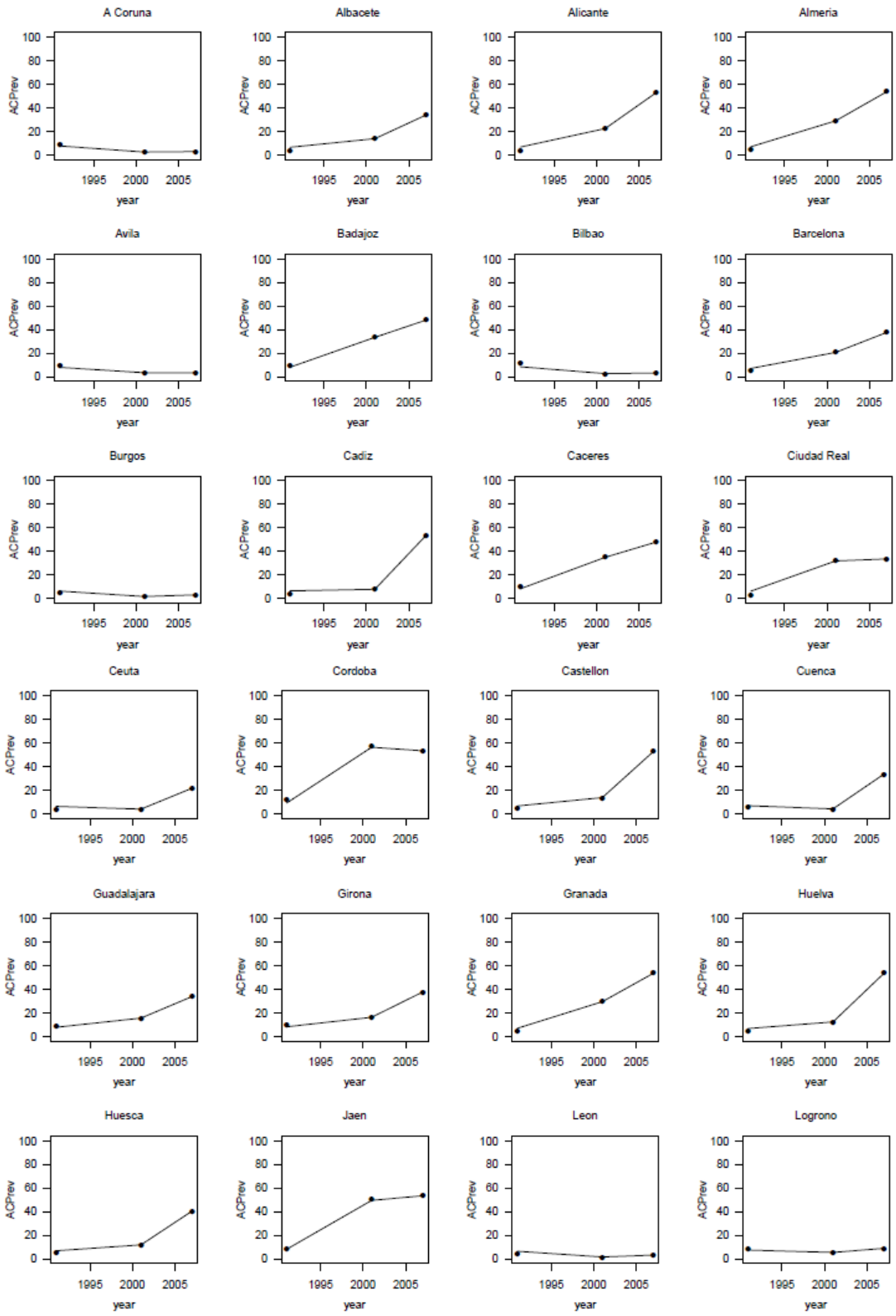


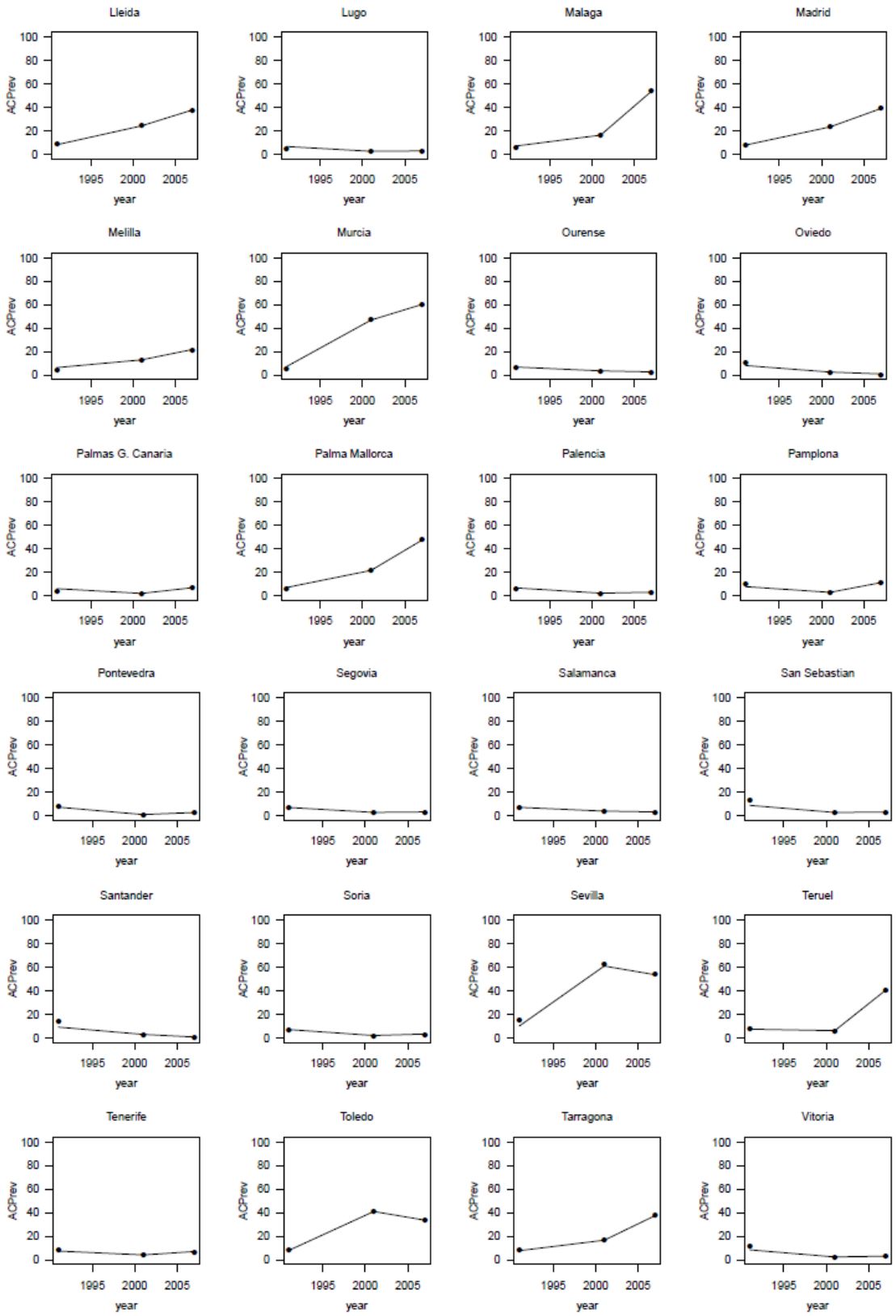
eFigure 1(b). Location specific air conditioning prevalence with the estimated smoothed trends. Japan, 47 locations, period 1972-2009.

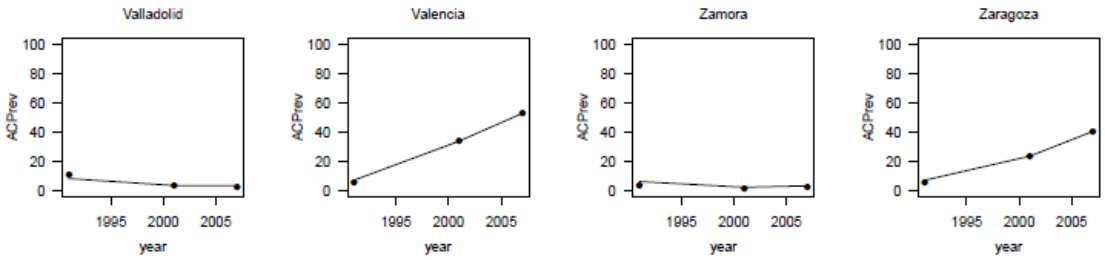



Mie
Miyagi
Miyazaki
Nagano
Nagasaki
Nara
Niigata
Oita
Okayama
Okinawa
Saga
Saitama
Shiga
Shimane
Shizuoka
Tochigi
Tokushima
Tottori
Toyama
Wakayama
Yamagata
Yamaguchi
Yamanashi
AC Prev
year
0
20
40
60
80
100
1975
1985
1995
2005

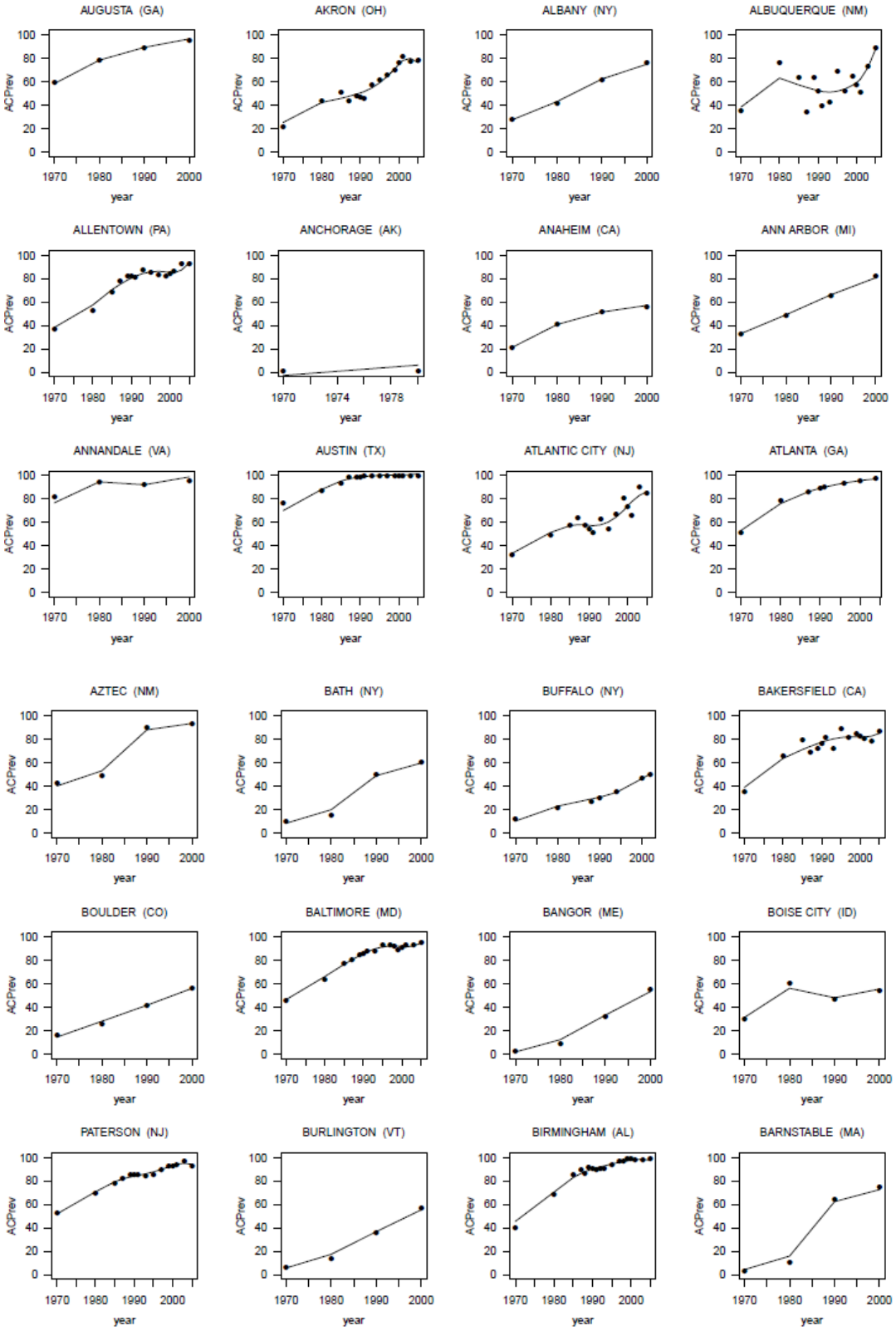
eFigure 1(c). Location specific air conditioning prevalence with the estimated smoothed trends. Spain, 52 locations, period 1990-2014.

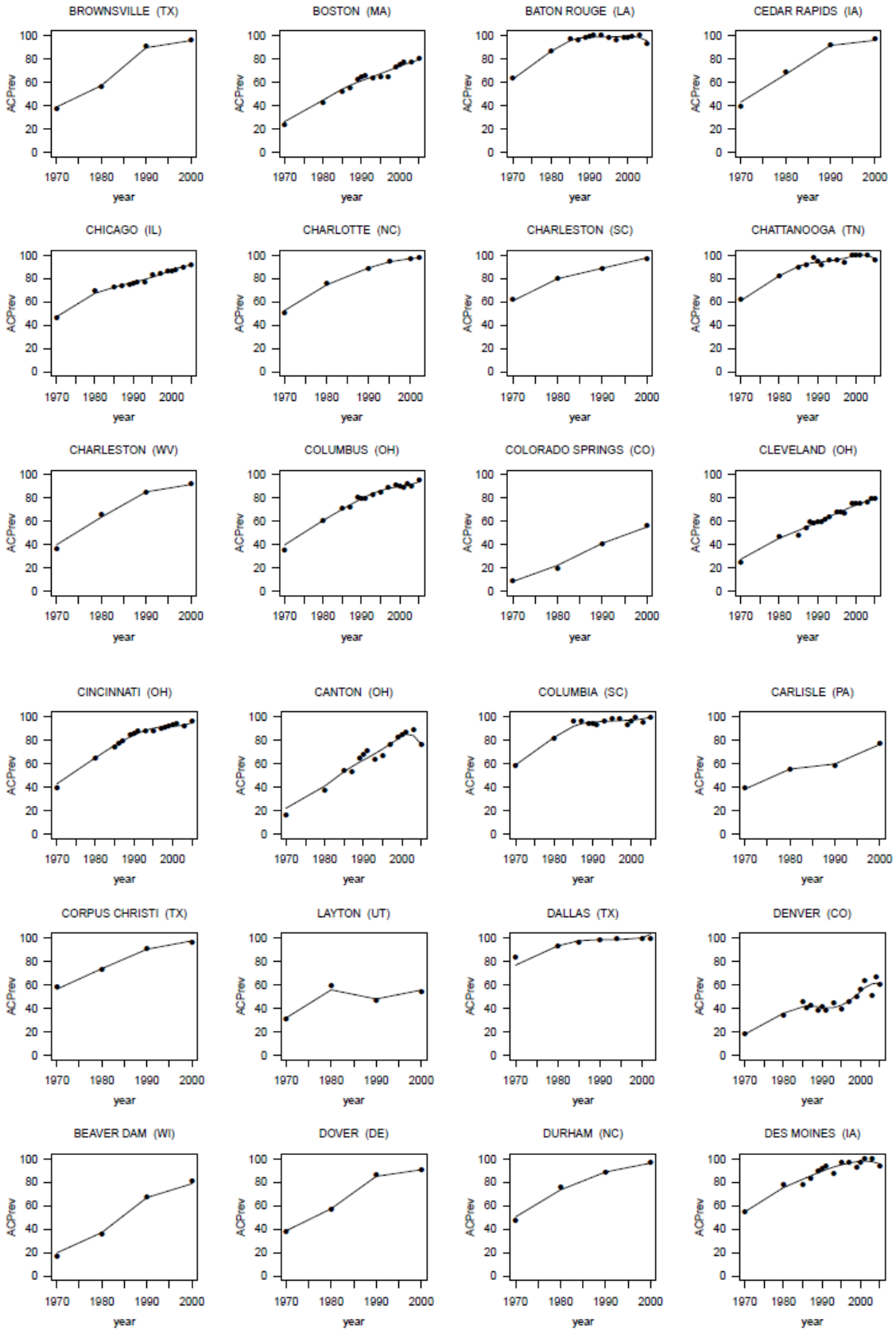


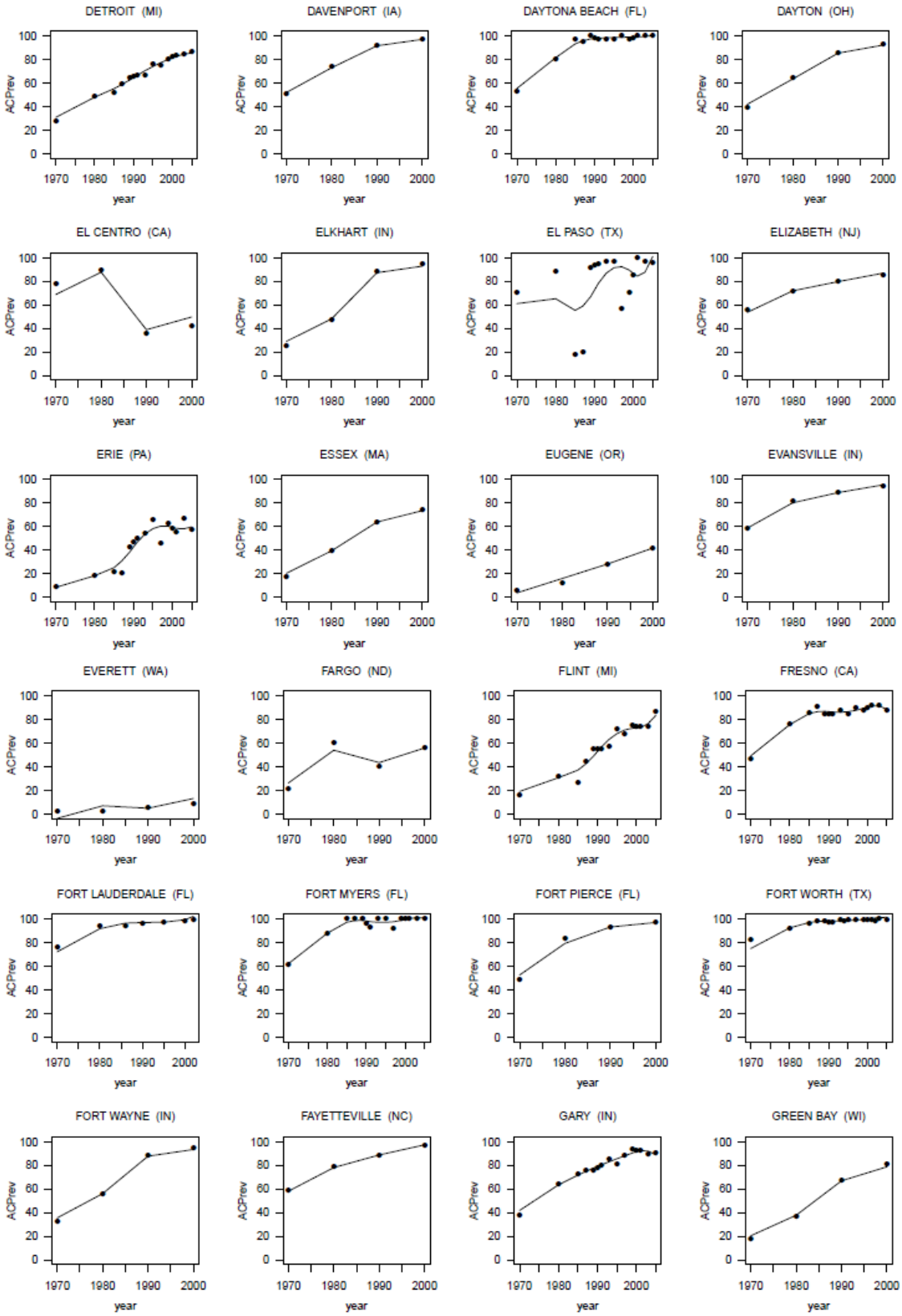

Lleida
Lugo
Malaga
Madrid
Melilla
Murcia
Ourense
Oviedo
Palmas G. Canaria
Palma Mallorca
Palencia
Pamplona
Pontevedra
Segovia
Salamanca
San Sebastian
Santander
Soria
Sevilla
Teruel
Tenerife
Toledo
Tarragona
Vitoria
ACPrev
year
0
20
40
60
80
100
1995
2000
2005

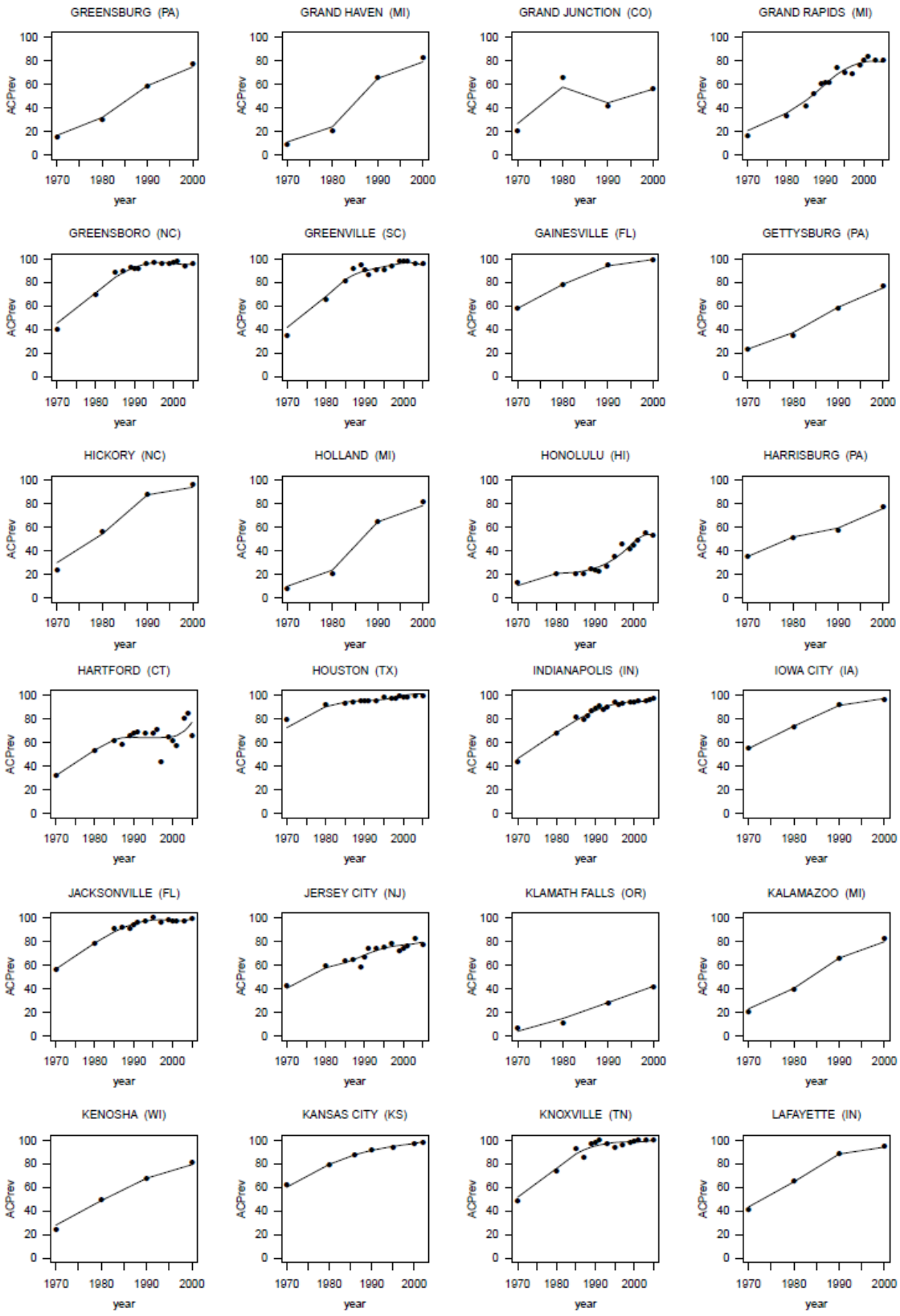

Valladolid
Valencia
Zamora
Zaragoza
ACPrev
year
100
80
60
40
20
0
1995
2000
2005

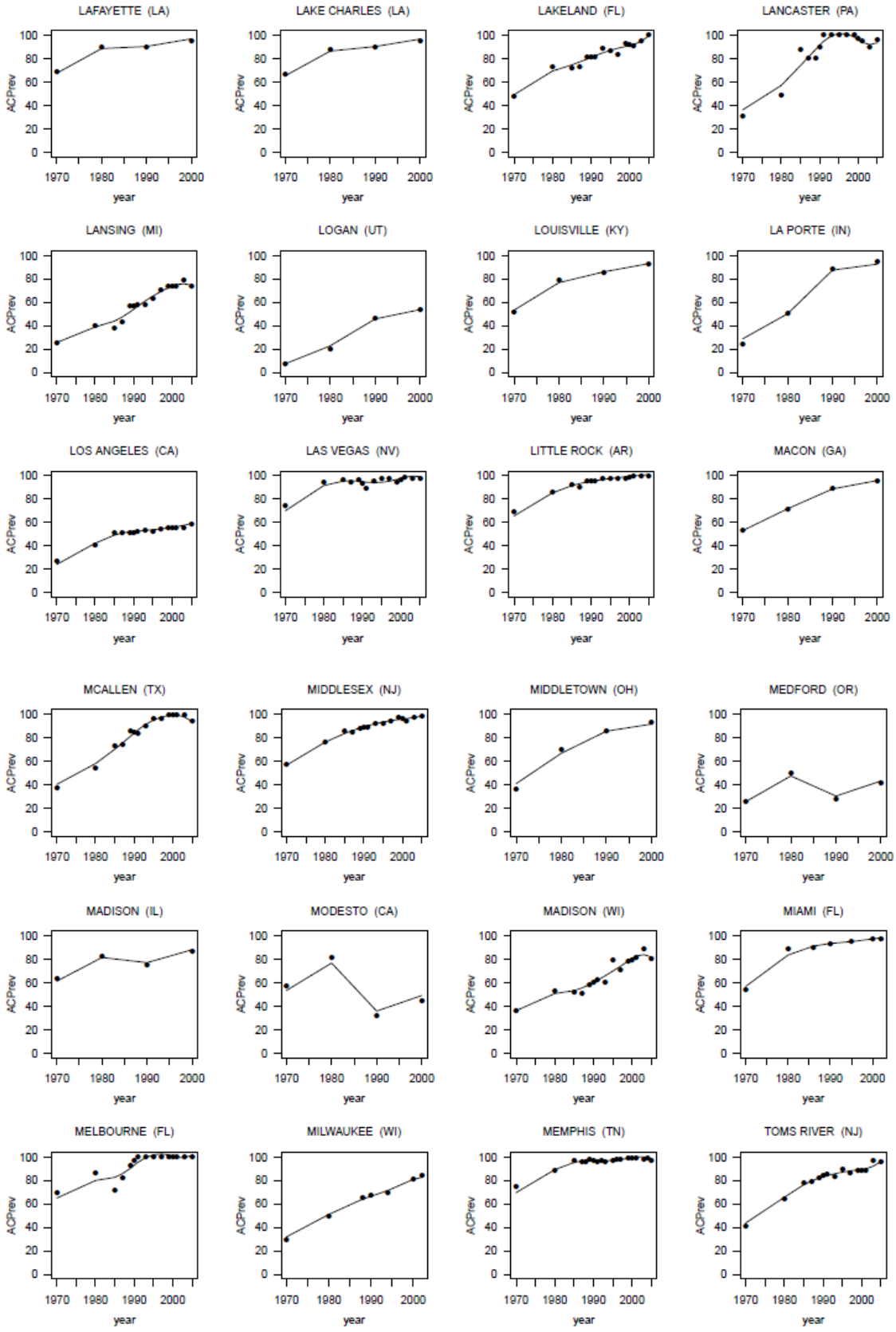
eFigure 1(d). Location specific air conditioning prevalence with the estimated smoothed trends. USA, 211 locations, period 1973-2006.

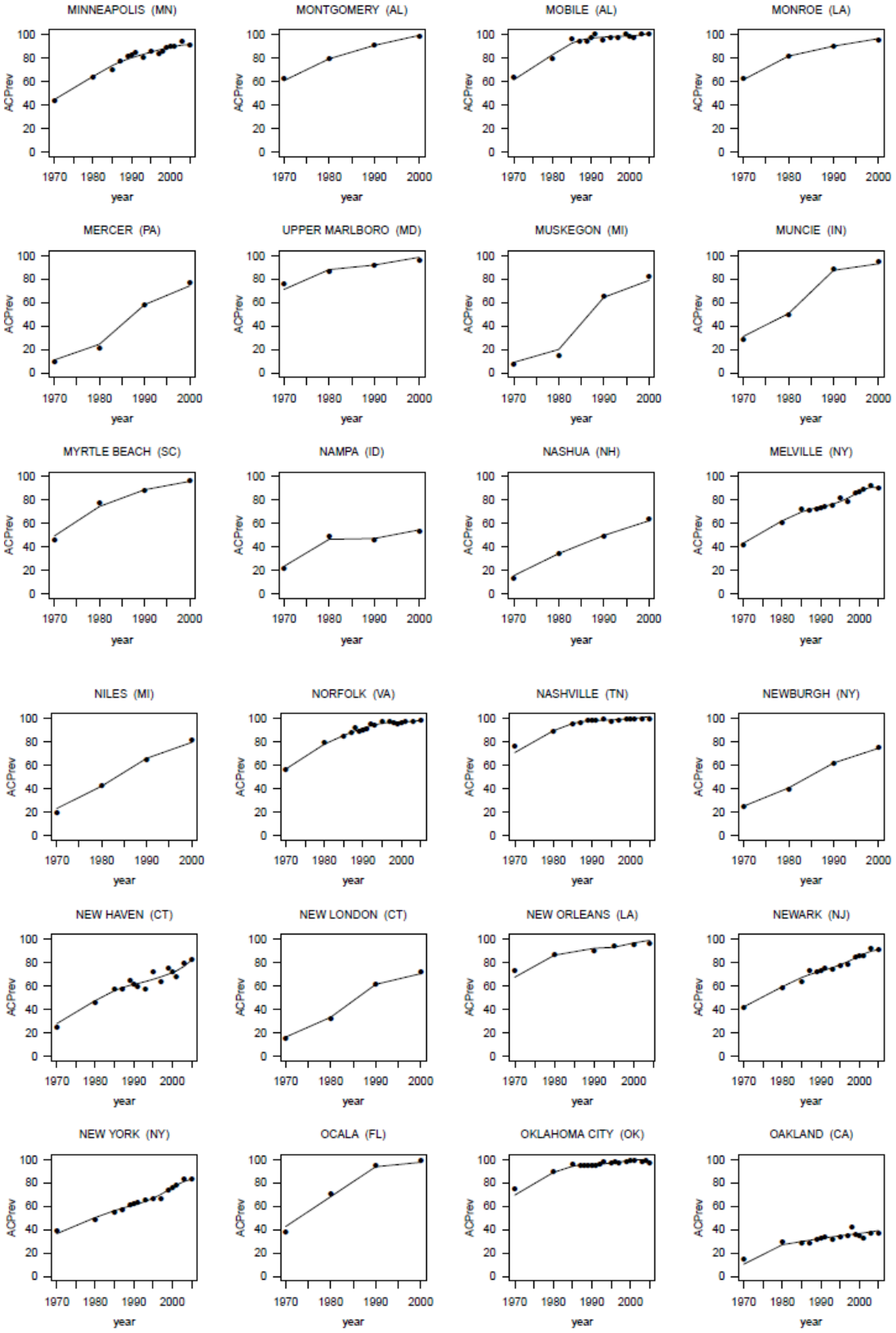


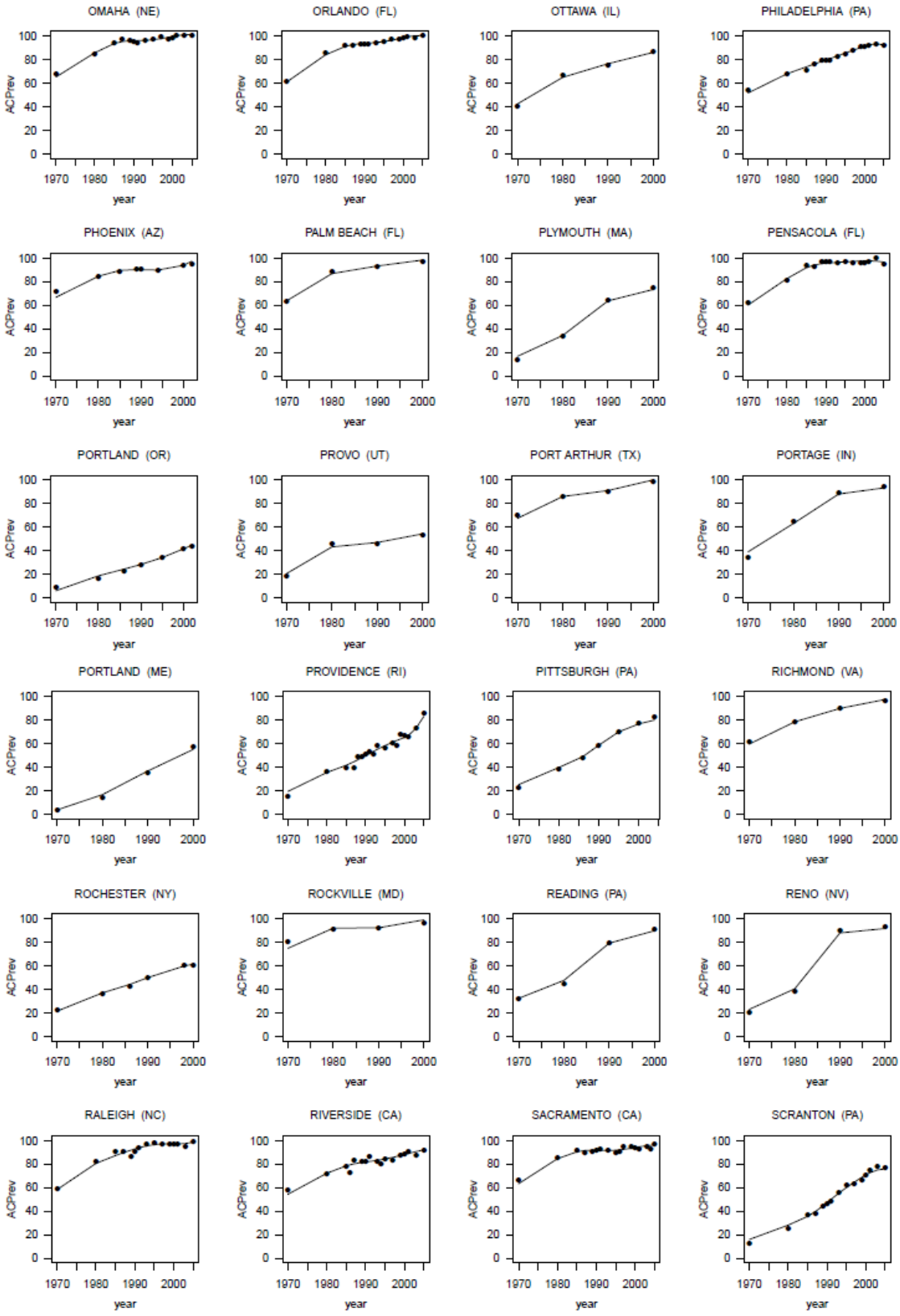

BROWNSVILLE (TX)
BOSTON (MA)
BATON ROUGE (LA)
CEDAR RAPIDS (IA)
CHICAGO (IL)
CHARLOTTE (NC)
CHARLESTON (SC)
CHATTANOOGA (TN)
CHARLESTON (WV)
COLUMBUS (OH)
COLORADO SPRINGS (CO)
CLEVELAND (OH)
CINCINNATI (OH)
CANTON (OH)
COLUMBIA (SC)
CARLISLE (PA)
CORPUS CHRISTI (TX)
LAYTON (UT)
DALLAS (TX)
DENVER (CO)
BEAVER DAM (WI)
DOVER (DE)
DURHAM (NC)
DES MOINES (IA)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000

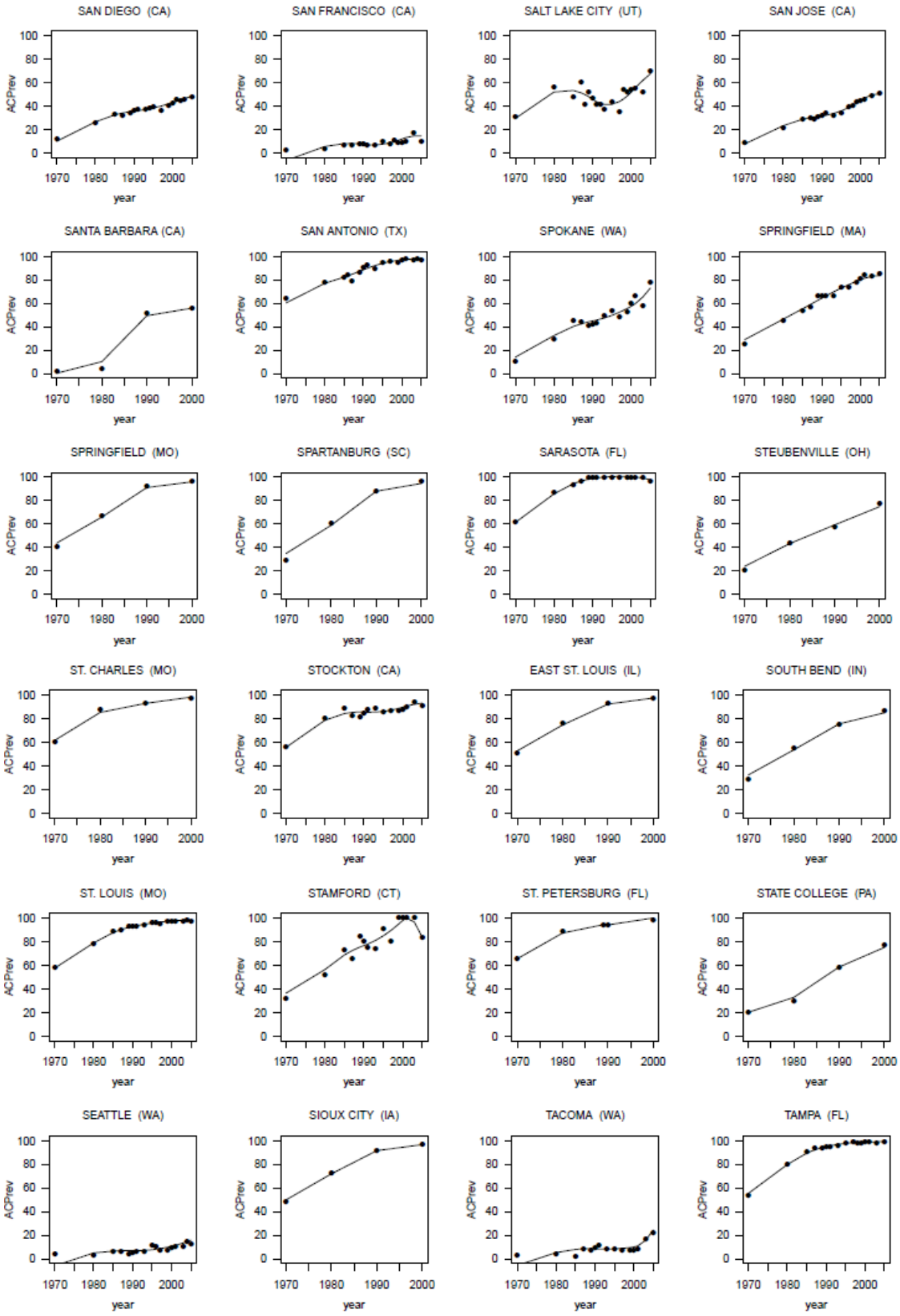

DETROIT (MI)
DAVENPORT (IA)
DAYTONA BEACH (FL)
DAYTON (OH)
EL CENTRO (CA)
ELKHART (IN)
EL PASO (TX)
ELIZABETH (NJ)
ERIE (PA)
ESSEX (MA)
EUGENE (OR)
EVANSVILLE (IN)
EVERETT (WA)
FARGO (ND)
FLINT (MI)
FRESNO (CA)
FORT LAUDERDALE (FL)
FORT MYERS (FL)
FORT PIERCE (FL)
FORT WORTH (TX)
FORT WAYNE (IN)
FAYETTEVILLE (NC)
GARY (IN)
GREEN BAY (WI)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000

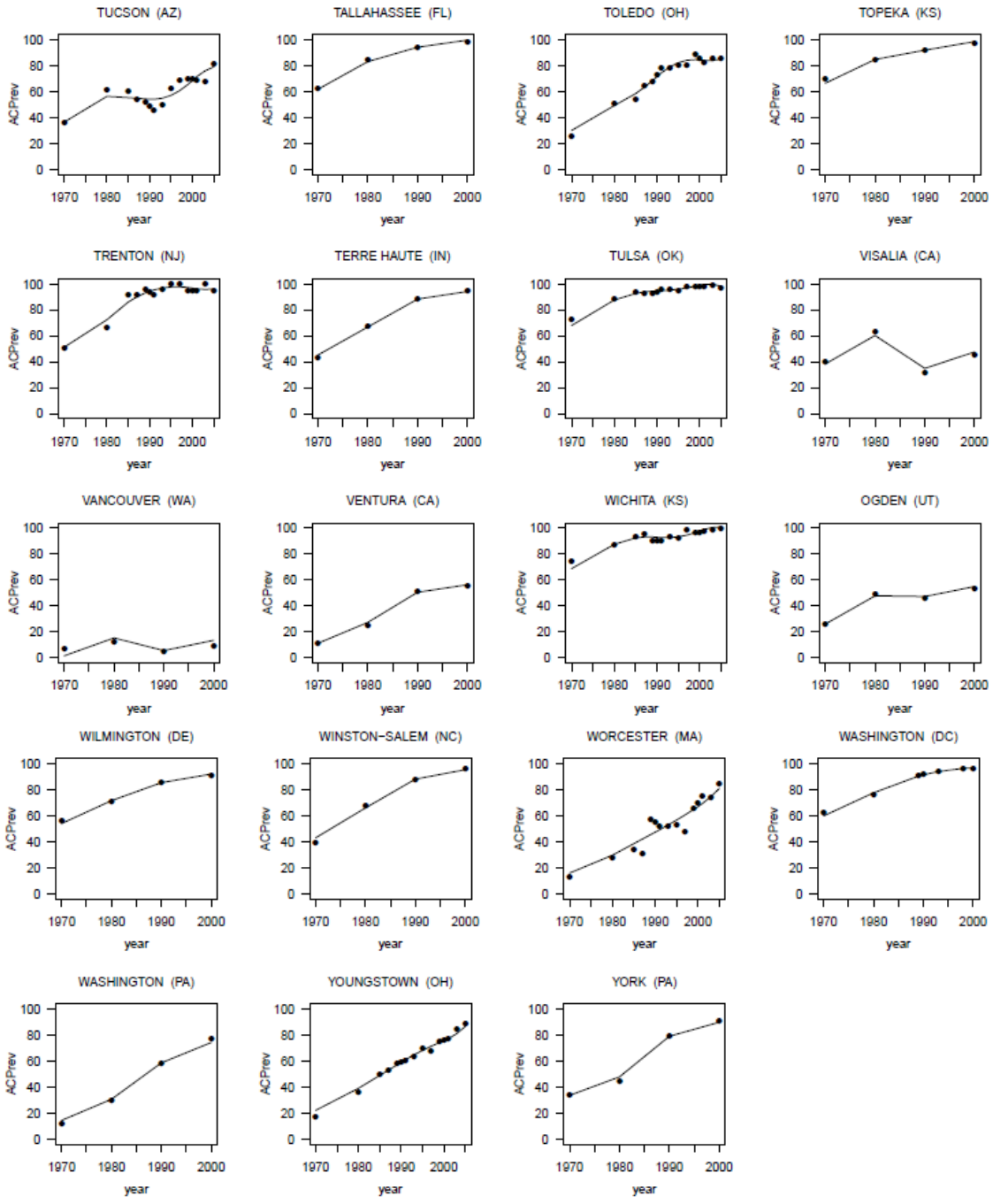

GREENSBURG (PA)
GRAND HAVEN (MI)
GRAND JUNCTION (CO)
GRAND RAPIDS (MI)
GREENSBORO (NC)
GREENVILLE (SC)
GAINESVILLE (FL)
GETTYSBURG (PA)
HICKORY (NC)
HOLLAND (MI)
HONOLULU (HI)
HARRISBURG (PA)
HARTFORD (CT)
HOUSTON (TX)
INDIANAPOLIS (IN)
IOWA CITY (IA)
JACKSONVILLE (FL)
JERSEY CITY (NJ)
KLAMATH FALLS (OR)
KALAMAZOO (MI)
KENOSHA (WI)
KANSAS CITY (KS)
KNOXVILLE (TN)
LAFAYETTE (IN)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000


LAFAYETTE (LA)
LAKE CHARLES (LA)
LAKELAND (FL)
LANCASTER (PA)
LANSING (MI)
LOGAN (UT)
LOUISVILLE (KY)
LA PORTE (IN)
LOS ANGELES (CA)
LAS VEGAS (NV)
LITTLE ROCK (AR)
MACON (GA)
MCALLEN (TX)
MIDDLESEX (NJ)
MIDDLETOWN (OH)
MEDFORD (OR)
MADISON (IL)
MODESTO (CA)
MADISON (WI)
MIAMI (FL)
MELBOURNE (FL)
MILWAUKEE (WI)
MEMPHIS (TN)
TOMS RIVER (NJ)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000

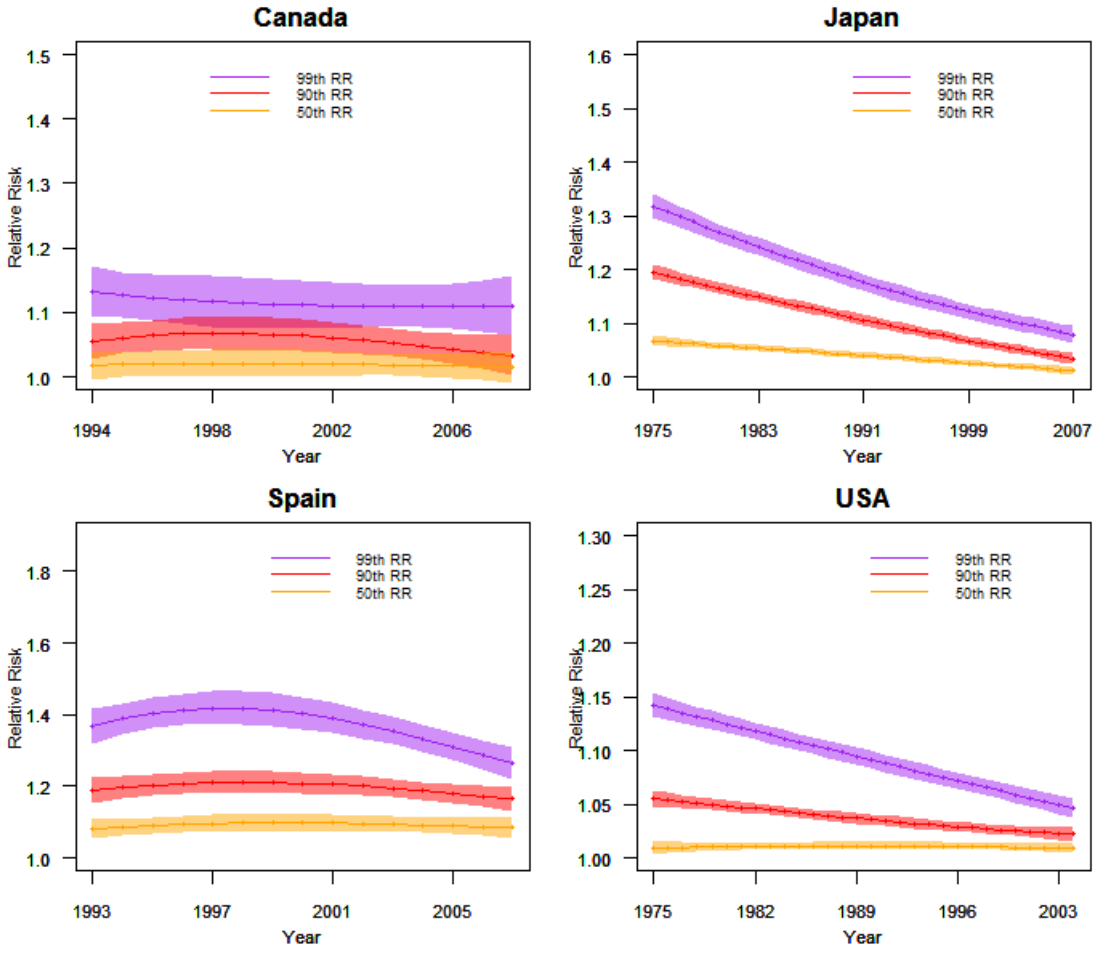

MINNEAPOLIS (MN)
MONTGOMERY (AL)
MOBILE (AL)
MONROE (LA)
MERCER (PA)
UPPER MARLBORO (MD)
MUSKEGON (MI)
MUNCIE (IN)
MYRTLE BEACH (SC)
NAMPA (ID)
NASHUA (NH)
MELVILLE (NY)
NILES (MI)
NORFOLK (VA)
NASHVILLE (TN)
NEWBURGH (NY)
NEW HAVEN (CT)
NEW LONDON (CT)
NEW ORLEANS (LA)
NEWARK (NJ)
NEW YORK (NY)
OCALA (FL)
OKLAHOMA CITY (OK)
OAKLAND (CA)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000


OMAHA (NE)
ORLANDO (FL)
OTTAWA (IL)
PHILADELPHIA (PA)
PHOENIX (AZ)
PALM BEACH (FL)
PLYMOUTH (MA)
PENSACOLA (FL)
PORTLAND (OR)
PROVO (UT)
PORT ARTHUR (TX)
PORTAGE (IN)
PORTLAND (ME)
PROVIDENCE (RI)
PITTSBURGH (PA)
RICHMOND (VA)
ROCHESTER (NY)
ROCKVILLE (MD)
READING (PA)
RENO (NV)
RALEIGH (NC)
RIVERSIDE (CA)
SACRAMENTO (CA)
SCRANTON (PA)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000

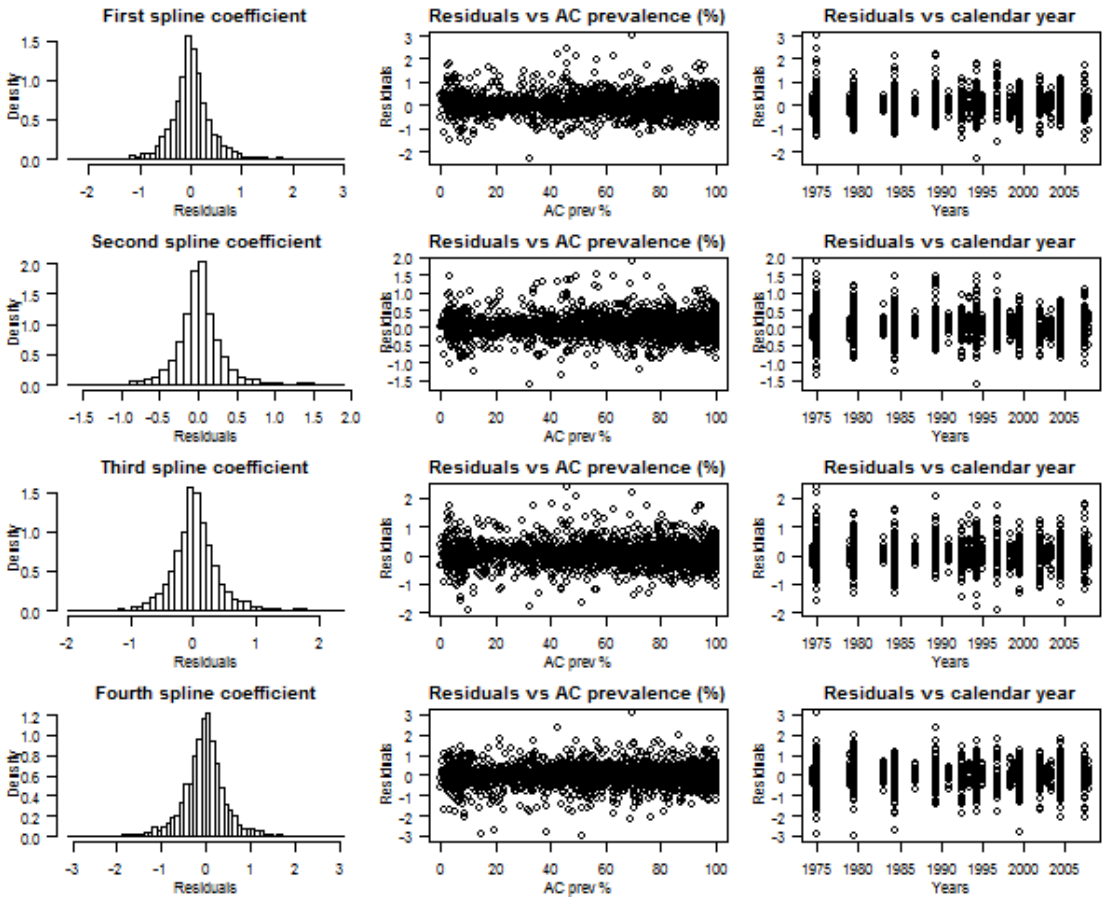

SAN DIEGO (CA)
SAN FRANCISCO (CA)
SALT LAKE CITY (UT)
SAN JOSE (CA)
SANTA BARBARA (CA)
SAN ANTONIO (TX)
SPOKANE (WA)
SPRINGFIELD (MA)
SPRINGFIELD (MO)
SPARTANBURG (SC)
SARASOTA (FL)
STEUBENVILLE (OH)
ST. CHARLES (MO)
STOCKTON (CA)
EAST ST. LOUIS (IL)
SOUTH BEND (IN)
ST. LOUIS (MO)
STAMFORD (CT)
ST. PETERSBURG (FL)
STATE COLLEGE (PA)
SEATTLE (WA)
SIOUX CITY (IA)
TACOMA (WA)
TAMPA (FL)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000


TUCSON (AZ)
TALLAHASSEE (FL)
TOLEDO (OH)
TOPEKA (KS)
TRENTON (NJ)
TERRE HAUTE (IN)
TULSA (OK)
VISALIA (CA)
VANCOUVER (WA)
VENTURA (CA)
WICHITA (KS)
OGDEN (UT)
WILMINGTON (DE)
WINSTON−SALEM (NC)
WORCESTER (MA)
WASHINGTON (DC)
WASHINGTON (PA)
YOUNGSTOWN (OH)
YORK (PA)
ACPrev
year
0
20
40
60
80
100
1970
1980
1990
2000

eFigure 2. Country specific trends of relative risks calculated at 90th, 95th and 99th percentile of the country specific mean temperature distribution in summer months.



eFigure 3. Analysis of the raw residuals of the multivariate multilevel meta-analysis model. For each outcome (spline coefficient) are shown the histogram of the residuals, and the scatterplot of the residuals (y axis) versus AC prevalence (%) and calendar year (x axes).



Chapter 7

Research paper V

Title: A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries.

Author(s): Francesco Sera, Ben Armstrong, Sam Abbott, Sophie Meakin, Kathleen O?Reilly, Rosa von Borries, Rochelle Schneider, Dominic Roye, Masahiro Hashizume, Mathilde Pascal, Aurelio Tobias, Ana Maria Vicedo-Cabrera, MCC Collaborative Research Network, CMMID COVID-19 working group, Antonio Gasparrini, Rachel Lowe.

Journal/Publisher: Nature Communications.

Type of publication: Research paper.

Stage of publication: Published online on October 13, 2021 as doi:10.1038/s41467-021-25914-8.

URL: `http://https://www.nature.com/articles/s41467-021-25914-8`.

Academic peer-reviewed: Yes.

Candidate's role: See Section 2.3.

Senior author: (Prof. Rachel Lowe)

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-25914-8 OPEN

A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries

Francesco Sera[1,2]✉, Ben Armstrong[1], Sam Abbott[3,4], Sophie Meakin[3,4], Kathleen O'Reilly[3,4], Rosa von Borries[5], Rochelle Schneider[1,6,7,8], Dominic Royé[9], Masahiro Hashizume[10,11,12], Mathilde Pascal[13], Aurelio Tobias[11,14], Ana Maria Vicedo-Cabrera[15,16], MCC Collaborative Research Network*, CMMID COVID-19 Working Group*, Antonio Gasparrini[1,6,17] & Rachel Lowe[3,4,6,18]✉

There is conflicting evidence on the influence of weather on COVID-19 transmission. Our aim is to estimate weather-dependent signatures in the early phase of the pandemic, while controlling for socio-economic factors and non-pharmaceutical interventions. We identify a modest non-linear association between mean temperature and the effective reproduction number (R_e) in 409 cities in 26 countries, with a decrease of 0.087 (95% CI: 0.025; 0.148) for a 10 °C increase. Early interventions have a greater effect on R_e with a decrease of 0.285 (95% CI 0.223; 0.347) for a 5th - 95th percentile increase in the government response index. The variation in the effective reproduction number explained by government interventions is 6 times greater than for mean temperature. We find little evidence of meteorological conditions having influenced the early stages of local epidemics and conclude that population behaviour and government interventions are more important drivers of transmission.

[1] Department of Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, London, UK. [2] Department of Statistics, Computer Science and Applications "G. Parenti", University of Florence, Florence, Italy. [3] Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK. [4] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK. [5] Charité Universitätsmedizin, Berlin, Germany. [6] Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, UK. [7] Forecast Department, European Centre for Medium-Range Weather Forecast (ECMWF), Reading, UK. [8] Φ-Lab, European Space Agency, Frascati, Italy. [9] Department of Geography, CIBER of Epidemiology and Public Health (CIBERESP), University of Santiago de Compostela, Santiago de Compostela, Spain. [10] Department of Paediatric Infectious Disease, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan. [11] School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan. [12] Department of Global Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [13] Santé Publique France, Department of Environmental and Occupational Health, French National Public Health Agency, Saint Maurice, France. [14] Institute of Environmental Assessment and Water Research (IDAEA), Spanish Council for Scientific Research (CSIS), Barcelona, Spain. [15] Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland. [16] Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland. [17] Centre for Statistical Modelling, London School of Hygiene & Tropical Medicine, London, UK. [18] Barcelona Supercomputing Center, Barcelona, Spain. *Lists of authors and their affiliations appear at the end of the paper. ✉email: francesco.sera@lshtm.ac.uk; rachel.lowe@lshtm.ac.uk

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has rapidly spread across the globe, traversing diverse climatic and environmental conditions. Sustained local transmission has occurred in most countries, leading to political, social and economic challenges and devastating loss of life. From the early phase of the pandemic, there has been speculation that weather conditions could modulate SARS-CoV-2 transmission patterns. The debate has been driven by analogy with existing seasonal endemic respiratory viral infections, such as influenza and other human coronaviruses, which tend to peak in the drier and colder winter months in temperate climates[1]. However, specific mechanisms behind this seasonality, in terms of host immunity and susceptibility, viral stability or weather-sensitive human behaviour are poorly understood[2]. Dynamic transmission modelling has shown that meteorological variables, such as temperature and humidity, are unlikely to have been a dominant transmission risk factor in the early stages of the COVID-19 pandemic, given high population susceptibility[3,4]. As SARS-CoV-2 is a new virus to humans, with <1 year of data available at the time of writing, ascertaining the potential for weather modulated transmission is challenging. Several studies have attempted such analyses. However, many such studies had methodological weaknesses and the results were at times conflicting[5,6]. Study findings for temperature resulted in either a positive[7,8], negative[9,10], non-linear[11,12] or non-significant association[13,14] with the COVID-19 response variable. For example, most studies did not control for key modulating factors, such as varying government restrictions, socio-economic indicators, population density or age structure[15–17].

In this study, we overcome methodological issues of previous approaches by using a two-stage ecological modelling approach to examine the impact of meteorological variables on SARS-CoV-2 transmission by comparing cities located across the globe, while accounting for confounding of non-pharmaceutical interventions (NPIs) and city-level covariates. The study is based on an extensive dataset, collected by the Multi-Country Multi-City MCC Collaborative Research Network (https://mccstudy.lshtm.ac.uk/), consisting of time series of daily COVID-19 cases registered between 11 January and 28 April 2020 in 409 locations (cities or small regions) in 26 countries. In the first stage, we estimated the effective reproduction number (R_e), in each city, over a city-specific time window early in the epidemic. We use a renewal equation-based approach that estimates latent infections and then map these infections to observed notifications via an incubation period, a report delay and a negative binomial observation model with a day of the week effect[18]. Focusing on the early phase of the pandemic allows us to minimise possible biases coming from factors impacting R_e (in particular non-pharmaceutical interventions (NPIs)), which developed as the pandemic progressed. These include change of ascertainment methods and strategies, the implementation of strong NPIs (e.g. travel bans, school closures and lockdown), the appearance of new variants and ultimately vaccination campaigns. Also, in the first stage we define our exposure variables as mean values of meteorological variables (including daily mean temperature, relative and absolute humidity, solar radiation, wind speed and precipitation), for each city, over the early-phase time window, using the ERA5 fifth-generation European Centre for Medium-Range Weather Forecast atmospheric reanalysis of the global climate[19]. In a second 'cross-sectional' stage, we estimate the association of city-level R_e, calculated for the city-specific window (allowing for standard errors), with each meteorological variable, controlling for confounding by total population, population density, gross domestic product (GDP) per capita, percentage of population >65 years, pollution levels (i.e. particulate matter, $PM_{2.5}$) and the lagged Oxford COVID-19 Government Response Tracker (OxCGRT) Government Response Index at the endpoint of the selected time window (lagged by 10 days), allowing for the two-level (cities and countries) structure of the data using a multilevel meta-regression model[20] (see 'Methods' for further details). We believe the data used and the analysis performed in this study improves upon previous approaches. Specifically, the fine spatial scale of the city-level data and the methodological design, accounting for confounding of NPIs and city-level covariates, allows us to accurately quantify the relationship between meteorological variables and R_e.

Results

Descriptive analysis of meteorological variables and R_e. The bivariate distribution of mean temperature and the effective reproduction number (R_e) across the 409 study cities is shown in Fig. 1, and the characteristics of the 26 countries are reported in Table 1. The mean effective reproduction number (R_e) across all cities was 1.4, ranging from 0.7 to 2.1, with all but ten cities experiencing an epidemic curve with a reproduction number >1. Mean temperatures over the observation period (between January and April 2020) reflect the late winter/early spring in 381 cities situated in the northern hemisphere and the summer/early autumn seasons in 28 cities in the southern hemisphere. Of the 136 cities classified as having high R_e values, 35 cities experienced low temperatures, 64 medium temperatures and 34 high temperatures (Fig. 1). When visualising the unadjusted association of R_e with mean temperature, relative humidity (RH), absolute humidity (AH), solar radiation at the surface and stratified by climate zone, we found no clear pattern (Fig. 2).

Associations between meteorological variables and R_e. Using a two-stage meta-regression model, we quantified the influence of meteorological variables, including mean temperature, on R_e between cities, while controlling for confounding factors including government interventions. After adjusting for the city-level characteristics (e.g. socio-economic and demographic factors) and the country's OxCGRT Government Response Index, we found a modest, non-linear association of mean temperature and AH with R_e (Table 2). Less strong evidence of association was found for RH, with no evidence of association for solar radiation, wind speed and precipitation (Table 2). The association between mean temperature and R_e is non-linear, with R_e initially rising to a peak at 10.2 °C, then falling to a trough at 20 °C, 0.087 (95% confidence interval (CI): 0.025; 0.148) lower than the peak, and finally rising again (Fig. 3). AH has a similar non-linear shape with a maximum difference of 0.061 (95% CI: 0.011; 0.111) between the peak at 6.6 g/m^3 and the trough at 11 g/m^3.

The effect of NPIs on R_e. Although we calculated R_e over a time window in which the OxCGRT Government Response Index, lagged by 10 days, had not yet reached 70, we included the value of the lagged OxCGRT Government Response Index at the end of the city-specific window in the model, to control for residual confounding. Despite being capped at 70, the OxCGRT Government Response Index had a strong association with the reproduction number ($p < 0.0001$) (Supplementary Table 4), explaining 13.8% of its variability (Fig. 3 and Supplementary Table 4, Models D1–D7) with an estimated reduction of R_e equal to 0.285 (95% CI: 0.223; 0.347) when levels of the Government Response Index increase from 21 (5th percentile) to 66 (95th percentile). Mean temperature explained 2.4% and AH 2.0% of the variation in R_e, and the five city-level characteristics explained 1.4% of the variability of the reproduction number (Supplementary Table 4, Models D1–D8).

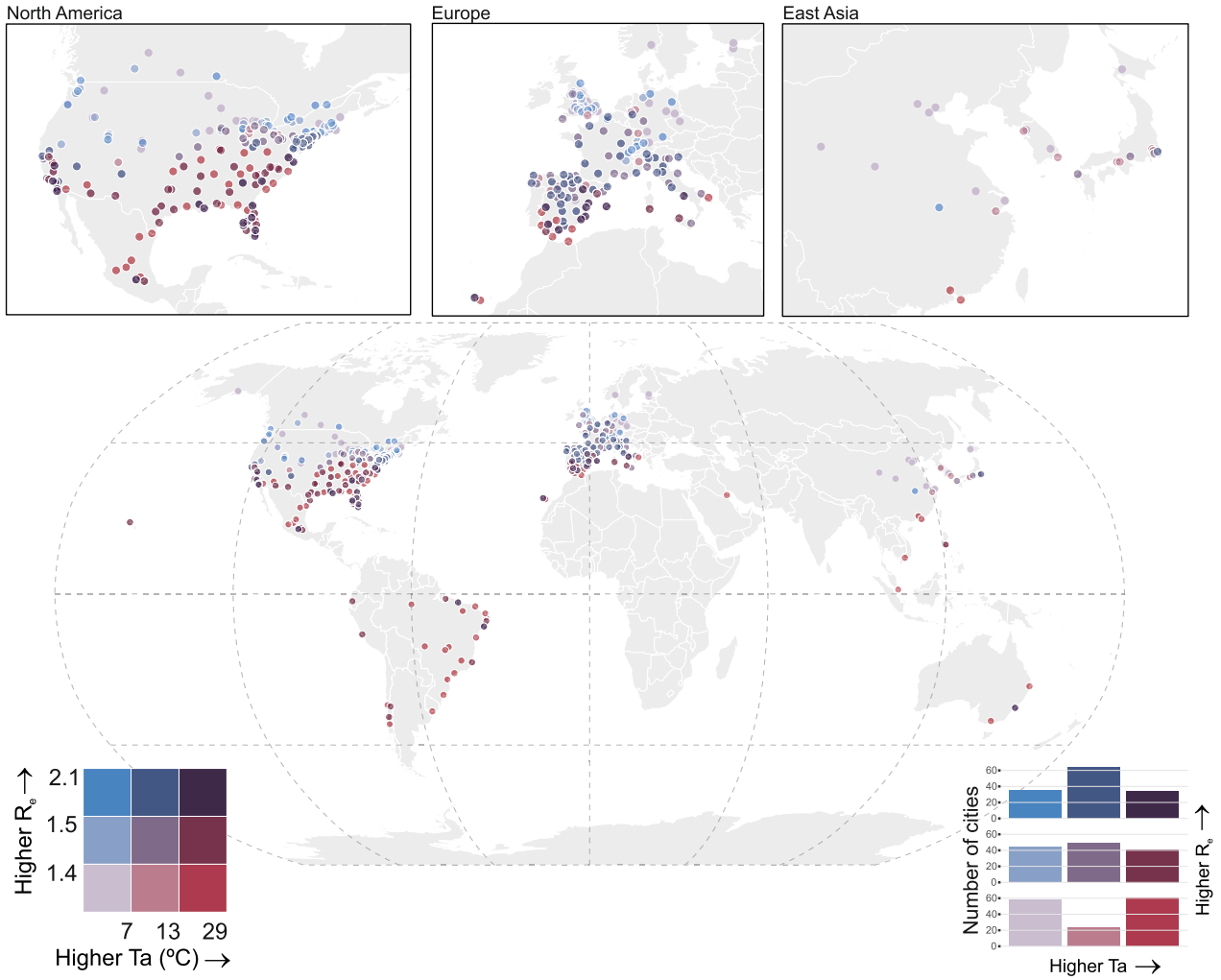


Fig. 1 Effective reproduction number and mean temperature in the observation window for 409 cities. Bivariate plot of effective reproduction number (R_e) and mean temperature (Ta) (°C) in the observation window for each of the 409 study cities. Dark purple circles represent cities with both high R_e and high Ta, while pale purple circles show areas with both low R_e and low Ta. Red circles represent cities with low R_e and high Ta and blue circles depict areas with high R_e and low Ta. The bar chart (bottom right) represents the number of cities in each category defined in the bivariate legend (bottom left).

Sensitivity analyses. We performed several sensitivity analyses to evaluate the robustness of the results considering alternative analytic or selection choices (see Supplementary Table 5). The main results are stable when including a country-level fixed effect in the meta-regression model, i.e. considering the only within-country variation of covariates and outcome. Restricting the analysis to cities with weaker interventions (OxCGRT Government Response Index <60) also gives similar results to the main analysis, apart from wind speed and precipitation also showing an association with R_e. The association between mean temperature and the effective reproduction number holds across all the sensitivity analyses, apart from in tropical and southern hemisphere cities, when stratifying by tropical and non-tropical or northern and southern hemisphere regions. However, this may be explained by the small number of cities and the resulting low power in the tropical and southern hemisphere sub-group. The association between AH and the effective reproduction number is somewhat less robust with no association observed when excluding tropical or southern hemisphere cities, when excluding China and Brazil (countries with earlier and later observation periods) and when considering meteorological variables lagged by 10 days. Excluding the ten cities with $R_e < 1$ shows a tendency of an increased R_e for cities with low RH ($p = 0.009$) and a lower R_e in cities with higher solar radiation at the surface ($p = 0.047$) (Supplementary Figure 5). We observed similar overall tendencies to our main results when we did not control for the OxCGRT Government Response Index in our model, although the effect of temperature and AH was enhanced (Supplementary Figure 6), and when considering meteorological variables lagged by 10 days (Supplementary Table 5). We found no evidence of an interaction between mean temperature and RH categorised in two levels (≤65% and >65%) using the median value of 65% as the category threshold ($p = 0.428$).

Discussion

We combined datasets of COVID-19 transmission with meteorological, demographic, socio-economic and intervention data for 409 cities in 26 countries across the world to estimate the association between meteorological factors and R_e in the early phase of the COVID-19 pandemic. We found evidence of a modest non-monotonic association of outdoor mean temperature and AH with early-phaseR_e, after controlling for potential confounders, including NPIs. Temperature explained 2.4% and AH 2.0% of the variation in R_e, compared to 13.8% explained by the OxCGRT Government Response Index in the adjusted analysis. The associations of temperature and AH with R_e were not

Table 1 Characteristics of the 26 countries included in the study.

Country	Number of cities	Reported COVID-19 cases	Mid-period	R_e	Government index
Australia	3	1747	20/03/2020	1.39	38.5
Brazil	18	17,179	10/04/2020	1.29	61.9
Canada	9	2709	21/03/2020	1.50	58.9
Chile	4	2587	27/03/2020	1.32	55.9
China	11	4178	03/02/2020	1.13	57.3
Czech Republic	1	358	21/03/2020	1.36	69.2
Ecuador	1	1014	20/03/2020	1.39	46.2
Estonia	1	209	20/03/2020	1.16	41.0
Finland	1	710	16/03/2020	1.37	30.1
France	17	5834	17/03/2020	1.51	55.8
Germany	12	7759	16/03/2020	1.43	41.1
Italy	19	11,796	11/03/2020	1.49	67.9
Japan	9	1178	12/03/2020	1.29	37.0
Kuwait	1	108	05/03/2020	1.31	21.8
Mexico	8	1894	25/03/2020	1.25	28.4
Norway	1	626	12/03/2020	1.32	16.7
Peru	1	428	18/03/2020	1.45	57.7
Philippines	2	215	21/03/2020	1.40	64.7
Singapore	1	56	15/02/2020	0.87	30.1
South Korea	4	5877	06/03/2020	1.17	54.8
Spain	52	43,331	11/03/2020	1.51	42.1
Switzerland	7	6908	13/03/2020	1.54	34.3
United Kingdom	45	9354	26/03/2020	1.41	58.0
United States	179	136,303	27/03/2020	1.45	60.7
Uruguay	1	271	19/03/2020	0.91	46.2
Vietnam	1	38	25/03/2020	1.10	45.5

The number of cities per country, total reported COVID-19 cases in the time window, mid-period of the pre-defined window of early transmission, effective reproduction number (R_e) and the lagged OxCGRT Government Response Index at the endpoint of the pre-defined window.

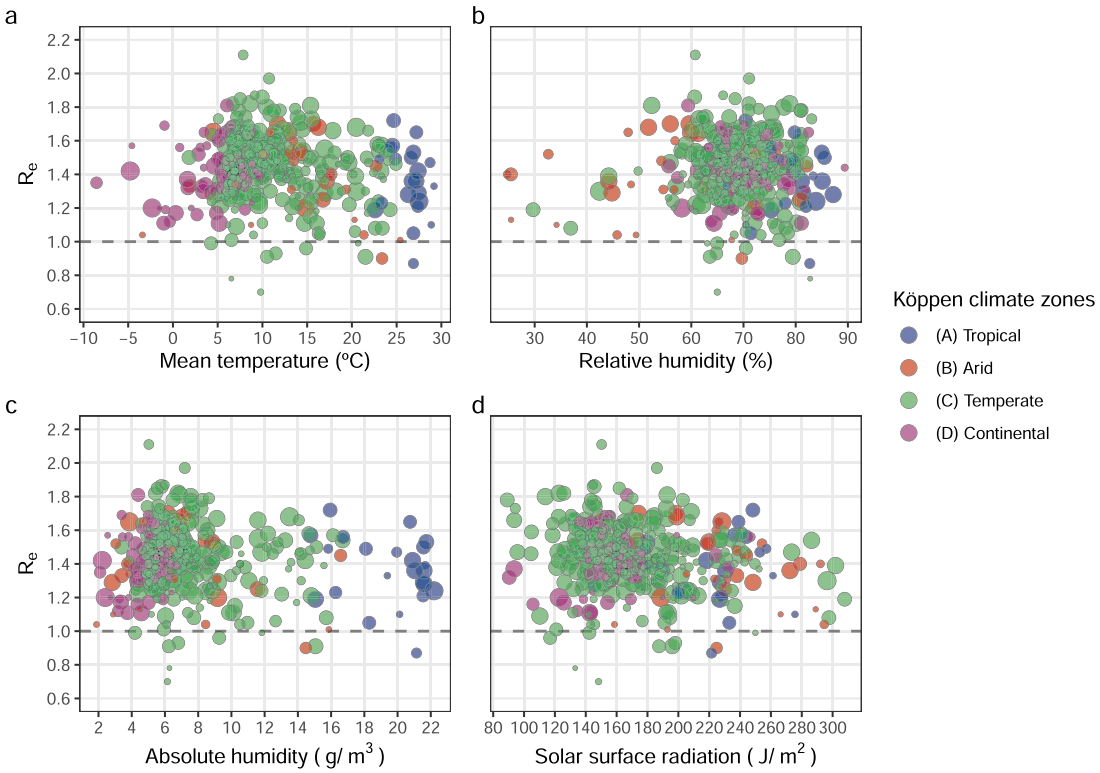


Fig. 2 Effective reproduction number vs key weather variables by climate zone. a Mean temperature (°C), **b** relative humidity (%), **c** absolute humidity (g/m^3) and **d** solar surface radiation (J/m^2) vs effective reproduction number (R_e) by climate zone (409 cities). The area of the circles is proportional to the precision (inverse of the variance) of R_e estimates.

Table 2 Association between weather variables and R_e.

+Variables	Contrast for which effect size is presented[a]	Effect size 95% CI	*P* value	Difference in the likelihood ratio R_{LR}^2 statistic
Mean temperature (°C)	10.2 vs 20	0.087 (0.025; 0.148)	0.014	+2.5
Absolute humidity (g/m^3)	6.6 vs 11	0.061 (0.011; 0.112)	0.036	+2.0
Relative humidity (%)	60 vs 75	0.043 (−0.001; 0.087)	0.058	+1.5
Surface solar radiation downwards (J/m^2)	248 vs 124	−0.053 (−0.117; 0.011)	0.208	+0.6
Wind speed (m/s)	1.1 vs 3.0	−0.038 (−0.090; 0.014)	0.152	+0.7
Total precipitation (m)	0.1 vs 6	−0.031 (−0.075; 0.014)	0.175	+0.4
OxCGRT (0–100)	21 vs 66	0.285 (0.223; 0.347)	<0.0001	+13.8

Effect size and variation explained by including, in turn, mean temperature (°C), absolute humidity (g/m^3), relative humidity (%), surface solar radiation downwards (J/m^2), wind speed (m/s), total precipitation (m) and OxCGRT (0–100) in the model of R_e. *P* values were obtained from a two-sided Wald test in the multivariable meta-regression multilevel models adjusted by population (log scale), population density (log scale), GDP (log scale), % population >65 years, PM2.5 (μg/m^3, log scale) and the OxCGRT Government Response Index, with cities nested within countries.
[a]The exposure contrast for which effect size is presented is that between the values predicting minimum and maximum R_e, where clear minima and maxima are observed (mean temperature, absolute humidity and relative humidity), otherwise the 5th to 95th percentiles.

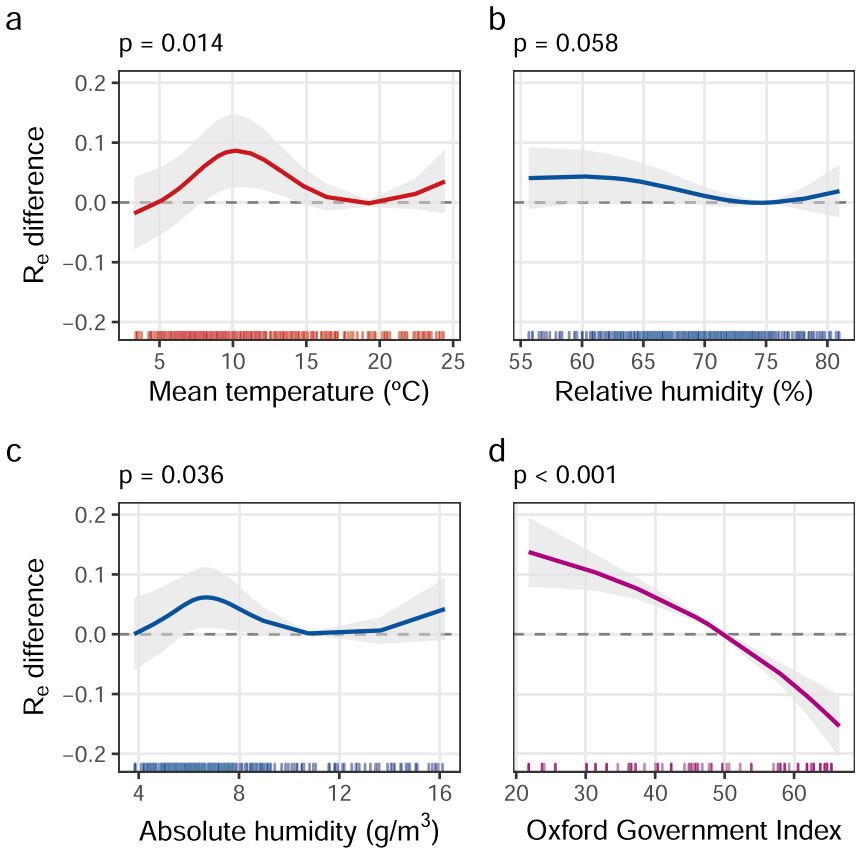


Fig. 3 Associations between weather variables, non-pharmaceutical interventions and the effective reproduction number. Non-linear associations between (**a**) mean temperature (°C), (**b**) relative humidity (%), (**c**) absolute humidity (g/m^3) and (**d**) OxCGRT Government Response Index and predicted R_e difference. Curves and their 95% confidence intervals show the predicted difference in R_e with respect to a reference value set to the value at the trough of the curve for meteorological variables (**a**–**c**), or for the OxCGRT Government Response Index = 50 (**d**). Two-sided Wald test *p* values and adjusted curves with 95% confidence intervals were obtained from multivariable meta-regression multilevel models adjusted by population (log scale), population density (log scale), GDP (log scale), % population >65 years of age, $PM_{2.5}$ (μg/m^3, log scale) and OxCGRT Government Response Index, with cities nested within countries. The marginal distribution along the *x*-axis represents the observed data for that covariate.

independent; the high correlation between them precluded control of one for the other. Overall, there was little evidence for any change in the R_e of COVID-19 associated with RH and no evidence for precipitation and wind speed.

Associations between temperature, humidity and SARS-CoV-2 transmission might be explained by three mechanisms. First, like other viruses with a lipid envelope, SARS-CoV-2 has been found to be sensitive to temperature, humidity and solar radiation under laboratory conditions[21–25], which affects its ability to survive on surfaces and in aerosols. The droplet behaviour in aerosols changes with different temperature and humidity levels. Low RH promotes the accumulation of aerosol particles (since evaporation leaves behind floating droplet nuclei) increasing the likelihood to be inhaled[26,27]. Second, innate and adaptive immune response mechanisms have been shown to be modulated by seasonal changes. Lower levels of vitamin D, mediated by decreased

ultraviolet B radiation exposure during winter might lead to impaired antiviral innate immune defences[28–30]. Breathing dry air can impair mucociliary clearance, reducing the ability of cilia cells to secrete mucus and remove viral particles (innate immune response)[27,31]. Interferon-stimulated genes, usually inducing an antiviral state as part of the innate immune response have been found to be impaired at low RH[32]. High temperatures have been shown to hinder virus-specific $CD8^+$ T cell responses and antibody production (adaptive immune system)[33]. Third, human mobility, contact patterns and time spent indoors are affected by weather conditions[34]. Very hot and very cold conditions can lead to more time spent in enclosed spaces, which might increase the likelihood of SARS-CoV-2 transmission.

Findings from this study are only partly consistent with findings from other global studies using statistical approaches to investigate meteorological effects on COVID-19 transmission. Meyer et al.[9] found that mean temperature had a modest negative association with COVID-19 incidence for temperatures above −15 °C based on a dataset of 100 countries, after controlling for surveillance capacity, time since first reported case, population density and median population age, whereas RH had a negative non-significant association with case incidence. Jüni et al.[13] covering 144 geopolitical areas showed that temperature and humidity measures were not significantly associated with epidemic growth while significant associations were found for restrictions of mass gatherings, school closures and measures of social distancing, which are consistent with our findings of a stronger impact of the OxCGRT Government Response Index compared to climatic conditions. Wu et al.[35] incorporating data from 166 countries found that a 1 °C increase in temperature and RH was associated with a 3% and 0.85% decrease in daily new cases, respectively, after controlling for wind speed, median population age, Global Health Index, Human Development Index and population density. Interestingly, non-linear associations between mean daily temperature and the instantaneous reproduction number (R_t) in the United States of America were found in a study by Rubin et al.[12] with R_t decreasing to a minimum as temperatures rose to 11 °C, increasing between 11 and 20 °C and then declining again at temperatures >20 °C. The shape of the association may be influenced by the indirect effect of weather in varying the likelihood of social interactions, e.g. at higher temperatures people may congregate in public cities, such as beaches and festivals[12], while colder temperatures could limit social activities, such as sporting events[34]. Runkle et al.[11] concluded from varying longitudinal associations in four cities that specific humidity in the range of 6–9 g/kg (i.e. AH range of 7.6–11.4 g/m^3) was a significant predictor of the COVID-19 growth rate, in line with our findings.

Unclear and inconsistent findings related to temperature and humidity may be due to methodological challenges and data limitations. Similar methodological challenges were highlighted when evaluating the association between air pollution and COVID-19 outbreaks[36,37]. The novelty of the virus, with less than a full annual cycle of data available in most places, makes it difficult to disentangle a seasonal signal or inter-annual trends from meteorological factors using time-series models[38]. Moreover, different interventions (e.g. restrictions of mass gatherings, international travel and school closures) adopted by countries at different times after the onset of local outbreaks potentially confound the association between weather variables and COVID-19 spread. These challenges have led us to consider an ecological approach where we compared the outbreak curve early in the epidemic, minimising the confounding effect of NPIs. Despite this, we found a strong association of the OxCGRT Government Response Index with R_e, confirming the importance of interventions implemented early on in the epidemic in controlling COVID-19 dynamics[39].

Comparing the early-phase outbreak curves in different countries is challenging given that countries have different case definitions, and early COVID-19 data only captured a small portion of cases, mainly hospitalised patients or individuals with severe symptoms. The estimated high proportion of asymptomatic cases compromises the use of COVID-19 case counts to estimate transmission dynamics[40]. We used an estimated response variable, i.e. the effective reproduction number, calculated accounting for reporting delays and other sources of uncertainty. The 20-day duration was chosen as a compromise between needing enough days for a more precise R_e estimation, while, at the same time, limiting the window to provide more constant weather, case ascertainment and R_e estimates within the window. A larger window would bias estimates in ways that cannot be readily adjusted for. Our meta-analysis approach accounts for the uncertainty in R_e estimates, which in turn reduces the level of certainty in the results. Further, 20 days is ~4 generations of infections, which, under most reporting scenarios, is sufficient to be confident about estimates in the level of transmission. We assume that within the 20-day time window, the case definition is constant within a city or country and R_e is not affected by differences in case definition between countries.

A clear strength of this study is the use of an extensive dataset of 409 cities, representing 44.8% of all cumulative reported COVID-19 cases registered by 31 May 2020 in the John Hopkins University Coronavirus Resource Center. Our analysis covers all major climate zones across the globe, ranging from temperate, continental to tropical and dry climate settings. Another strength is our flexible methodological and statistical approach. We used multilevel meta-analytic models that take into account uncertainty of the response variable, i.e. the effective reproduction number. The model allowed for possible non-linearity of the exposures, and we adjusted for a selection of key socio-economic and demographic factors, as well as using a random effect to account for the country- and city-level differences. We chose covariates based on their potential impact on viral transmission that might confound the examined association of weather and COVID-19 dynamics. Indeed, population density leads to higher contact rates, potentially increasing the likelihood of transmission[41]. The age structure of a population is relevant given that elderly people were found to be more susceptible to infection and more likely to experience clinical symptoms of COVID-19 compared to younger age groups, increasing the likelihood of seeking medical care and getting tested[42]. Moreover, differences in contact patterns among different age groups can further affect the number of COVID-19 cases in each age group[42]. Socio-economic indicators, such as GDP per capita, are important to consider as more deprived populations might be at higher risk of infection due to potential conditions of overcrowded accommodation, inability to work from home or limited access to medical care[43]. Also, among air quality factors, a positive association between $PM_{2.5}$ and COVID-19 incidence and mortality has been reported[44,45].

We investigated model uncertainties with several sensitivity analyses, e.g. excluding cities with $R < 1$, excluding China and Brazil, cities in the southern hemisphere, cities with a latitude lower than 45°, and cities with an OxCGRT Government Response Index of more than 60. Previous studies compared cities within a country or considered large geographical units[13,35,46], which could lead to a limited exposure range with narrow temperature and humidity variability reported during winter seasons, or high measurement errors for meteorological variables defined over large geographical units. We considered small area/city units distributed among 26 countries worldwide, allowing a good exposure range and minimising the measurement error of the exposures.

Our study has several important limitations in addition to those already discussed. Cities in the northern hemisphere were overrepresented compared to southern hemisphere cities, which indicates that the findings might be more representative for cities in the global north. Our results need to be put into the context of complex uncertainties surrounding characteristics of the novel virus, such as incomplete knowledge on possible underlying mechanisms between weather conditions and the virus itself, the role of host immunity and the potential influence of weather-sensitive human behaviour, such as indoor crowding[47]. However, AH was found to demonstrate the strongest indoor-to-outdoor correlation, indicating that outdoor AH measures could reflect indoor conditions[48,49]. Data limitations regarding the novel virus, including varying accuracy of COVID-19 case numbers, limited data availability across cities and temporal constraints of an incomplete seasonal cycle of SARS-CoV-2 contribute to the limitations of this analysis.

Despite these limitations, the associations of weather with R_e in this study suggests that such effects are likely to be small compared to other drivers of transmission. NPIs had a stronger impact on variation in transmission between cities than meteorological variables. We found no weather conditions in which transmission is impeded if precautions (social distancing, mask use, etc.) are not taken. These results support the statement that, to date, COVID-19 interventions are critical regardless of meteorological conditions.

Methods

Data. Data in this study were obtained from a well-established MCC Collaborative Research Network[50]. The current MCC network covers 750 locations (cities or regions) in 43 countries/regions. For this study, 26 countries provided a daily time series of COVID-19 cases for a total of 502 locations (cities or small regions). COVID-19 data were downloaded from a publicly available repository or obtained from health agencies (Supplementary Table 1) and data management was performed using Microsoft Excel 2019. The time series from 1 January 2020 to 31 May 2020 comprises 2,771,137 COVID-19 cases, representing 44.8% of the cumulative cases registered by 31 May 2020 in the Johns Hopkins database (https://coronavirus.jhu.edu/). Supplementary Table 1 shows the sources used for each country along with the definition of COVID-19 cases.

To limit potential confounding by NPIs and temporal variation in case ascertainment, we selected a 10–20-day window early in the epidemic, starting after at least ten confirmed cases had occurred in a 10-day period, to reduce noise introduced by imported cases. We excluded days for which the OxCGRT Government Response Index exceeded 70, accepting reduced windows down to 10 days in length. The OxCGRT collates publicly available information on 18 indicators about governments' policy responses to the COVID-19 pandemic (https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker). These indicators are categorised as containment or closure policies (e.g. school and workplace closures, restrictions on gatherings and movement), economic policies (e.g. income support) or health policies (e.g. COVID-19 testing programmes). The OxCGRT Government Response Index aggregates these indicators into a single score between 0 and 100 and provides a measure of how many policies a government has enacted, and to what degree. We chose 70 as the maximum value of OxCGRT Government Response Index allowable as a compromise between limiting confounding by government interventions and including enough cities to enable estimation of the associations studied (see Supplementary text 1). Applying these conditions/restrictions reduced our dataset to 409 cities or small regions in 26 countries with an observation period between 11 January 2020 and 28 April 2020.

Most of the 409 cities are situated in the northern hemisphere ($n = 381$), and in temperate (n = 292) or continental (n = 65) climatic zones, with few cities located in tropical ($n = 23$) and dry ($n = 29$) climatic zones. The COVID-19 cases were observed in the early phase of the epidemics, ranging from the first week of February 2020 in China to mid-April 2020 in Brazil (Supplementary Figure 1). This early epidemic phase is characterised in many countries (except Uruguay and Singapore) by a reproduction number >1 (Table 1).

We estimated R_e for infections in the time window of interest using EpiNow2 1.3.2[18]. This R package implements a Bayesian latent variable method for estimating R_e, where infections at time t are estimated using the sum of previous infections, weighted by an uncertain, gamma-distributed, generation time probability mass function, and multiplied by an estimate of R_e[51,52]. Initial infections (prior to the first reported case) were estimated using a log-linear model with priors based on the observed growth in cases. Complete infection trajectories were then mapped to reported case counts by first convolving over the incubation period distribution and an estimated distribution representing the delay between symptom onset and case report (both assumed to be log-normal). Reporting noise was then added using a negative binomial observation model combined with a multiplicative day of the week effect (modelled using a simplex). R_e was considered to be piecewise constant with a breakpoint 3 days into the time window. The R_e estimate from the first 3 days of the window was discarded with the R_e estimate from the remainder of the window used in all analyses. Each region was fitted independently using Markov chain Monte Carlo. Four chains were used with a warmup of 1000 samples and 4000 samples post warmup. Convergence was assessed using the R hat diagnostic[53].

We used a gamma-distributed generation time with a mean of 3.6 days (standard deviation (SD) 0.7) and a SD of 3.1 days (SD 0.8)[54,55]. This generation time was slightly shorter than the consensus estimate reported by Ferretti et al.[56], leading to our R_e estimates and subsequent effect sizes being conservative. We used a log-normally distributed prior for the incubation period with a mean of 5.2 days (SD 1.1) and SD of 1.52 days (SD 1.1)[57]. The log-normal prior for the delay from symptom onset to case report was estimated globally using a subsampled Bayesian bootstrapping approach (with 100 subsamples each using 250 samples) using data from an international line list of cases. The resulting distribution had a mean of 6.4 days and a standard deviation (SD) of 17 days (or a log mean of 0.83 (SD 0.15) and a log SD 1.44 (SD 0.12). The subsampled bootstrap approach was used to incorporate both the temporal and spatial uncertainty in the reporting distribution as data specific to each setting and time point was not available.

To define our exposures, we considered the following time series from the ERA5 dataset: 2 m temperature, 2 m dewpoint temperature, surface solar radiation downwards, precipitation, and 10 m eastward (u) and northward (v) components of wind. These are published by the Copernicus Climate Change service on a regular latitude/longitude grid of 0.25° (~25 km × 25 km) in NetCDF format[19]. The hourly 2 m temperature, 2 m dewpoint temperature and surface solar radiation downwards were averaged for each day to derive daily mean temperature, dewpoint temperature and surface solar radiation. From dewpoint temperature and the corresponding temperature (T; °C) we obtained RH (%) using the R 'humidity' 0.1.5 package[58]. The following formula was used to calculate AH (g/m^3), which represents the mass of water vapour in the air mixture[59]:

$$\mathrm{AH} = (6.112 \times e^{(17.67 \times T)/(T+243.5)} \times 2.1674 \times \mathrm{RH})/(273.15 + T).$$

The hourly 10 m u and v components of wind were averaged for each day, and the daily average u and v components were used to compute the wind speed using the formula wind speed = sqrt($u^2 + v^2$). Hourly precipitation data were summed to derive daily totals. The daily variables were calculated for each 25 km^2 grid cell and assigned to a city if the city centroid fell within the grid cell.

Mean temperature (and other meteorological variables, Supplementary Table 3) observed during the city-specific time window reflect the late winter/early spring observation period in cities situated in the northern hemisphere and in temperate or continental climatic zones. We found a high correlation between mean temperature and AH (Supplementary Figure 2). Socio-economic and demographic characteristics were extracted from the OECD Regional and Metropolitan database[60] and Worldcities database[61] (Supplementary Table 2). We selected, a priori, the following set of confounders: total population, population density, % elderly population (>65 years) and GDP (per capita). Pollution data ($PM_{2.5}$) for the observation period (10–20 days) was obtained from the Copernicus Atmosphere Monitoring Service global near-real-time service[62–64]. This product provides hourly modelled values of surface $PM_{2.5}$ (μg/m^3) at a 0.4 × 0.4 arc degrees grid cell resolution. The hourly time series were averaged over the observation period and linked to the city using the city centroid coordinates. Cities vary in terms of socio-demographic characteristics (Supplementary Table 3). The correlation between socio-demographic characteristics is shown in Supplementary Figure 3 and the correlation between meteorological variables, OxCGRT Government Response Index, day of the year and R_e in Supplementary Figure 4. To account for differences in NPIs we used the OxCGRT Government Response Index[65]. In this study, we considered the 10 days lagged value of the OxCGRT Government Response Index, and for each city, we assigned the index on the last day of the specified window for each city[39]. Note, in our analysis, meteorological variables and socio-demographic covariates were collated and summarised at the city level, while the COVID-19 time series were defined at the smallest administrative level containing the city. We only included cities for which the COVID-19 time series were available for an area in which most of the population resided in that city. We, therefore, refer to our unit of analysis as a city.

Statistical analysis. For descriptive purposes, the following statistics (mean, standard deviation and range) were calculated for meteorological variables (mean temperature, AH, RH, surface solar radiation, wind speed, total precipitation) and covariates considered in this study (total population, population density, % elderly population (>65 years), GDP (per capita), $PM_{2.5}$, OxCGRT Government Response Index). We also calculated the correlation (Pearson coefficient) among meteorological variables and among covariates.

The association between city-level covariates and climatic variables with the effective reproduction number were evaluated using multilevel meta-regression models with two levels (cities nested within countries)[20] using the R 'mixmeta' 1.1.0 package. The inclusion of country as a random effect allowed the model to

account for country differences (e.g. data reporting) with efficient use of the within- and between-country information. Moreover, the meta-regression models allowed us to consider the precision of the R_e estimates, as estimated by its variance, giving less weight to more imprecise estimates for shorter time windows.

Firstly, we used two-level meta-regression models to evaluate the possible non-linear association between each meteorological variable and the reproduction number R_e. We considered possible non-linearity in the association with R_e using a natural spline parameterisation of the meteorological variables with a variable number of internal knots from 0 (linear term) to 5, placed at respective percentiles of the variable. We compare the models with different non-linear parameterisations of the meteorological variable using the Akaike Information Criteria (AIC), choosing models with the lowest AIC.

We fitted the following two-level random-effects meta-regression models with cities nested within countries and an increasing number of predictors; Model A with two random effects (cities and countries) and the intercepts, Model B including the OxCGRT Government Response Index, Model C considering also total population, density, GDP, % population older than 65 years, and $PM_{2.5}$ (total population, density, GDP and $PM_{2.5}$ were log-transformed due to the skewness of their distribution).

Then for each meteorological variable, we fitted two-level meta-regression models (D1–D6) with the meteorological variable as exposure and total population, density, GDP, % population older than 65 years, $PM_{2.5}$ and the OxCGRT Government Response Index as covariates. We considered non-linearity in the association with R_e using a natural spline parameterisation of the climatic variables with the number of internal knots as determined in the univariate analysis. The coefficients of the natural spline parameterisation of the meteorological variable were used to derive the plot of the association between the meteorological variable and R_e in the 5–95th percentile of the meteorological variable distribution (Fig. 3 and Supplementary Figures 5 and 6). The coefficients of the natural spline parameterisation of the meteorological variable were also used to test the association between the meteorological variable and R_e using the multivariate Wald test. All the tests were two-sided. Given the small number of pre-defined exposures variables, no adjustment was made for multiple comparisons.

We quantified heterogeneity between cities with standard measures of I^2 [66]. These measures are estimated once from a meta-regression model without meta-predictors (Model A) and once from the meta-regression models (Models B, C and D1–D6) to assess the reduction in residual heterogeneity provided by the different set of predictors. For each model, we calculated the likelihood ratio test (R_{LR}^2) statistic[67]. R_{LR}^2 is calculated as $1 - \exp(-2/409 \times (\log \mathrm{Lik_m} - \log \mathrm{Lik_0})$, where $\log \mathrm{Lik_m}$ is the log-likelihood of the model of interest and $\log \mathrm{Lik_0}$ is the log-likelihood from a null model including only city and country random effect (i.e. Model A). For each meteorological variable, we calculated the difference in the likelihood ratio test R^2 (R_{LR}^2) with respect to Model C (including random effects, OxCGRT Government Response Index and city-level covariates). For OxCGRT and city-level covariates, the R_{LR}^2 represents the reduction in R_{LR}^2 when dropping OxCGRT or city-level covariates from Model D1 with temperature and all other terms (i.e. random effects, OxCGRT and city-level covariates).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

COVID-19 data were downloaded from publicly available repositories or obtained from health agencies (Supplementary Table 1). COVID-19 data for Australia, Brazil, Canada, Chile, China, Czech Republic, Estonia, Finland, Germany, Italy, Kuwait, Mexico, Norway, Peru, Philippines, Romania, South Korea, Spain, United Kingdom, United States and Vietnam are publicly available. COVID-19 data for Japan and Singapore are available upon request. COVID-19 Data for France, Switzerland and Uruguay were obtained by a specific request to health agencies and are not publicly available.

Meteorological variables (mean temperature, dewpoint temperature, solar radiation, wind components and precipitation) were derived from ERA5 reanalysis product 'https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset'.

Pollution levels (PM2.5) was derived from CAMS near real time 'https://apps.ecmwf.int/datasets/data/cams-nrealtime/levtype=sfc/'.

The OxCGRT Government Response Index was downloaded from the public repository: https://github.com/OxCGRT/covid-policy-tracker/raw/master/data/OxCGRT_latest.csv (downloaded 3 August 2020).

Socio-economic and demographic characteristics were extracted from the OECD Regional and Metropolitan database 'https://www.oecd.org/regional/regional-policy/regionalstatisticsandindicators.htm' and Worldcities database.

Data were processed and harmonised at the city level. The city-level data used in the main and supplementary analysis of the paper are available in the GitHub directory: https://github.com/fsera/COVIDWeather/ [68].

Code availability

The code developed in the study to perform the city-level main analysis is available in the following GitHub repository[68].

For each meteorological variable, the effect size was calculated using predicted curves from multivariable meta-regression multilevel models. We calculated the difference in the likelihood ratio test R^2 (R_{LR}^2) with respect to a model including random effects, OxCGRT Government Response Index, and city-level covariates (Model C, Supplementary Table 4). R_{LR}^2 is calculated as $1 - \exp(-2/409 \times (\log \mathrm{Lik_m} - \log \mathrm{Lik_0})$, where $\log \mathrm{Lik_m}$ is the log-likelihood of the model of interest and $\log \mathrm{Lik_0}$ is the log-likelihood from a null model including only city and country random effect (i.e., Model A, Supplementary Table 4). For OxCGRT, the R_{LR}^2 represents the reduction in R_{LR}^2 when dropping OxCGRT from the model with temperature and all other terms (i.e., random-effects and city-level covariates).

Received: 18 February 2021; Accepted: 8 September 2021;
Published online: 13 October 2021

References

1. Moriyama, M., Hugentobler, W. J. & Iwasaki, A. Seasonality of respiratory viral infections.*Annu. Rev. Virol.* **7**, 83–101 (2020).
2. Lowen, A. C. & Steel, J. Roles of humidity and temperature in shaping influenza seasonality. *J. Virol.* **88**, 7692–7695 (2014).
3. Carlson, C. J., Gomez, A. C. R., Bansal, S. & Ryan, S. J. Misconceptions about weather and seasonality must not misguide COVID-19 response. *Nat. Commun.* **11**, 4312 (2020).
4. Baker, R. E., Yang, W., Vecchi, G. A., Metcalf, C. J. E. & Grenfell, B. T. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. *Science* **369**, 315–319 (2020).
5. O'Reilly, K. M. et al. Effective transmission across the globe: the role of climate in COVID-19 mitigation strategies. *Lancet Planet. Health* **4**, e172 (2020).
6. Zeka, A. et al. Responding to COVID-19 requires strong epidemiological evidence of environmental and societal determining factors. *Lancet Planet. Health* **4**, e375–e376 (2020).
7. Adhikari, A. & Yin, J. Short-term effects of ambient ozone, PM2.5, and meteorological factors on COVID-19 confirmed cases and deaths in Queens, New York. *Int. J. Environ. Res. Public Health* **17**, 4047 (2020).
8. Hoang, T. & Tran, T. T. A. Ambient air pollution, meteorology, and COVID-19 infection in Korea. *J. Med. Virol.* **93**, 878–885 (2021).
9. Meyer, A., Sadler, R., Faverjon, C., Cameron, A. R. & Bannister-Tyrrell, M. Evidence that higher temperatures are associated with a marginally lower incidence of COVID-19 cases. *Front. Public Health* **8**, 367 (2020).
10. Pequeno, P. et al. Air transportation, population density and temperature predict the spread of COVID-19 in Brazil. *PeerJ* **8**, e9322 (2020).
11. Runkle, J. D. et al. Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. *Sci. Total Environ.* **740**, 140093 (2020).
12. Rubin, D. et al. Association of social distancing, population density, and temperature with the instantaneous reproduction number of SARS-CoV-2 in counties across the United States. *JAMA Netw. Open* **3**, e2016099 (2020).
13. Jüni, P. et al. Impact of climate and public health interventions on the COVID-19 pandemic: a prospective cohort study. *Can. Med. Assoc. J.* **192**, E566–E573 (2020).
14. Carleton, T., Cornetet, J., Huybers, P., Meng, K. C. & Proctor, J. Global evidence for ultraviolet radiation decreasing COVID-19 growth rates. *Proc. Natl Acad. Sci. USA* **118**, e2012370118 (2021).
15. Smit, A. J. et al. Winter is coming: a Southern Hemisphere perspective of the environmental drivers of SARS-CoV-2 and the potential seasonality of COVID-19. *Int. J. Environ. Res. Public Health* **17**, 5634 (2020).
16. Mecenas, P., Bastos, R. T., da, R. M., Vallinoto, A. C. R. & Normando, D. Effects of temperature and humidity on the spread of COVID-19: A systematic review. *PLoS ONE* **15**, e0238339 (2020).
17. Briz-Redón, Á. & Serrano-Aroca, Á. The effect of climate on the spread of the COVID-19 pandemic: a review of findings, and statistical and modelling techniques. *Prog. Phys. Geogr. Earth Environ.* **44**, 591–604 (2020).
18. Abbott, S., Hellewell, J., Munday, J., Thompson, R. & Funk, S. EpiNow: estimate realtime case counts and time-varying epidemiological parameters. *Zenodo* https://doi.org/10.5281/zenodo.3957489 (2020).
19. Hersbach, H. et al. *ERA5 Hourly Data on Single Levels from 1979 to Present* (Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2018).
20. Sera, F., Armstrong, B., Blangiardo, M. & Gasparrini, A. An extended mixed-effects framework for meta-analysis. *Stat. Med.* **38**, 5429–5444 (2019).
21. Ratnesar-Shumate, S. et al. Simulated sunlight rapidly inactivates SARS-CoV-2 on surfaces. *J. Infect. Dis.* **222**, 214–222 (2020).
22. Schuit, M. et al. Airborne SARS-CoV-2 is rapidly inactivated by simulated sunlight. *J. Infect. Dis.* **222**, 564–571 (2020).
23. Biryukov, J. et al. Increasing temperature and relative humidity accelerates inactivation of SARS-CoV-2 on surfaces. *mSphere* **5**, e00441–20 (2020).

24. Chan, K.-H. et al. Factors affecting stability and infectivity of SARS-CoV-2. *J. Hosp. Infect.* **106**, 226–231 (2020).
25. Dabisch, P. et al. The influence of temperature, humidity, and simulated sunlight on the infectivity of SARS-CoV-2 in aerosols. *Aerosol Sci. Technol.* **55**, 142–153 (2021).
26. Zhao, L., Qi, Y., Luzzatto-Fegiz, P., Cui, Y. & Zhu, Y. COVID-19: effects of environmental conditions on the propagation of respiratory droplets. *Nano Lett.* **20**, 7744–7750 (2020).
27. Lowen, A. C., Mubareka, S., Steel, J. & Palese, P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* **3**, e151 (2007).
28. Cannell, J. J. et al. Epidemic influenza and vitamin D. *Epidemiol. Infect.* **134**, 1129–1140 (2006).
29. Grant, W. B. et al. Evidence that vitamin D supplementation could reduce risk of influenza and COVID-19 infections and deaths. *Nutrients* **12**, 988 (2020).
30. Tamerius, J. et al. Global influenza seasonality: reconciling patterns across temperate and tropical regions. *Environ. Health Perspect.* **119**, 439–445 (2011).
31. Sun, Z., Thilakavathy, K., Kumar, S. S., He, G. & Liu, S. V. Potential factors influencing repeated SARS outbreaks in China. *Int. J. Environ. Res. Public. Health* **17**, 1633 (2020).
32. Kudo, E. et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. *Proc. Natl Acad. Sci. USA* **116**, 10905–10910 (2019).
33. Moriyama, M. & Ichinohe, T. High ambient temperature dampens adaptive immune responses to influenza A virus infection. *Proc. Natl Acad. Sci. USA* **116**, 3118–3125 (2019).
34. Fares, A. Factors influencing the seasonal patterns of infectious diseases. *Int. J. Prev. Med.* **4**, 128–132 (2013).
35. Wu, Y. et al. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci. Total Environ.* **729**, 139051 (2020).
36. Villeneuve, P. J. & Goldberg, M. S. Methodological considerations for epidemiological studies of air pollution and the SARS and COVID-19 coronavirus outbreaks. *Environ. Health Perspect.* **128**, 095001 (2020).
37. Heederik, D. J. J., Smit, L. A. M. & Vermeulen, R. C. H. Go slow to go fast: a plea for sustained scientific rigour in air pollution research during the COVID-19 pandemic. *Eur. Respir. J.* **56**, 2001361 (2020).
38. Imai, C., Armstrong, B., Chalabi, Z., Mangtani, P. & Hashizume, M. Time series regression model for infectious disease and weather. *Environ. Res.* **142**, 319–327 (2015).
39. Liu, Y., Morgenstern, C., Kelly, J., Lowe, R. & Jit, M. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BMC Med.* **19**, 40 (2021).
40. Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
41. Rocklöv, J. & Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* **27**, taaa038 (2020).
42. Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).
43. Patel, J. A. et al. Poverty, inequality and COVID-19: the forgotten vulnerable. *Public Health* **183**, 110–111 (2020).
44. Borro, M. et al. Evidence-based considerations exploring relations between SARS-CoV-2 pandemic and air pollution: involvement of PM2.5-mediated up-regulation of the viral receptor ACE-2. *Int. J. Environ. Res. Public. Health* **17**, 5573 (2020).
45. Pozzer, A. et al. Regional and global contributions of air pollution to risk of death from COVID-19. *Cardiovasc. Res.* **116**, 2247–2253 (2020).
46. Jiang, Y., Wu, X.-J. & Guan, Y.-J. Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infect. Control Hosp. Epidemiol.* **41**, 1011–1015 (2020).
47. Ryti, N. R. I., Korpelainen, A., Seppänen, O. & Jaakkola, J. J. K. Paradoxical home temperatures during cold weather: a proof-of-concept study. *Int. J. Biometeorol.* **64**, 2065–2076 (2020).
48. Marr, L. C., Tang, J. W., Van Mullekom, J. & Lakdawala, S. S. Mechanistic insights into the effect of humidity on airborne influenza virus survival, transmission and incidence. *J. R. Soc. Interface* **16**, 20180298 (2019).
49. Nguyen, J. L., Schwartz, J. & Dockery, D. W. The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. *Indoor Air* **24**, 103–112 (2014).
50. Gasparrini, A. et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet* **386**, 369–375 (2015).
51. Abbott, S. et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Res.* **5**, 112 (2020).
52. Sherratt, K. et al. Exploring surveillance data biases when estimating the reproduction number: with insights into subpopulation transmission of Covid-19 in England. *Philos. Trans. R. Soc. B* **376**, 20200283 (2021).
53. Stan Development Team. *RStan: The R interface to Stan* (Stan Development Team, 2020).
54. Abbott, S. et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Res.* **5**, 112 (2020).
55. Ganyani, T. et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Eurosurveillance* **25**, 2000257 (2020).
56. Ferretti, L. et al. The timing of COVID-19 transmission. Preprint at *medRxiv* https://doi.org/10.1101/2020.09.04.20188516 (2020).
57. Lauer, S. A. et al. The incubation period of Coronavirus Disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
58. Jun, C. Calculate water vapor measures from temperature and dew point. https://github.com/caijun/humidity (2019).
59. Shi, P. et al. The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak - evidence from China. Preprint at *medRxiv* https://doi.org/10.1101/2020.03.22.20038919 (2020).
60. OECD. OECD regions at a glance 2016. https://doi.org/10.1787/reg_glance-2016-en (2016).
61. Simplemaps. World cities database. https://simplemaps.com/data/world-cities (2016).
62. Christophe, Y. et al. *Validation Report of the CAMS Near-Real-Time Global Atmospheric Composition Service: Period March–May 2019*. Copernicus Atmosphere Monitoring Service (CAMS) Report (2019).
63. Morcrette, J.-J. et al. Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: forward modeling. *J. Geophys. Res. Atmos.* **114**, https://doi.org/10.1029/2008JD011235 (2009).
64. Benedetti, A. et al. Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. *J. Geophys. Res. Atmos.* **114**, https://doi.org/10.1029/2008JD011115 (2009).
65. Hale, T. et al. *Variation in Government Responses to COVID-19*. Blavatnik School of Government Working Paper. www.bsg.ox.ac.uk/covidtracker (2021).
66. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
67. Kramer, M. R2 statistics for mixed models. *Conf. Appl. Stat. Agric.* https://doi.org/10.4148/2475-7772.1142 (2005).
68. Sera, F., Abbott, S. & Royé. Data and code to replicate the analysis of the paper. 'A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries'. https://doi.org/10.5281/zenodo.5215842 (2021).

Acknowledgements

This work was generated using Copernicus Climate Change Service (C3S) and Copernicus Atmosphere Monitoring Service (CAMS) information [2020]. The authors would like to thank the European Centre for Medium-Range Weather Forecasts (ECMWF) that implements the C3S and CAMS on behalf of the European Union. D.R. was supported by a postdoctoral research fellowship of the Xunta de Galicia (Spain). A.G. was funded by the Medical Research Council-UK (Grant ID: MR/R013349/1), the Natural Environment Research Council UK (Grant ID: NE/R009384/1) and the European Union's Horizon 2020 Project Exhaustion (Grant ID: 820655). R.L. was supported by a Royal Society Dorothy Hodgkin Fellowship. S.A. and S.M. were funded by the Wellcome Trust (grant 210758/Z/18/Z210758/Z/18/Z). The following funding sources are acknowledged as providing funding for the MCC Collaborative Research Network authors: J.K. and A.U. were supported by the Czech Science Foundation, project 18-22125S. S.T. was supported by the Shanghai Municipal Science and Technology Commission (Grant 18411951600). N.S. is supported by the NIEHS-funded HERCULES Center (P30ES019776). H.K. was supported by the National Research Foundation of Korea (BK21 Center for Integrative Response to Health Disasters, Graduate School of Public Health, Seoul National University). A.S., F.D.R. and S.R. were funded by the European Union's Horizon 2020 Project Exhaustion (Grant ID: 820655). Each member of the CMMID COVID-19 Working Group contributed to processing, cleaning and interpretation of data, interpreted findings, contributed to the manuscript and approved the work for publication. The following funding sources are acknowledged as providing funding for the CMMID COVID-19 working group authors. This research was partly funded by the Bill & Melinda Gates Foundation (INV-001754: M.Q; INV-003174: K.P., M.J., Y.L., J.L.; NTD Modelling Consortium OPP1184344: C.A.B.P., G.M.; OPP1180644: S.R.P.; OPP1183986: E.S.N.). BMGF (OPP1157270: K.E.A.). DFID/Wellcome Trust (Epidemic Preparedness Coronavirus research programme 221303/Z/20/Z: C.A.B.P.). EDCTP2 (RIA2020EF-2983-CSIGN: H.P.G.). ERC Starting Grant (#757699: M.Q.). This project has received funding from the European Union's Horizon 2020 research and innovation programme—project EpiPose (101003688: K.P., M.J., P.K., R.C.B., W.J.E., Y.L.). This research was partly funded by the Global Challenges Research Fund (GCRF) project 'RECAP' managed through RCUK and ESRC (ES/P010873/1: A.G., C.I.J., T.J.).

HDR UK (MR/S003975/1: R.M.E.). MRC (MR/N013638/1: N.R.W.; MR/V027956/1: W.W.). Nakajima Foundation (A.E.). This research was partly funded by the National Institute for Health Research (NIHR) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care (16/136/46: B.J.Q.; 16/137/109: B.J.Q., F.Y.S., M.J., Y.L.; Health Protection Research Unit for Immunisation NIHR200929: N.G.D.; Health Protection Research Unit for Modelling Methodology HPRU-2012-10096: T.J.; NIHR200908: R.M.E.; NIHR200929: F.G.S., M.J.; PR-OD-1017-20002: A.R., W.J.E.). Royal Society (Dorothy Hodgkin Fellowship: R.L.; RP \EA\180004: P.K.). UK DHSC/UK Aid/NIHR (PR-OD-1017-20001: H.P.G.). UK MRC (MC_PC_19065—Covid 19: Understanding the dynamics and drivers of the COVID-19 epidemic using real-time outbreak analytics: A.G., N.G.D., R.M.E., S.C., T.J., W.J.E., Y.L.; MR/P014658/1: G.M.K.). Authors of this research receive funding from the UK Public Health Rapid Support Team funded by the United Kingdom Department of Health and Social Care (T.J.). Wellcome Trust (206250/Z/17/Z: A.J.K., T.W.R.; 206471/Z/17/Z: O.B.; 208812/Z/17/Z: S.C.; 210758/Z/18/Z: J.D.M., J.H., N.I.B.; UNS110424: F.K.). No funding (A.M.F., A.S., C.J.V.-A., D.C.T., J.W., K.E.A., Y.-W.D.C.). LSHTM, DHSC/UKRI COVID-19 Rapid Response Initiative (MR/V028456/1: Y.L.). Innovation Fund of the Joint Federal Committee (01VSF18015: F.K.). Foreign, Commonwealth and Development Office/Wellcome Trust (221303/Z/20/Z: M.K.).

Author contributions

Conceptualisation—ideas; formulation or evolution of overarching research goals and aims: F.S., B.A., S.A., K.O'R., M.H., M.P., A.T., A.M.V.-C., A.G., R.L. Data curation—management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use: F.S., B.A., S.A. and S.M. Formal analysis—application of statistical, mathematical, computational or other formal techniques to analyse or synthesise study data: F.S., B.A. and S.A. Funding acquisition—acquisition of the financial support for the project leading to this publication: AG. Investigation—conducting a research and investigation process, specifically performing the experiments or data/evidence collection: F.S., S.A., S.M., R.L., R.S. and MCC Collaborative Research Network. Methodology—development or design of methodology; creation of models: F.S., B.A., S.A. and A.G. Project administration—management and coordination responsibility for the research activity planning and execution: F.S. and R.L. Resources—provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources or other analysis tools: A.G. Software—programming, software development; designing computer programmes; implementation of the computer code and supporting algorithms; testing of existing code components: F.S., B.A., S.A. and A.G. Supervision—oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team: F.S., B.A. and R.L. Validation—verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs: F.S. and S.A. Visualisation—preparation, creation and/or presentation of the published work, specifically visualisation/data presentation: F.S., B.A., S.A., R.L. and D.R. Writing—original draft—preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation): F.S., R.L., B.A. and R.v.B. Writing—review and editing—preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision—including pre- or post-publication stages: F.S., B.A., S.A., S.M., K.O'R., R.v.B., R.S., D.R., M.H., M.P., A.T., A.M.V.-C., A.G., R.L., MCC Collaborative Research Network and CMMID COVID-19 Working Group.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-25914-8.

Correspondence and requests for materials should be addressed to Francesco Sera or Rachel Lowe.

Peer review information *Nature Communications* thanks Uriel Kitron, Olivier Restif and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

MCC Collaborative Research Network

Wenbiao Hu[19], Shilu Tong[19,20,21,22], Eric Lavigne[23,24], Patricia Matus Correa[25], Xia Meng[26], Haidong Kan[26], Jan Kynčl[27,28], Aleš Urban[29,30], Hans Orru[31], Niilo R. I. Ryti[32,33], Jouni J. K. Jaakkola[32,33], Simon Cauchemez[34], Marco Dallavalle[35], Alexandra Schneider[35], Ariana Zeka[36], Yasushi Honda[37,38], Chris Fook Sheng Ng[11], Barrak Alahmad[39], Shilpa Rao[40], Francesco Di Ruscio[40], Gabriel Carrasco-Escobar[41,42], Xerxes Seposo[11], Iulian Horia Holobâcă[43], Ho Kim[44], Whanhee Lee[44], Carmen Íñiguez[45], Martina S. Ragettli[46,47], Alicia Aleman[48], Valentina Colistro[49], Michelle L. Bell[50], Antonella Zanobetti[39], Joel Schwartz[39], Tran Ngoc Dang[51], Noah Scovronick[52], Micheline de Sousa Zanotti Stagliorio Coélho[53], Magali Hurtado Diaz[54] & Yuzhou Zhang[55,56]

[19]School of Public Health and Social Work, Queensland University of Technology, Brisbane, QLD, Australia. [20]Shanghai Children's Medical Centre, School of Medicine, Shanghai Jiao-Tong University, Shanghai, China. [21]School of Public Health, Institute of Environment and Human Health, Anhui Medical University, Hefei, China. [22]Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China. [23]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada. [24]Air Health Science Division, Health Canada, Ottawa, ON, Canada. [25]Department of Public Health, Universidad de los Andes, Santiago, Chile. [26]Key Lab of Public Health Safety of the Ministry of Education and NHC Key Lab of Health Technology Assessment, School of Public Health, Fudan University, Shanghai, China. [27]Department of Infectious Diseases Epidemiology, National Institute of Public Health, Prague, Czech Republic. [28]Department of Epidemiology and Biostatistics,

Third Faculty of Medicine, Charles University, Prague, Czech Republic. [29]Institute of Atmospheric Physics of the Czech Academy of Sciences, Prague, Czech Republic. [30]Faculty of Environmental Sciences, Czech University of Life Sciences, Prague, Czech Republic. [31] Institute of Family Medicine and Public Health, University of Tartu, Tartu, Estonia. [32]Center for Environmental and Respiratory Health Research (CERH), University of Oulu, Oulu, Finland. [33]Medical Research Center Oulu (MRC Oulu), Oulu University Hospital and University of Oulu, Oulu, Finland. [34]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France. [35]Institute of Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health (GmbH), Neuherberg, Germany. [36]Institute of Environment, Health and Societies, Brunel University London, London, UK. [37]Center for Climate Change Adaptation, National Institute for Environmental Studies, Tsukuba, Japan. [38]Faculty of Health and Sport Sciences, University of Tsukuba, Tsukuba, Japan. [39]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA. [40]Norwegian Institute of Public Health, Oslo, Norway. [41]Health Innovation Laboratory, Institute of Tropical Medicine "Alexander von Humboldt", Universidad Peruana Cayetano Heredia, Lima, Peru. [42]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. [43]Faculty of Geography, Babes-Bolyai University, Cluj-Napoca, Romania. [44]Department of Public Health Science, Graduate School of Public Health, Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea. [45]Department of Statistics and Computational Research, Universitat de València, València, Spain. [46]Swiss Tropical and Public Health Institute, Basel, Switzerland. [47]University of Basel, Basel, Switzerland. [48]Departament of Preventive and Social Medicine, School of Medicine, Universidad de la República, Montevideo, Uruguay. [49]Departmente of Cuantitative Methods, School of Medicine, Universidad de la República, Montevideo, Uruguay. [50]School of the Environment, Yale University, New Haven, CT, USA. [51]Department of Environmental Health, Faculty of Public Health, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam. [52]Gangarosa Department of Environmental Health. Rollins School of Public Health, Emory University, Atlanta, GA, USA. [53]Department of Pathology, Faculty of Medicine, University of São Paulo, São Paulo, Brazil. [54]Department of Environmental Health, National Institute of Public Health, Cuernavaca Morelos, Mexico. [55]College of Computer Science and Technology, Zhejiang University, Hangzhou, China. [56]Department of Research, Baolue Technology (Zhejiang) Co., Ltd, Hangzhou, China.

CMMID COVID-19 Working Group

Timothy W. Russell[3], Mihaly Koltai[3], Adam J. Kucharski[3], Rosanna C. Barnard[3], Matthew Quaife[3], Christopher I. Jarvis[3], Jiayao Lei[3], James D. Munday[3], Yung-Wai Desmond Chan[3], Billy J. Quilty[3], Rosalind M. Eggo[3], Stefan Flasche[3], Anna M. Foss[3], Samuel Clifford[3], Damien C. Tully[3], W. John Edmunds[3], Petra Klepac[3], Oliver Brady[3], Fabienne Krauer[3], Simon R. Procter[3], Thibaut Jombart[3], Alicia Rosello[3], Alicia Showering[3], Sebastian Funk[3], Joel Hellewell[3], Fiona Yueqian Sun[3], Akira Endo[3], Jack Williams[3], Amy Gimma[3], Naomi R. Waterlow[3], Kiesha Prem[3], Nikos I. Bosse[3], Hamish P. Gibbs[3], Katherine E. Atkins[3], Carl A. B. Pearson[3], Yalda Jafari[3], C. Julian Villabona-Arenas[3], Mark Jit[3], Emily S. Nightingale[3], Nicholas G. Davies[3], Kevin van Zandvoort[3], Yang Liu[3], Frank G. Sandmann[3], William Waites[3], Kaja Abbas[3], Graham Medley[3] & Gwenan M. Knight[3]

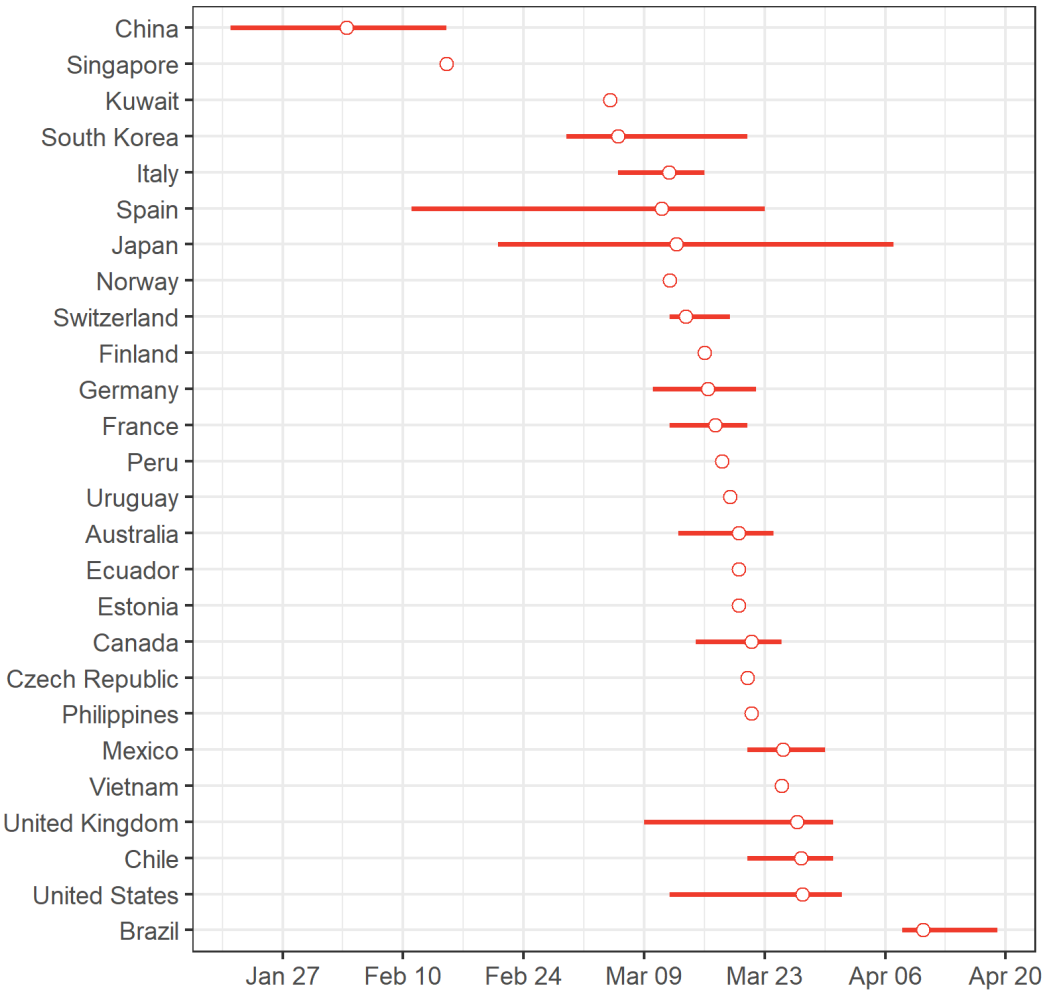
A cross-sectional analysis of meteorological factors and SARS-CoV-2 transmission in 409 cities across 26 countries

Supplementary materials

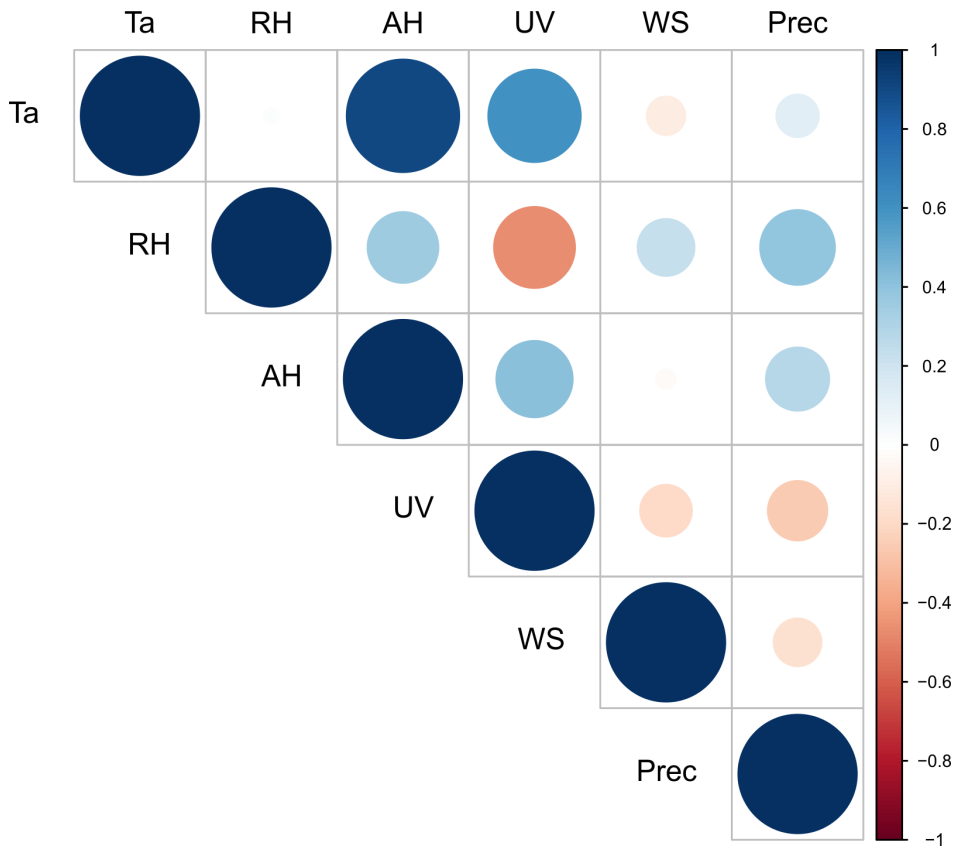
Supplementary Methods

Dependence of sample size on maximum Oxford Government Response Tracker

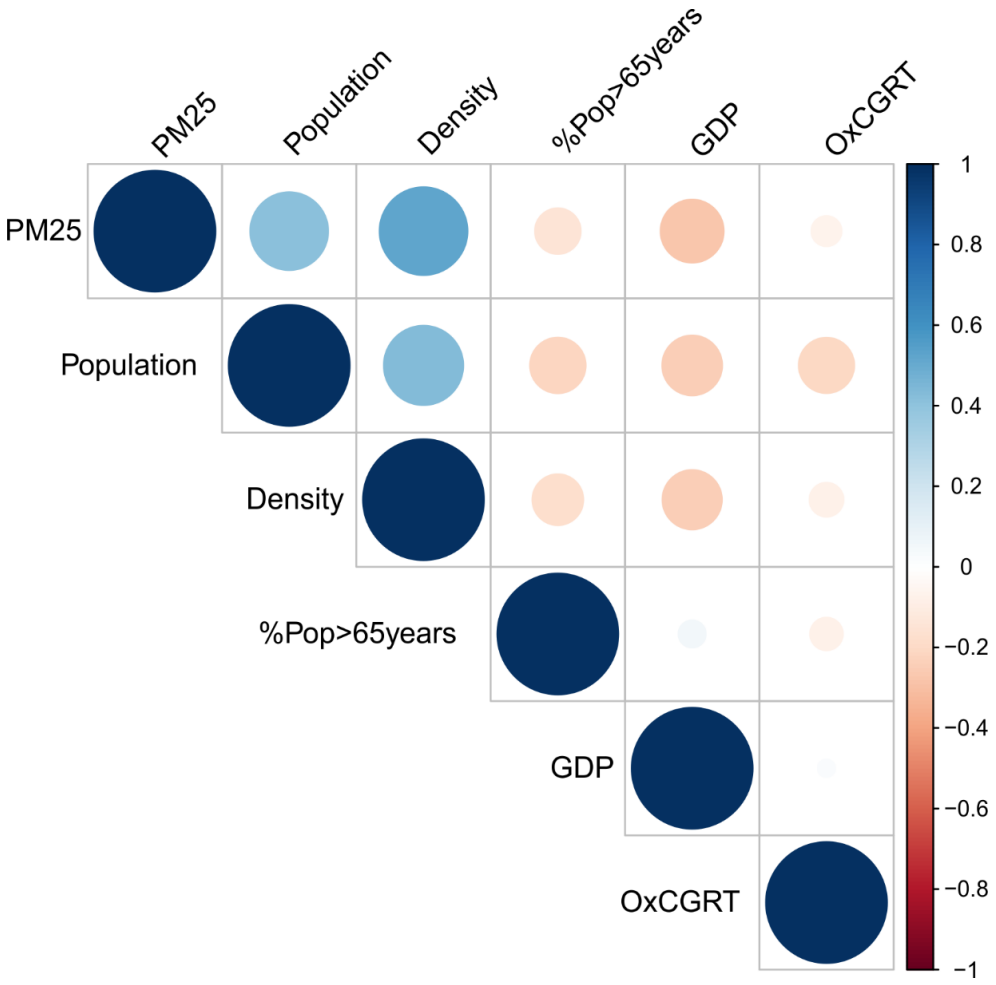
We chose 70 as the maximum value of OxCGRT Government Response Index allowed in included days as a compromise between limiting confounding by government interventions and including enough cities to enable estimation of the associations studied. This choice was informed by the preliminary evaluation (see Supplementary Figure 7). Supplementary Figure 7a shows the % of the 502 total cities for which data was available, according to the chosen criteria (window length between 10-20 days and there were at least 10 cases) for given maximum OxCGRT Government Response Index values (ranging from 60 to 100). Supplementary Figure 7b shows the dependence of the number of days included in windows for different cut-off values. Windows were also required to include at least 10 days and to begin only when 10 cases had occurred. For this purpose, each day's OxCGRT Government Response Index value was lagged ten days, to allow for the incubation period and reporting delays. As the OxCGRT Government Response Index cut-off was lowered, the number of cities included and of days included in windows diminished. The sharp rise in the number of cities included by increasing the maximum allowed OxCGRT Government Response Index from 60 to 70 with diminishing increases beyond that suggested 70 as a sensible compromise. During the analysis we checked the possible residual confounding role of the capped OxCGRT index by including the value at the end of the time window (lagged by 10 days) as covariate in our model. After observing its strong effect, we retained this variable for all further analyses.



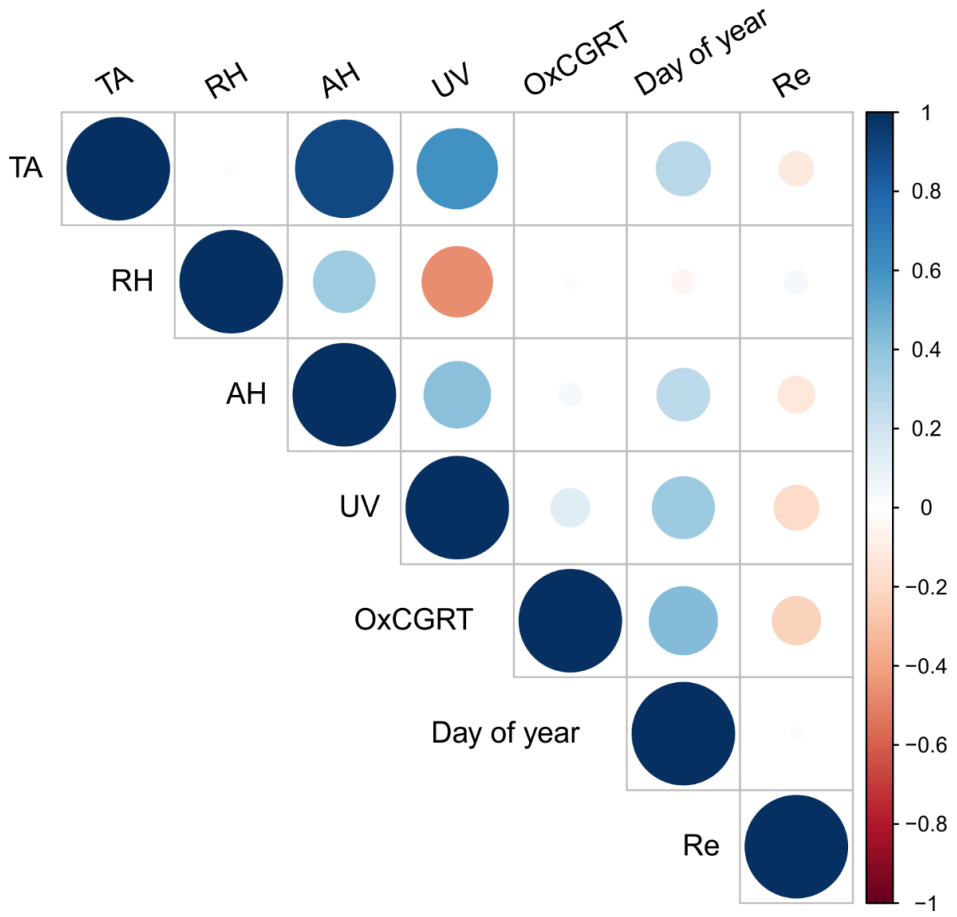
Supplementary Figure 1. Range (line) and mean (dot) observation day (midpoint of the time-window) for the cities within each of the 26 countries.



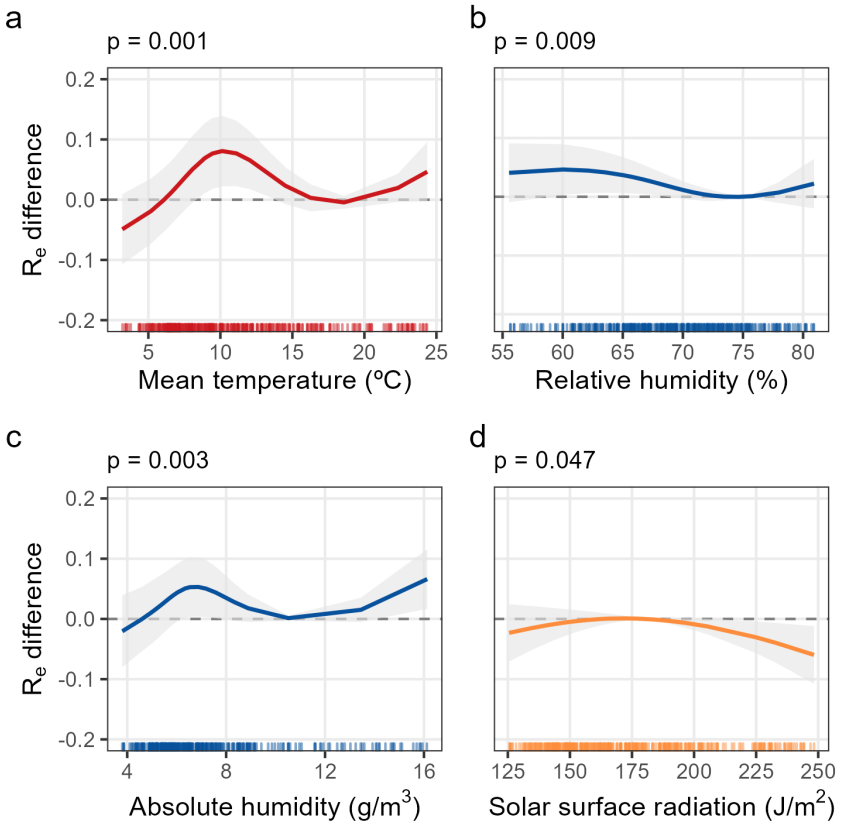
Supplementary Figure 2. Correlations between meteorological variables (Ta = Temperature, RH = Relative Humidity, AH=Absolute Humidity, UV=Surface solar radiation, WS=Wind speed, Prec=Total precipitations).



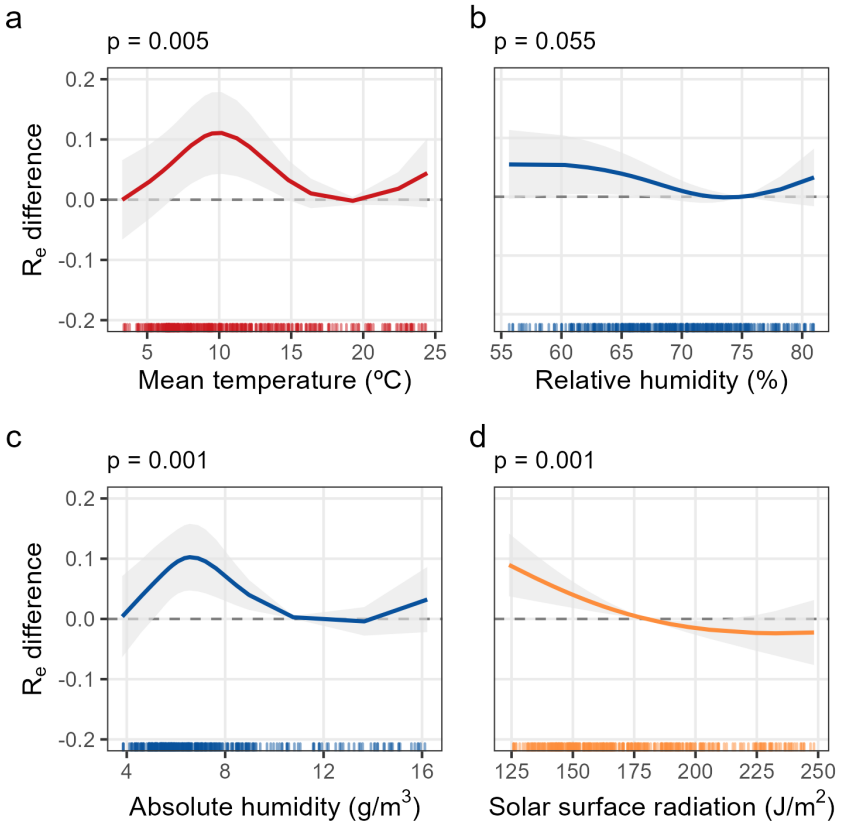
Supplementary Figure 3. Correlations a between city-level socio-demographic variables.



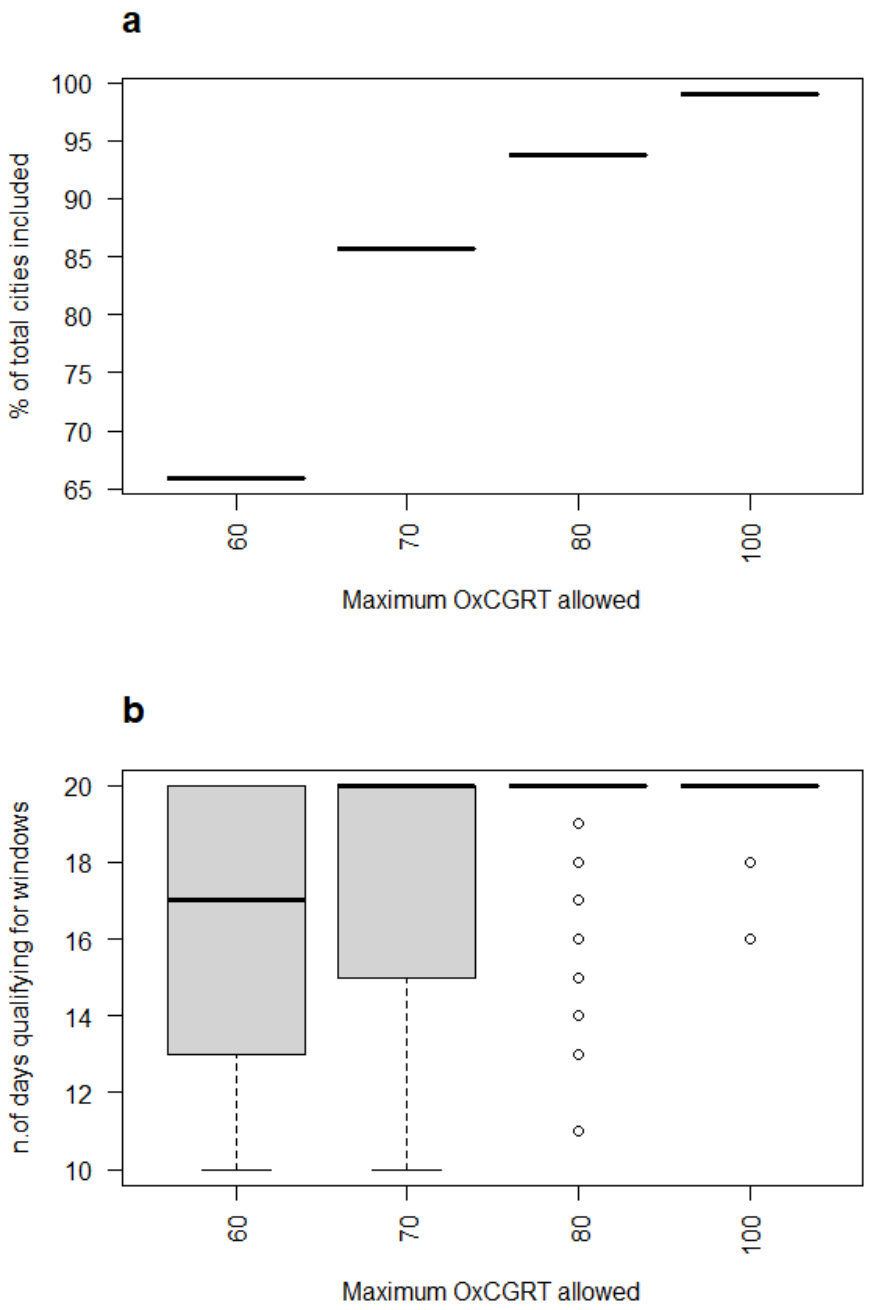
Supplementary Figure 4. Correlations between meteorological variables (Ta = Air temperature, RH = Relative humidity, AH=Absolute humidity, UV = Surface solar radiation), OxCGRT Government Response Index, day of the year, and Reproduction number (R_e).



Supplementary Figure 5. Associations (with 95% confidence intervals) between (a) mean temperature (°C), (b) relative humidity (%), (c) absolute humidity (g/m3), and (d) Solar surface radiation with predicted Re difference when cities with $R_e < 1$ were excluded. Two-sided Wald test p-values and adjusted curves with 95% confidence intervals were obtained from multivariable meta-regression multilevel models adjusted by population (log scale), population density (log scale), GDP (log scale), % population >65 years of age, PM2.5 (µg/m^3, log scale) and OxCGRT Government Response Index, with cities nested within countries. The marginal distribution along the x axis represents the observed data for that covariate.



Supplementary Figure 6. Associations (with 95% confidence intervals) between (a) mean temperature (°C), (b) relative humidity (%), (c) absolute humidity (g/m3), and (d) Solar surface radiation with predicted Re difference when non-pharmaceutical interventions were not controlled for in the model. Two-sided Wald test p-values and adjusted curves with 95% confidence intervals were obtained from multivariable meta-regression multilevel models adjusted by population (log scale), population density (log scale), GDP (log scale), % population >65 years of age, PM2.5 (µg/m^3, log scale), with cities nested within countries. The marginal distribution along the x axis represents the observed data for that covariate.



Supplementary Figure 7. (a) % of the 502 total cities for which data was available according to the chosen criteria (window length between 10-20 days and there were at least 10 cases) for given maximum OxCGRT Government Response Index values (ranging from 60 to 100). (b) number of days included in windows for a given cut-off value.

Supplementary Table 1. Source and COVID-19 case definition for the different countries.

Country	Start date	End date	City definition	No. of cities	Source	Case Definition
Australia	22/01/2020	04/06/2020	City	3	Health department website	Confirmed COVID-19 cases
Brazil	25/02/2020	04/06/2020	Municipality	18	https://covid.saude.gov.br/	The new confirmed COVID-19 numbers take into account the cases recorded from the previous day
Canada	25/01/2020	06/06/2020	Health Regions	17	https://github.com/ishaberry/Covid19Canada Berry I, Soucy J-PR, Tuite A, Fisman D. Open access epidemiologic data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. CMAJ. 2020 Apr	The COVID-19 data includes confirmed and presumptive positive (i.e, probable) cases of COVID-19.
Chile	03/03/2020	12/06/2020	Regions	4	https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Chile	Confirmed COVID-19 cases

China	22/01/2020	04/06/2020	City	17	nCov19 package in R	Confirmed COVID-19 cases
Czech Republic	29/02/2020	09/06/2020	Regions	1	The Ministry of Health of the Czech Republic - https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19 Komenda M., Bulhart V., Karolyi M., et al. Complex reporting of coronavirus disease (COVID-19) epidemic in the Czech Republic: use of interactive web-based application in practice. *Journal of Medical Internet Research. 2020*, **22**(5), e19367.	RT-PCR confirmed cases per day
Ecuador	12/03/2020	15/05/2020	Provinces	2	Health authority	Confirmed COVID-19 cases
Estonia	26/03/2020	03/06/2020	County	1	Estonian Health Board - https://www.terviseamet.ee/et/koroonaviirus/avaandmed	Confirmed cases by clinical laboratory diagnostic tests.

Finland	01/03/2020	31/05/2020	Hospital districts	1	Finnish institute of health and welfare (THL)	All cases confirmed by laboratory testing. The date in the time-series refers to the date of taking the test
France	28/01/2020	08/06/2020	Departments	17	Santé publique France; data.gouv.fr	Until 19/3/2020 Confirmed cases. From 20/3/2020 Daily number of newly hospitalized persons

Germany	28/01/2020	31/05/2020	City	12	„Fallzahlen in Deutschland" of the Robert Koch-Institut (RKI) - Link to the dataset: https://www.arcgis.com/home/item.html?id=f10774f1c63e40168479a1feb6c7ca74	"Confirmed cases by clinical laboratory diagnostic tests. Infections confirmed by laboratory diagnostic evidence in case of a non-matching clinical picture (e.g. asymptomatic) are also included.
Italy	24/02/2020	04/06/2020	Provinces	23	Protezione civile	Confirmed COVID-19 cases
Japan	16/01/2020	31/05/2020	Prefectures	10	Health authority	Confirmed COVID-19 cases
Kuwait	22/01/2020	04/06/2020	Country	1	COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University	RT-PCR positive nasopharyngeal swab

Mexico	01/01/2020	04/06/2020	States	8	https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico	Confirmed COVID-19 cases
Norway	21/02/2020	26/05/2020	City	1	https://www.fhi.no/sv/smittsomme-sykdommer/corona/dags--og-ukerapporter/dags--og-ukerapporter-om-koronavirus/	Confirmed COVID-19 cases
Peru	06/03/2020	05/06/2020	Departments	18	Ministry of Health Peru (https://www.datosabiertos.gob.pe/group/datos-abiertos-de-covid-19)	Confirmed COVID-19 cases; test date
Philippines	09/03/2020	11/06/2020	City	4	https://doh.gov.ph/covid19tracker	RT-PCR confirmed cases per day
Romania	22/03/2020	31/05/2020	County	8	PRESS RELEASE, Strategic Communication Group, MINISTRY OF INTERNAL AFFAIRS	new cases of people infected with SARS – CoV – 2 (COVID – 19) these being cases that had not previously had a positive test

Singapore	23/01/2020	16/06/2020	City	1	Ministry of Health Singapore. (https://www.moh.gov.sg/covid-19/past-updates , https://www.moh.gov.sg/covid-19/situation-report)	Dates of confirmed COVID-19 cases
South Korea	20/01/2020	31/05/2020	Provinces	7	From http://ncov.mohw.go.kr/	People who diagnostic test positive for the virus, regardless of clinical manifestations. All confirmed cases were registered in the KCDC Health and Disease Integrated Management System.

Spain	31/01/2020	21/06/2020	Provinces	52	https://cnecovid.isciii.es/covid19/#documentación-y-datos	Confirmed cases with clinical symptoms of acute respiratory infection of any severity with fever, cough or feeling of shortness of breath (other symptoms such as onychophagia, anosmia, ageustia, muscle pain, diarrhoea, chest pain or headache can also be considered) and with a positive result from a Diagnostic Test of Active Infection by SARS-CoV-2

Switzerland	01/01/2020	25/05/2020	Cantons	8	Federal Office of Public Health (FOPH, https://www.bag.admin.ch/bag/en/home.html) ; Federal Statistical Office (FSO, https://www.bfs.admin.ch/bfs/en/home.html)	Confirmed COVID-19 cases; date of testing
United Kingdom	30/01/2020	31/05/2020	LTLA	54	Public health England	The date in the time-series refers to the date the specimen was taken from the person being tested
United States	22/01/2020	04/06/2020	City	211	COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University	Confirmed COVID-19 cases
Uruguay	24/02/2020	15/06/2020	Departments	1	Epidemiology Section of the Ministry of Health	Date start of symptoms
Vietnam	23/01/2020	19/06/2020	Provinces	2	Health authority	Confirmed COVID-19 cases

Supplementary Table 2. City-level socio-economic, demographic and pollution indicators.

Indicator	**Source**	**Year**
Total population (persons)	Worldcities database	2015
Population density (persons per km^2)	Worldcities database	2015
Population, % (population > 65 years)	OECD	2018
GDP per capita (US$)	OECD	2016;2018
$PM_{2.5}$ ($\mu g/m^3$)	CAMS Near real time	2020 (Covid window)

Supplementary Table 3. Mean, standard deviation (SD) and range (min and max) of the effective reproduction number, meteorological and city-level variables calculated in the 409 cities.

Variable	Mean	SD	min	max
Outcome				
R_e	1.43	0.19	0.70	2.11
Meteorological				
Mean temperature (°C)	11.27	6.66	-8.54	29.18
Relative humidity (%)	68.49	8.86	24.74	89.38
Absolute humidity (g/m^3)	7.65	3.97	1.88	22.19
Surface solar radiation downwards (J/m^2)	175.36	40.17	89.08	307.79
Wind speed (km/h)	2.88	1.15	0.63	7.30
Total precipitation (m/day)	2.46	2.23	0.00	21.07
City characteristics				
$PM_{2.5}$ ($\mu g/m^3$)	10.0	9.5	3.4	87.1
Total population (persons)	1,309,744.7	3,154,340.3	3,478.0	2,6174,599.0
Population density (persons per km^2)	4,157.4	5,397.8	42.0	53,108.1
Population, % pop >65 years	13.1	4.5	3.0	27.2
GDP per capita (US$)	37,752.2	14,922.3	3,168.0	101,375.0
Gross Value Added per capita (US$)	70,981.0	21,227.4	14,647.0	366,027.5

Non pharmaceutical Interventions				
Oxford Government Index	55.2	13.8	5.8	69.9

Supplementary Table 4. Sequence of multilevel meta-regression models.

Model	**Model terms**	**LogLik**	**likelihood ratio test R^2**	**I^2 (%)**	**Two-sided p-value**
Model A	Only random effects (city and country)	107.6483		66.2	
Model B	Model A + Oxford government index	149.0627	18.3	62.7	<0.0001
Model C	Model B + City-level covariates	153.8145	20.2	55.8	<0.0001
Model D1	Model C + Mean temperature (°C)	160.1787	22.7	53.3	0.014
Model D2	Model C + Relative humidity (%)	157.6298	21.7	55.3	0.058
Model D3	Model C + Absolute humidity (g/m^3)	159.0228	22.2	53.3	0.036
Model D4	Model C + Surface solar radiation downwards (J/m^2)	155.4314	20.8	55.7	0.208
Model D5	Model C + Wind speed (m/s)	155.7435	21.0	55.0	0.152
Model D6	Model C + Total precipitation (m/day)	154.7431	20.6	55.4	0.175
Model D7	Model D1 without Oxford government index	126.6178	8.9	58.6	
Model D8	Model D1 without City-level covariates	156.6395	21.3	59.3	

likelihood ratio test R^2 calculated as $1-\exp(-2/409 \times (\log Lik_m - \log Lik_0))$, where logLikm is the log-likelihood of the model of interest and logLik0 is the log-likelihood from a null model including only city and country random effect (i.e. Model A).

Supplementary Table 5. Sensitivity analysis: p values for each experiment.

	Mean temperature (°C)	Absolute humidity (g/m^3)	Relative humidity (%)	Surface solar radiation downwards (J/m^2)	Wind speed (m/s)	Total precipitation (m)
Model presented in Table 2 (main text) (n = 409)	0.014	0.036	0.058	0.208	0.152	0.175
Cities with OxCGRT <60 (n = 129)	0.001	0.038	0.454	0.370	0.018	0.036
No Adjustment by OxCGRT (n = 409)	0.005	0.001	0.055	0.001	0.202	0.158
No lagged OxCGRT (n = 409)	0.016	0.035	0.107	0.260	0.132	0.256
Country as fixed effect (n = 409)	0.018	0.011	0.058	0.297	0.148	0.155
10 days lagged exposure variables (n = 409)	0.001	0.126	0.037	0.009	0.722	0.209
Models also adjusted by day of the year (n = 409)*	0.015	0.036	0.060	0.210	0.151	0.174
Only Cities with R>=1 (n = 399)	0.001	0.003	0.009	0.047	0.186	0.182
Excluding China and Brazil (n = 380)	0.011	0.156	0.049	0.332	0.189	0.880
Non tropical cities (n = 386)	0.019	0.201	0.063	0.199	0.185	0.699
Tropical cities (n = 23)	0.198	0.667	0.882	0.501	0.633	0.880
Northern hemisphere (n = 381)	0.036	0.355	0.054	0.294	0.192	0.774
Southern hemisphere (n = 28)	0.456	0.666	0.606	0.992	0.223	0.992
Only cities with latitude < 45 degrees (n = 308)	0.021	0.055	0.066	0.211	0.028	0.221

p values were obtained from multivariable meta-regression multilevel models adjusted by population (log scale), density (log scale) , GDP (log scale), % population > 65 years, $PM_{2.5}$ (log scale) and OxCGRT government response index with cities nested within countries.

*p values were obtained from multivariable meta-regression multilevel models adjusted by population (log scale), population density (log scale), GDP (log scale), % population > 65 years, $PM_{2.5}$ (log scale), OxCGRT oxford government response index and day of the year, with cities nested within countries.

Part III

Discussion

Chapter 8

Final comments

In this final chapter, I provide some conclusive comments about my research on the extended random-effects framework for complex meta-analysis and its application in environmental epidemiological studies. In the first Section 8.1 I will present the main outputs of this PhD and their relationship with the PhD's objectives. In the second Section 8.2 I will present the contribution to the scientific literature of the output of this PhD project. In the third Section 8.3, I will examine possible future developments from computational, theoretical, and applied perspectives. A final discussion is then provided in Section 8.4.

8.1 Outputs of the PhD and their relationship with PhD Objectives

The first objective of this PhD was to develop an extended random-effects framework for complex meta-analysis. Coherently with this aim, the first methodological output of this PhD is the research paper (Sera *et al.*, 2019a) reported in Chapter 3. In this paper, I illustrate a general framework for meta-analysis based on linear mixed-effects models, where potentially complex patterns of effect sizes are modelled through an extended and flexible structure of fixed and random terms. As shown in the next section 8.2 the development of a general framework with its implementation in freely

available statistical software has facilitated the application of complex meta-analytic models in several fields. The second methodological paper of this PhD (Sera and Gasparrini, 2022) has been presented in Chapter 4. According to objective 2 of this PhD, this paper illustrates extensions of the classical two-stage study design used in environmental epidemiology. I demonstrate the advantages of applying various extensions of the two-stage design in multiple case studies using multi-location analyses of environmental exposures. Among the benefits a more robust characterisation of the non-linear association curve at different analysis units (e.g. cities, countries), a more flexible parametrisation of the temporal structure of the association curves and the possibility of exploring differences of association curves among sub-group (e.g. age, sex). In Chapters 5, 6 and 7, I presented three applications of the extended framework for complex meta-analysis (Sera *et al.*, 2019b, 2020, 2021). Coherently with Objective 4 of this PhD, these papers illustrate the application of the extended random-effect framework for complex meta-analysis in environmental epidemiology studies. These studies are characterised by complex settings with multiple levels of hierarchies, non-linearity, spatial and temporal structure and demonstrate the advantages of the extended two-stage design using the complex meta-analytic framework. Following Objective 3 of this PhD, The extended random-effects framework for complex meta-analysis has been implemented in the R package mixmeta. The availability of fully-documented packages in a freely-available software has been fundamental for this PhD. The flexibility of the package allows to represent of different parametrisations of complex meta-analytic problems as shown in the methodological (Sera and Gasparrini, 2022; Sera *et al.*, 2019a) and more applied papers (Sera *et al.*, 2020, 2021) of this PhD. These papers include as supplementary material methodological notes, data, and R scripts for reproducing the examples presented in the papers. The availability of detailed documentation has promoted the application of the techniques by other research teams, as shown in the next section 8.2 and I am confident that the material can support future research implementing extensions of the classical two-stage study design in environmental epidemiology.

8.2 Impact of the PhD research project

The first methodological output of this PhD is the research paper in Chapter 3 presenting the general framework for meta-analysis based on linear mixed effects. The development of a general framework

with its implementation in freely available statistical software has facilitated the application of complex meta-analytic models in several fields. At the time of writing, I identified 53 papers that cited the research paper Sera *et al.* (2019a), 21 of which I co-authored. The extended framework was used to evaluate the short term effects of environmental stressors within the Multi-Country Multi-City (MCC) scientific collaboration network (Chen *et al.*, 2021; Liu *et al.*, 2022; Madaniyazi *et al.*, 2022; Masselot *et al.*, 2021; Meng *et al.*, 2021; Mistry *et al.*, 2022; Royé *et al.*, 2021; Tobías *et al.*, 2021; Urban *et al.*, 2021; Vicedo-Cabrera *et al.*, 2020, 2021). The MCC is a scientific collaboration network that shares resources and data to investigate the health effect of environmental stressors. In this setting characterised by time-series data for more than 750 cities in 40 countries, the flexibility of the extended framework allowed to deal with non-linear exposure-response curve and multilevel structure of the data (with cities nested within countries), e.g., fitting multivariate multilevel meta-regression models. Using these models, we obtained more reliable country and city-level BLUPs health impact estimates. Moreover, the extended framework was used in COVID-19, environmental and clinical related projects within my research network (Donzelli *et al.*, 2022; Formica *et al.*, 2021; Gasparrini *et al.*, 2022; Huber *et al.*, 2022; Lavigne *et al.*, 2022; Nottmeyer and Sera, 2021; Onozuka *et al.*, 2022; de Schrijver *et al.*, 2022; Scortichini *et al.*, 2020; Sim *et al.*, 2020).

Outside my research network, the complex meta-analyitc framework and software have been used in environmental (Bär *et al.*, 2022; Chu *et al.*, 2021; Chua *et al.*, 2021; Fong and Smith, 2022; Hasegawa *et al.*, 2022; He *et al.*, 2022; Madaniyazi *et al.*, 2021; Martínez-Solanas *et al.*, 2021; Wen *et al.*, 2022a,b), and ecological (Beck *et al.*, 2022; Gerli *et al.*, 2020; James *et al.*, 2021; Junqueira *et al.*, 2022; Tuttle and Donahue, 2022; Wohner *et al.*, 2022) studies.

The methods presented in the methodological paper Sera *et al.* (2019a) have been also applied in meta-analysis performed in clinical setting Agrawal *et al.* (2022); Allinson *et al.* (2022); Borbón *et al.* (2022); Dickerson-Young *et al.* (2022); Filippini *et al.* (2021); Jeon *et al.* (2021); Wieland *et al.* (2021), and in behavioural sciences Barth *et al.* (2022); Lintner (2022); Tarai *et al.* (2022); Walton *et al.* (2021).

Lastly, my work has been discussed in several methodological papers (Joshi *et al.*, 2021; Orsini, 2021; Srinivasjois, 2021; Wu *et al.*, 2022; Yu *et al.*, 2021).

The second methodological paper of this PhD has been presented in Chapter 4. This paper illustrates

extensions of the classical two-stage study design used in environmental epidemiology. The paper has just been published, and no impact metric is available, but there is some interest from the scientific community with more than 1100 accesses to the paper according to the journal metrics, and I am confident that the paper can support future research implementing extensions of the classical two-stage study design in environmental epidemiology.

In Chapters 5, 6 and 7 I presented three applications of the extended framework for complex meta-analysis. These papers have been well received from the scientific community, and in spite of being recently published, they already feature a high number of citations. In particular, the paper on modifiers of heat health effects Sera *et al.* (2019b), published in 2019, now has 65 citations according to SCUPUS (at time of writing). Notably, it was included in the 2020 report of the Lancet Countdown on health, and climate change (Watts *et al.*, 2021). The paper on the longitudinal health impact of air conditioning Sera *et al.* (2020) has 25 citations according to Dimensions Citations, and the ecological analysis of COVID-19 Sera *et al.* (2021), published in the autumn of 2021, has 22 citations according to Dimensions Citations.

This impact shows the importance of the research I carried out within my PhD project and a strong motivation to develop the work further.

8.3 Future developments

In this PhD project, I developed the extended random-effects framework for complex meta-analysis and defined extensions of the two-stage design in environmental epidemiology. This methodological framework has been applied in several epidemiological projects. This work has stimulated some reflections on the possible future development of this research area. These future developments can be divided into the need to improve the feasibility of using the proposed methodology in a big data setting, the possibility of considering the uncertainty of the random terms during the inferential process, and further applications in environmental epidemiology and other research areas. These directions will be discussed in the following three sections.

Big data applications

In the research paper Sera *et al.* (2019a), I showed how several types of meta-analysis (univariate, multivariate, multilevel, and dose-response) could be modeled as mixed effects models. I used a likelihood-based computational approach to obtain the estimates, and this approach is implemented in the package R package mixmeta.

The computational approach implemented in mixmeta uses the block-diagonal structure of the within and between error terms at the outer level. On applying the method, I observed that by increasing the number of inner level units, the computational time increases non-linearly. The computational time becomes an issue when the number of inner level units is on the order of hundreds. For example, in a two-level model with ten outer level units and 500 inner level units, the computational time is 75 seconds for each iteration of the Newton-Ramphson-based procedure. This problem is even more relevant with other software that could be used to fit complex meta-analytic models. For example, the R package metafor that uses the full marginal covariance matrix in the Fisher-scoring estimation procedure, or the R package nlme based on the QR decomposition operating at each nested level Pinheiro and Bates (2006). For example, if we considered a two levels models with 10 outer level units and 100 inner level units the R package mixmeta converge in 18.3 seconds, while R package metafor in 193.3 seconds, and the R package nlme in 1121.9 seconds. Some improvement could be reached writing C++ version of some functions used internally by R package mixmeta. A preliminary attempt shows that using the C++ function and other optimization functions reduces the computational time by 30%. For example, for a problem with 500 inner level units and ten outer level units, the computational time for each iteration became 48 seconds instead of 75 seconds.

The estimation procedure implemented in the mixmeta package uses a hybrid approach that considers the estimates of the random terms obtained after the first ten iterations of the IGLS and RIGLS algorithm proposed by Goldstein (Goldstein, 1986, 1989) under the assumption of unconstrained covariance among the random terms. These estimated represent the starting point of the Newton-Raphson minimisation procedures implemented in the optim package. The implementation in the mixmeta package of the (R)IGLS algorithm can be made faster using the procedure presented in (Goldstein and Rasbash, 1992). I made preliminary attempts at implementing such a procedure with encouraging results.

The computational time increased linearly with the number of inner level units, e.g., the problem with 500 inner level units and ten outer level units requires 18.3 seconds for each iteration to run, respect the 75 seconds of the current version. These preliminary results are encouraging and suggest further research for implementing the complex meta-analytic framework characterised by big data settings.

Uncertainity on the estimates of the variance-components

As described previously, I implemented the extended framework for complex meta-analysis using a frequentist inferential procedure. This procedure is based on the ML and REML estimation of the between-study variance components. More specifically, numerical iterative methods (Newton-Raphson) are used to estimate the variance component. These are plugged into the generalised least square formula for point estimates and standard error for the fixed effects coefficients. In the current version, no correction for the uncertainty of the estimates of the variance components is implemented.
I started to evaluate different approaches that consider the uncertainty of the variance-covariance component estimates in the inferential procedures for the fixed-effects terms. In particular, I considered three different possible methods. The first is based on likelihood-based confidence intervals as proposed by Hardy and Thompson (Hardy and Thompson, 1996). This approach requires the specification of the full likelihood, and it would allow for obtaining confidence intervals for the fixed effects coefficients and the between studies variance components. A different approach is based on the Hartung-Knapp method (Hartung, 1999; Hartung and Knapp, 2001a,b; Sidik and Jonkman, 2002). Interestingly, recently a paper shows the equivalence of the Hartung-Knapp method with the weighted least square regression (van Aert and Jackson, 2019), and this equivalence can be used to generalise the correction in complex meta-analytic setting, e.g. deriving the weights from the inverse of the marginal covariance matrix. Lastly, within a finite-sample inference with REML estimation context, Kenward and Roger proposed an approximate correction of the covariance matrix and degree of freedom (Kenward and Roger, 1997). This correction has been recently applied to standard univariate meta-analysis (Morris *et al.*, 2018), and can be extended to multivariate and other complex meta-analytic regression models. Some of these methods require the analytic formulation of first and second derivatives of the likelihood functions (Gumedze and Dunne, 2011; Stroup, 2012; Wang and Merkle, 2018). I made the first

attempts to develop these three procedures using Matlab considering simple multivariate multilevel settings (e.g., two levels with four outer and five inner level units). Still, these procedures need to be developed in more general settings, and the performance of the different approaches need to be compared with simulation studies. I think this will be an important research area. Papers have compared methods that take into account uncertainty on the estimation of the (co)variances components in the simple univariate settings Sidik and Jonkman (2007); Veroniki *et al.* (2016), but there is no specific contribution that evaluates this issue in complex meta-analysis.

Applications in environmental epidemiology and other research area

A possible research direction is to explore developments of the two-stage design to evaluate the health effects of environmental hazards in multi-centre studies using a "causal" perspective based on Rubin's potential outcomes approach and its implementation using propensity scores. Rubin's potential outcomes approach and its implementation using propensity scores have been widely used in observational studies to remove bias due to the imbalance of confounders between exposure groups. It was firstly proposed for a binary treatment (exposure) and then extended to consider continuous treatment (Hirano and Imbens, 2004; Imai and Van Dyk, 2004; Zhao *et al.*, 2020). It has also been extended to consider multiple treatments (Egger and Von Ehrlich, 2013; Williams and Crespi, 2020). Several approaches have been proposed and can be broadly classified into approaches based on matching, stratification, weights, and regression. Recently these approaches have been introduced in environmental epidemiology to investigate the short-term effect of PM2.5 using a matching Baccini *et al.* (2017) and a regression approach Forastiere *et al.* (2020). The potential outcome approach could give some insights into the causal structure of the epidemiological relationship and could more efficiently remove unbalances due to confounders/exposure associations. In this setting, two methodologically challenges can be identified; the first is related to the meta-analytic problem of pooling dose-response "causal" dose-response curves obtained using a potential outcome regression approach in each study area, and the second is related to the possibility of using propensity score weights to evaluate the "causal" effect of contextual meta-regressor in the second-stage meta-analysis.

One of the specifications of the extended random effect framework for meta-analysis presented in

this PhD thesis is the multilevel meta-analysis that allows to characterise a possible hierarchical geographical structure of estimates to be pooled. This has been applied in one of the case-example of the paper presenting the extended two-stage design in which the association between ozone and mortality was assessed in US cities nested in the US States. The multilevel parametrisation implies a common correlation structure between the lower level units (e.g., cities) nested within the highest group-level (US state). This assumption could be mitigated by imposing a spatial correlation structure on which the correlation is a function of some distance metrics across the lower-level units (cities). The possibility of defining the spatial autocorrelation covariance matrix would improve the robustness of extended two-stage design in a study setting characterised by a spatial structure of units on which first-stage estimates are estimated. In this PhD I showed applications of the extended two stage-design in environmental epidemiologic studies, but the extended random-effect framework can be applied in other multi-centre observational or clinical studies in which the two-stage design could give computational advantages. In particular, it could be applied to evaluate the effect of modifiers, e.g performing sub-group or dose-response meta-analysis based on estimates obtained through first-stage stratified analysis, e.g. by sex or age.

8.4 Conclusions

During the last years, complex meta-analytic problems arose in several research fields. In this PhD thesis, I developed an extended random effects framework for meta-analysis that provides some tools to improve the analytical approaches in these fields. In addition, the implementation in a freely available statistical software facilitates the application of these methods among applied researchers. I also showed how the extended framework could be applied in environmental epidemiology two stages studies allowing a clearer characterisation of the short-term effects of environmental stressors. Although recently proposed and published, these statistical methods and related software seem to represent a useful and valuable tool for the research community. This is reassuring about the importance of the research I carried out within my PhD project and a strong motivation to develop it further in the future.

Bibliography

van Aert, R. C. and Jackson, D. (2019) A new justification of the hartung-knapp method for random-effects meta-analysis based on weighted least squares regression. *Research synthesis methods*, **10**(4), 515–527.

Agrawal, M., Petralia, F., Tepler, A., Durbin, L., Reinisch, W., Colombel, J.-F. and Shah, S. C. (2022) Gender-based differences in response to tumor necrosis factor inhibitor therapies for ulcerative colitis: Individual participant data meta-analyses of clinical trials. *Inflammatory Bowel Diseases*.

Ahn, J. E. and French, J. L. (2010) Longitudinal aggregate data model-based meta-analysis with nonmem: approaches to handling within treatment arm correlation. *Journal of pharmacokinetics and pharmacodynamics*, **37**(2), 179–201.

Aitkin, M. (1999) A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, **55**(1), 117–128.

Allinson, J. P., Afzal, S., Çolak, Y., Jarvis, D., Backman, H., van den Berge, M., Boezen, H. M., Breyer, M.-K., Breyer-Kohansal, R., Brusselle, G. *et al.* (2022) Changes in lung function in european adults born between 1884 and 1996 and implications for the diagnosis of lung disease: a cross-sectional analysis of ten population-based studies. *The Lancet Respiratory Medicine*, **10**(1), 83–94.

Altman, D. G. (2015) Some reflections on the evolution of meta-analysis. *Research synthesis methods*, **6**(3), 265–267.

Arends, L. R., Hunink, M. and Stijnen, T. (2008) Meta-analysis of summary survival curve data. *Statistics in medicine*, **27**(22), 4381–4396.

Armstrong, B. G., Gasparrini, A., Tobias, A. and Sera, F. (2020) Sample size issues in time series regressions of counts on environmental exposures. *BMC medical research methodology*, **20**(1), 1–9.

Baccini, M., Biggeri, A., Accetta, G., Kosatsky, T., Katsouyanni, K., Analitis, A., Anderson, H. R., Bisanti, L., D'Ippoliti, D., Danova, J. *et al.* (2008) Heat effects on mortality in 15 european cities. *Epidemiology*, pp. 711–719.

Baccini, M., Mattei, A., Mealli, F., Bertazzi, P. A. and Carugno, M. (2017) Assessing the short term impact of air pollution on mortality: a matching approach. *Environmental Health*, **16**(1), 1–12.

Bagos, P. G. (2015) Meta-analysis in stata using gllamm. *Research synthesis methods*, **6**(4), 310–332.

Bagos, P. G. and Liakopoulos, T. D. (2010) A multipoint method for meta-analysis of genetic association studies. *Genetic epidemiology*, **34**(7), 702–715.

Bailar III, J. C. (1995) The practice of meta-analysis. *Journal of clinical epidemiology*, **48**(1), 149–157.

Bär, S., Bundo, M., de Schrijver, E., Müller, T. J. and Vicedo-Cabrera, A. M. (2022) Suicides and ambient temperature in switzerland: A nationwide time-series analysis. *Swiss medical weekly*, (9).

Barth, M., Güllich, A., Macnamara, B. N. and Hambrick, D. Z. (2022) Predictors of junior versus senior elite performance are opposite: A systematic review and meta-analysis of participation patterns. *Sports Medicine*, pp. 1–18.

Basagana, X., Pedersen, M., Barrera-Gomez, J., Gehring, U., Giorgis-Allemand, L., Hoek, G., Stafoggia, M., Nieuwenhuijsen, M. J., Brunekreef, B., Slama, R. *et al.* (2018) Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. *International Journal of Epidemiology*, **47**(4), 1343–1354.

Beck, M. W., de Valpine, P., Murphy, R., Wren, I., Chelsky, A., Foley, M. and Senn, D. B. (2022) Multi-scale trend analysis of water quality using error propagation of generalized additive models. *Science of the Total Environment*, **802**, 149927.

Berhane, K. and Thomas, D. C. (2002) A two-stage model for multiple time series data of counts. *Biostatistics*, **3**(1), 21–32.

Berkey, C., Hoaglin, D., Antczak-Bouckoms, A., Mosteller, F. and Colditz, G. (1998) Meta-analysis of multiple outcomes by regression with random effects. *Statistics in medicine*, **17**(22), 2537–2550.

Berkey, C. S., Hoaglin, D. C., Mosteller, F. and Colditz, G. A. (1995) A random-effects regression model for meta-analysis. *Statistics in medicine*, **14**(4), 395–411.

Berlin, J. A., Longnecker, M. P. and Greenland, S. (1993) Meta-analysis of epidemiologic dose-response data. *Epidemiology*, pp. 218–228.

Bobb, J. F., Peng, R. D., Bell, M. L. and Dominici, F. (2014) Heat-related mortality and adaptation to heat in the united states. *Environmental health perspectives*, **122**(8), 811–816.

Borbón, T. Y., Qu, P., Coleman-Satterfield, T. T., Kearney, R. and Klein, E. J. (2022) Digital nerve blocks: A systematic review and meta-analysis. *Journal of the American College of Emergency Physicians Open*, **3**(4), e12753.

Borenstein, M., Hedges, L. V., Higgins, J. P. and Rothstein, H. R. (2021) *Introduction to meta-analysis*. John Wiley & Sons.

Brockwell, S. E. and Gordon, I. R. (2001) A comparison of statistical methods for meta-analysis. *Statistics in medicine*, **20**(6), 825–840.

Bujkiewicz, S., Thompson, J. R., Sutton, A. J., Cooper, N. J., Harrison, M. J., Symmons, D. P. and Abrams, K. R. (2013) Multivariate meta-analysis of mixed outcomes: a bayesian approach. *Statistics in Medicine*, **32**(22), 3926–3943.

Burke, D. L., Ensor, J. and Riley, R. D. (2017) Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Statistics in medicine*, **36**(5), 855–875.

Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D. and Geys, H. (2000) The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*, **1**(1), 49–67.

Chen, K., Breitner, S., Wolf, K., Stafoggia, M., Sera, F., Vicedo-Cabrera, A., Guo, Y., Tong, S., Lavigne, E., Matus, P., Valdés, N., Kan, H., Jaakkola, J., Ryti, N., Huber, V., Scortichini, M., Hashizume, M., Honda, Y., Nunes, B., Madureira, J., Holobâcă, I., Fratianni, S., Kim, H., Lee, W., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Ragettli, M., Guo, Y.-L., Chen, B.-Y., Li, S., Milojevic, A., Zanobetti, A., Schwartz, J., Bell, M., Gasparrini, A. and Schneider, A. (2021) Ambient carbon monoxide and daily mortality: a global time-series study in 337 cities. *The Lancet Planetary Health*, **5**(4), e191–e199. cited By 6.

Chen, R., Kan, H., Chen, B., Huang, W., Bai, Z., Song, G. and Pan, G. (2012) Association of particulate air pollution with daily mortality: the china air pollution and health effects study. *American journal of epidemiology*, **175**(11), 1173–1181.

Cheung, M. W.-L. (2014a) Fixed-and random-effects meta-analytic structural equation modeling: Examples and analyses in r. *Behavior Research Methods*, **46**(1), 29–40.

Cheung, M. W.-L. (2014b) Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological Methods*, **19**(2), 211.

Cheung, M. W.-L. (2015) metasem: An r package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, **5**, 1521.

Cheung, M. W.-L. and Jak, S. (2016) Analyzing big data in psychology: A split/analyze/meta-analyze approach. *Frontiers in psychology*, **7**, 738.

Chu, L., Du, H., Li, T., Lu, F., Guo, M., Dubrow, R. and Chen, K. (2021) Short-term associations between particulate matter air pollution and hospital admissions through the emergency room for urinary system disease in beijing, china: A time-series study. *Environmental Pollution*, **289**, 117858.

Chua, P., Ng, C., Rivera, A., Salva, E., Salazar, M., Huber, V. and Hashizume, M. (2021) Association between ambient temperature and severe diarrhoea in the national capital region, philippines. *International Journal of Environmental Research and Public Health*, **18**(15). cited By 0.

Crippa, A., Discacciati, A., Bottai, M., Spiegelman, D. and Orsini, N. (2019) One-stage dose–response meta-analysis for aggregated data. *Statistical methods in medical research*, **28**(5), 1579–1596.

Crowther, M. J., Look, M. P. and Riley, R. D. (2014) Multilevel mixed effects parametric survival models using adaptive gauss–hermite quadrature with application to recurrent events and individual participant data meta-analysis. *Statistics in medicine*, **33**(22), 3844–3858.

DerSimonian, R. and Laird, N. (1986) Meta-analysis in clinical trials. *Controlled clinical trials*, **7**(3), 177–188.

Dickerson-Young, T., Uspal, N. G., Prince, W. B., Qu, P. and Klein, E. J. (2022) Racial and ethnic differences in ondansetron use for acute gastroenteritis in children. *Pediatric Emergency Care*, pp. 10–1097.

Dominici, F., Samet, J. M. and Zeger, S. L. (2000) Combining evidence on air pollution and daily mortality from the 20 largest us cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **163**(3), 263–302.

Dominici, F., Zanobetti, A., Zeger, S. L., Schwartz, J. and Samet, J. M. (2004) Hierarchical bivariate time series models: a combined analysis of the effects of particulate matter on morbidity and mortality. *Biostatistics*, **5**(3), 341–360.

Donzelli, G., Biggeri, A., Tobias, A., Nottmeyer, L. N. and Sera, F. (2022) Role of meteorological factors on sars-cov-2 infection incidence in italy and spain before the vaccination campaign. a multi-city time series study. *Environmental research*, p. 113134.

Egger, M. and Smith, G. D. (1995) Misleading meta-analysis. *Bmj*, **311**(7007), 753–754.

Egger, M., Schneider, M. and Smith, G. D. (1998) Spurious precision? meta-analysis of observational studies. *BMJ: British Medical Journal*, **316**(7125), 140.

Egger, P. H. and Von Ehrlich, M. (2013) Generalized propensity scores for multiple continuous treatment variables. *Economics letters*, **119**(1), 32–34.

Eysenck, H. J. (1994) Systematic reviews: Meta-analysis and its problems. *Bmj*, **309**(6957), 789–792.

Filippini, T., Malavolti, M., Whelton, P., Naska, A., Orsini, N. and Vinceti, M. (2021) Blood pressure effects of sodium reduction: Dose-response meta-analysis of experimental studies. *Circulation*, pp. 1542–1567. cited By 24.

Follmann, D. A. and Proschan, M. A. (1999) Valid inference in random effects meta-analysis. *Biometrics*, **55**(3), 732–737.

Fong, F. C. and Smith, D. R. (2022) Exposure-lag response of air temperature on covid-19 incidence in twelve italian cities: A meta-analysis. *Environmental research*, p. 113099.

Forastiere, L., Carugno, M. and Baccini, M. (2020) Assessing short-term impact of pm10 on mortality using a semiparametric generalized propensity score approach. *Environmental Health*, **19**(1), 1–13.

Formica, V., Sera, F., Cremolini, C., Riondino, S., Morelli, C., Arkenau, H.-T. and Roselli, M. (2021) Kras and braf mutations in stage ii and iii colon cancer: A systematic review and meta-analysis. *JNCI: Journal of the National Cancer Institute*.

Gasparrini, A. (2014) Modeling exposure–lag–response associations with distributed lag non-linear models. *Statistics in medicine*, **33**(5), 881–899.

Gasparrini, A., Armstrong, B. and Kenward, M. (2012) Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in medicine*, **31**(29), 3821–3839.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklöv, J., Forsberg, B. *et al.* (2015) Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet*, **386**(9991), 369–375.

Gasparrini, A., Sera, F. and Gasparrini, M. A. (2021) Package 'mixmeta'.

Gasparrini, A., Masselot, P., Scortichini, M., Schneider, R., Mistry, M. N., Sera, F., Macintyre, H. L., Phalkey, R. and Vicedo-Cabrera, A. M. (2022) Small-area assessment of temperature-related mortality risks in england and wales: a case time series analysis. *The Lancet Planetary Health*, **6**(7), e557–e564.

Gerli, A., Centanni, S., Miozzo, M., Virchow, J., Sotgiu, G., Canonica, G. and Soriano, J. (2020) Covid-19 mortality rates in the european union, switzerland, and the uk: Effect of timeliness, lockdown rigidity, and population density. *Minerva Medica*, **111**(4), 308–314. cited By 22.

Glass, G. V. (1976) Primary, secondary, and meta-analysis of research. *Educational researcher*, **5**(10), 3–8.

Goldstein, H. (1986) Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, **73**(1), 43–56.

Goldstein, H. (1989) Restricted unbiased iterative generalized least-squares estimation. *Biometrika*, **76**(3), 622–623.

Goldstein, H. (2011) *Multilevel statistical models*, volume 922. John Wiley & Sons.

Goldstein, H. and Rasbash, J. (1992) Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares. *Computational Statistics & Data Analysis*, **13**(1), 63–71.

Goldstein, H., Yang, M., Omar, R., Turner, R. and Thompson, S. (2000) Meta-analysis using multilevel models with an application to the study of class size effects. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **49**(3), 399–412.

Gumedze, F. and Dunne, T. (2011) Parameter estimation and inference in the linear mixed model. *Linear Algebra and its Applications*, **435**(8), 1920–1944.

Guolo, A. and Varin, C. (2017) Random-effects meta-analysis: the number of studies matters. *Statistical methods in medical research*, **26**(3), 1500–1518.

Hardy, R. J. and Thompson, S. G. (1996) A likelihood approach to meta-analysis with random effects. *Statistics in medicine*, **15**(6), 619–629.

Hartung, J. (1999) An alternative method for meta-analysis. *Biometrical Journal*, **41**(8), 901–916.

Hartung, J. and Knapp, G. (2001a) A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine*, **20**(24), 3875–3889.

Hartung, J. and Knapp, G. (2001b) On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in medicine*, **20**(12), 1771–1782.

Hasegawa, K., Tsukahara, T. and Nomiyama, T. (2022) Short-term associations of ambient air pollution with hospital admissions for ischemic stroke in 97 japanese cities. *Environmental Science and Pollution Research*, pp. 1–11.

He, C., Kim, H., Hashizume, M., Lee, W., Honda, Y., Kim, S. E., Kinney, P. L., Schneider, A., Zhang, Y., Zhu, Y. *et al.* (2022) The effects of night-time warming on mortality burden under future climate change scenarios: a modelling study. *The Lancet Planetary Health*, **6**(8), e648–e657.

Hedges, L. V., Tipton, E. and Johnson, M. C. (2010) Robust variance estimation in meta-regression with dependent effect size estimates. *Research synthesis methods*, **1**(1), 39–65.

Hemming, K., Bowater, R. J. and Lilford, R. J. (2012) Pooling systematic reviews of systematic reviews: a bayesian panoramic meta-analysis. *Statistics in Medicine*, **31**(3), 201–216.

Hemming, K., Pinkney, T., Futaba, K., Pennant, M., Morton, D. G. and Lilford, R. J. (2013) A systematic review of systematic reviews and panoramic meta-analysis: staples versus sutures for surgical procedures. *PloS one*, **8**(10), e75132.

Higgins, J. and Thompson, S. G. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, **21**(11), 1539–1558.

Higgins, J., Whitehead, A., Turner, R. M., Omar, R. Z. and Thompson, S. G. (2001) Meta-analysis of continuous outcome data from individual patients. *Statistics in medicine*, **20**(15), 2219–2241.

Higgins, J., Thompson, S. G. and Spiegelhalter, D. J. (2009) A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **172**(1), 137–159.

Higgins, J. P., Thompson, S. G., Deeks, J. J. and Altman, D. G. (2003) Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, **327**(7414), 557.

Hirano, K. and Imbens, G. W. (2004) The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, **226164**, 73–84.

Hong, F. and Breitling, R. (2008) A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics*, **24**(3), 374–382.

Huber, V., Ortiz, C. P., Puyol, D. G., Lange, S. and Sera, F. (2022) Evidence of rapid adaptation integrated into projections of temperature-related excess mortality. *Environmental Research Letters*, **17**(4), 044075.

Imai, K. and Van Dyk, D. A. (2004) Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, **99**(467), 854–866.

Ishak, K. J., Platt, R. W., Joseph, L., Hanley, J. A. and Caro, J. J. (2007) Meta-analysis of longitudinal studies. *Clinical Trials*, **4**(5), 525–539.

Jackson, D. (2006) The power of the standard test for the presence of heterogeneity in meta-analysis. *Statistics in Medicine*, **25**(15), 2688–2699.

Jackson, D., White, I. R. and Thompson, S. G. (2010) Extending dersimonian and laird's methodology to perform multivariate random effects meta-analyses. *Statistics in medicine*, **29**(12), 1282–1297.

Jackson, D., Riley, R. and White, I. R. (2011) Multivariate meta-analysis: Potential and promise. *Statistics in medicine*, **30**(20), 2481–2498.

Jackson, D., Rollins, K. and Coughlin, P. (2014) A multivariate model for the meta-analysis of study level survival data at multiple times. *Research synthesis methods*, **5**(3), 264–272.

Jackson, D., White, I. R., Price, M., Copas, J. and Riley, R. D. (2017) Borrowing of strength and study weights in multivariate and network meta-analysis. *Statistical methods in medical research*, **26**(6), 2853–2868.

James, J., Page-Dumroese, D., Busse, M., Palik, B., Zhang, J., Eaton, B., Slesak, R., Tirocke, J. and Kwon, H. (2021) Effects of forest harvesting and biomass removal on soil carbon and nitrogen: Two complementary meta-analyses. *Forest Ecology and Management*, **485**. cited By 5.

Jeon, J. P., Chen, C.-H., Tsuang, F.-Y., Liu, J., Hill, M. D., Zhang, L., Yang, P., Wang, G., Cho, B.-H., Kim, J.-T. *et al.* (2021) Impact of renal impairment on short-term outcomes following endovascular thrombectomy for acute ischemic stroke: A systematic review and meta-analysis. *International Journal of Stroke*, p. 17474930211047337.

Joshi, M., Pustejovsky, J. E. and Beretvas, S. N. (2021) Clusterwild bootstrapping to handle dependent effect sizes in meta-analysis with a small number of studies. *Research Synthesis Methods*.

Junqueira, C. N., Pereira, R. A. S., da Silva, R. C., Alves Cardoso Kobal, R. O., Araújo, T. N., Prato, A., Pedrosa, J., Martínez-Martínez, C. A., Castrillon, K. P., Felício, D. T. *et al.* (2022) Do apis and non-apis bees provide a similar contribution to crop production with different levels of pollination dependency? a review using meta-analysis. *Ecological Entomology*, **47**(1), 76–83.

Kalaian, H. A. and Raudenbush, S. W. (1996) A multivariate mixed linear model for meta-analysis. *Psychological methods*, **1**(3), 227.

Kenward, M. G. and Roger, J. H. (1997) Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, pp. 983–997.

Konstantopoulos, S. (2011) Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*, **2**(1), 61–76.

Larose, D. T. and Dey, D. K. (1997) Grouped random effects models for bayesian meta-analysis. *Statistics in Medicine*, **16**(16), 1817–1829.

Lavigne, E., Ryti, N., Gasparrini, A., Sera, F., Weichenthal, S., Chen, H., To, T., Evans, G. J., Sun, L., Dheri, A. *et al.* (2022) Short-term exposure to ambient air pollution and individual emergency department visits for covid-19: a case-crossover study in canada. *Thorax*.

Lintner, T. (2022) A social network perspective on formation of peer relationships in czech lower-secondary classrooms. *Issues in Educational Research*, **32**(1), 182–204.

Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A. M., Guo, Y., Tong, S., Coelho, M. S., Saldiva, P. H., Lavigne, E., Matus, P. *et al.* (2019) Ambient particulate air pollution and daily mortality in 652 cities. *New England Journal of Medicine*, **381**(8), 705–715.

Liu, C., Cai, J., Chen, R., Sera, F., Guo, Y., Tong, S., Li, S., Lavigne, E., Correa, P. M., Ortega, N. V. *et al.* (2022) Coarse particulate air pollution and daily mortality: A global study in 205 cities. *American journal of respiratory and critical care medicine*, (ja).

Liu, Q., Cook, N. R., Bergström, A. and Hsieh, C.-C. (2009) A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose–response data. *Computational Statistics & Data Analysis*, **53**(12), 4157–4167.

Lopes, H. F., Müller, P. and Rosner, G. L. (2003) Bayesian meta-analysis for longitudinal data models using multivariate mixture priors. *Biometrics*, **59**(1), 66–75.

Ma, X., Nie, L., Cole, S. R. and Chu, H. (2016) Statistical methods for multivariate meta-analysis of diagnostic tests: an overview and tutorial. *Statistical methods in medical research*, **25**(4), 1596–1619.

Ma, Y. and Mazumdar, M. (2011) Multivariate meta-analysis: a robust approach based on the theory of u-statistic. *Statistics in medicine*, **30**(24), 2911–2929.

Madaniyazi, L., Chung, Y., Kim, Y., Tobias, A., Ng, C. F. S., Seposo, X., Guo, Y., Honda, Y., Gasparrini, A., Armstrong, B. *et al.* (2021) Seasonality of mortality under a changing climate: a time-series analysis of mortality in japan between 1972 and 2015. *Environmental health and preventive medicine*, **26**(1), 1–9.

Madaniyazi, L., Armstrong, B., Chung, Y., Ng, C. F. S., Seposo, X., Kim, Y., Tobias, A., Guo, Y., Sera, F., Honda, Y. *et al.* (2022) Seasonal variation in mortality and the role of temperature: a multi-country multi-city study. *International journal of epidemiology*, **51**(1), 122–133.

Martínez-Solanas, È., Quijal-Zamorano, M., Achebak, H., Petrova, D., Robine, J.-M., Herrmann, F. R., Rodó, X. and Ballester, J. (2021) Projections of temperature-attributable mortality in europe: a time series analysis of 147 contiguous regions in 16 countries. *The Lancet Planetary Health*, **5**(7), e446–e454.

Masselot, P., Sera, F., Schneider, R., Kan, H., Lavigne, É., Stafoggia, M., Tobias, A., Chen, H., Burnett, R. T., Schwartz, J. *et al.* (2021) Differential mortality risks associated with pm2. 5 components: a multi-country, multi-city study. *Epidemiology*.

Meng, X., Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A. M., Milojevic, A., Guo, Y., Tong, S., Coelho, M. d. S. Z. S., Saldiva, P. H. N. *et al.* (2021) Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities. *bmj*, **372**.

Mistry, M. N., Schneider, R., Masselot, P., Royé, D., Armstrong, B., Kyselỳ, J., Orru, H., Sera, F., Tong, S., Lavigne, É. *et al.* (2022) Comparison of weather station and climate reanalysis data for modelling temperature-related mortality. *Scientific reports*, **12**(1), 1–14.

Morris, T. P., Fisher, D. J., Kenward, M. G. and Carpenter, J. R. (2018) Meta-analysis of gaussian individual patient data: two-stage or not two-stage? *Statistics in medicine*.

Musekiwa, A., Manda, S. O., Mwambi, H. G. and Chen, D.-G. (2016) Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *PloS one*, **11**(10), e0164898.

Nam, I.-S., Mengersen, K. and Garthwaite, P. (2003) Multivariate meta-analysis. *Statistics in Medicine*, **22**(14), 2309–2333.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F. and Sánchez-Meca, J. (2013) Three-level meta-analysis of dependent effect sizes. *Behavior research methods*, **45**(2), 576–594.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F. and Sánchez-Meca, J. (2015) Meta-analysis of multiple outcomes: a multilevel approach. *Behavior research methods*, **47**(4), 1274–1294.

Nottmeyer, L. and Sera, F. (2021) Influence of temperature, and of relative and absolute humidity on covid-19 incidence in england - a multi-city time-series study. *Environmental Research*, **196**. cited By 9.

Olkin, I. and Gleser, L. (2009) Stochastically dependent effect sizes. *The handbook of research synthesis and meta-analysis*, pp. 357–376.

Onozuka, D., Tanoue, Y., Nomura, S., Kawashima, T., Yoneoka, D., Eguchi, A., Ng, C. F. S., Matsuura, K., Shi, S., Makiyama, K. *et al.* (2022) Reduced mortality during the covid-19 outbreak in japan, 2020: a two-stage interrupted time-series design. *International journal of epidemiology*, **51**(1), 75–84.

Orsini, N. (2021) Weighted mixed-effects dose–response models for tables of correlated contrasts. *Stata Journal*, **21**(2), 320–347. cited By 5.

Orsini, N., Li, R., Wolk, A., Khudyakov, P. and Spiegelman, D. (2011) Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *American journal of epidemiology*, **175**(1), 66–73.

Peters, J. L. and Mengersen, K. L. (2008) Meta-analysis of repeated measures study designs. *Journal of evaluation in clinical practice*, **14**(5), 941–950.

Peto, R., Pike, M., Armitage, P., Breslow, N. E., Cox, D., Howard, S., Mantel, N., McPherson, K., Peto, J. and Smith, P. (1977) Design and analysis of randomized clinical trials requiring prolonged observation of each patient. ii. analysis and examples. *British journal of cancer*, **35**(1), 1.

Pinheiro, J. and Bates, D. (2006) *Mixed-effects models in S and S-PLUS*. Springer science & business media.

Polanin, J. R., Hennessy, E. A. and Tanner-Smith, E. E. (2017) A review of meta-analysis packages in r. *Journal of Educational and Behavioral Statistics*, **42**(2), 206–242.

Prevost, T. C., Abrams, K. R. and Jones, D. R. (2000) Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. *Statistics in medicine*, **19**(24), 3359–3376.

Rabe-Hesketh, S., Pickles, A. and Skrondal, A. (2001) Gllamm: A general class of multilevel models and a stata program. *Multilevel modelling newsletter*, **13**(1), 17–23.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I. and Lewis, T. (2000) A user's guide to mlwin. *London: Institute of Education*, **286**.

Raudenbush, S. W., Becker, B. J. and Kalaian, H. (1988) Modeling multivariate effect sizes. *Psychological Bulletin*, **103**(1), 111.

Riley, R. D. (2009) Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **172**(4), 789–811.

Riley, R. D., Abrams, K., Lambert, P., Sutton, A. and Thompson, J. (2007a) An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Statistics in medicine*, **26**(1), 78–97.

Riley, R. D., Thompson, J. R. and Abrams, K. R. (2007b) An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics*, **9**(1), 172–186.

Riley, R. D., Lambert, P. C. and Abo-Zaid, G. (2010) Meta-analysis of individual participant data: rationale, conduct, and reporting. *Bmj*, **340**, c221.

Riley, R. D., Higgins, J. P. and Deeks, J. J. (2011) Interpretation of random effects meta-analyses. *Bmj*, **342**, d549.

Riley, R. D., Jackson, D., Salanti, G., Burke, D. L., Price, M., Kirkham, J. and White, I. R. (2017) Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *bmj*, **358**, j3932.

Ritz, J., Demidenko, E. and Spiegelman, D. (2008) Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference*, **138**(7), 1919–1933.

Romieu, I., Gouveia, N., Cifuentes, L. A., de Leon, A. P., Junger, W., Vera, J., Strappa, V., Hurtado-Díaz, M., Miranda-Soberanis, V., Rojas-Bracho, L. *et al.* (2012) Multicity study of air pollution and mortality in latin america (the escala study). *Res Rep Health Eff Inst*, **171**, 5–86.

Rota, M., Bellocco, R., Scotti, L., Tramacere, I., Jenab, M., Corrao, G., La Vecchia, C., Boffetta, P. and Bagnardi, V. (2010) Random-effects meta-regression models for studying nonlinear dose–response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statistics in medicine*, **29**(26), 2679–2687.

Royé, D., Sera, F., Tobías, A., Lowe, R., Gasparrini, A., Pascal, M., De'Donato, F., Nunes, B. and Teixeira, J. (2021) Effects of hot nights on mortality in southern europe. *Epidemiology*, pp. 487–498. cited By 2.

Rue, H. and Held, L. (2005) *Gaussian Markov random fields: theory and applications*. CRC press.

Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H. R., Sunyer, J., Bisanti, L., Zmirou, D., Vonk, J. M., Pekkanen, J. *et al.* (2005) Estimating the exposure–response relationships between particulate matter and mortality within the aphea multicity project. *Environmental health perspectives*, **113**(1), 88–95.

de Schrijver, E., Bundo, M., Ragettli, M. S., Sera, F., Gasparrini, A., Franco, O. H. and Vicedo-Cabrera, A. M. (2022) Nationwide analysis of the heat-and cold-related mortality trends in switzerland between 1969 and 2017: The role of population aging. *Environmental health perspectives*, **130**(3), 037001.

Schwartz, J. (2000) Assessing confounding, effect modification, and thresholds in the association between ambient particles and daily deaths. *Environmental health perspectives*, **108**(6), 563–568.

Schwarzer, G., Carpenter, J. R. and Rücker, G. (2015) *Meta-analysis with R*. Springer.

Scortichini, M., Schneider Dos Santos, R., De' Donato, F., De Sario, M., Michelozzi, P., Davoli, M., Masselot, P., Sera, F. and Gasparrini, A. (2020) Excess mortality during the covid-19 outbreak in italy: A two-stage interrupted time-series analysis. *International Journal of Epidemiology*, **49**(6), 1909–1917. cited By 37.

Sera, F. and Gasparrini, A. (2022) Extended two-stage designs for environmental research. *Environmental health*, **21**(1), 1–13.

Sera, F., Armstrong, B., Blangiardo, M. and Gasparrini, A. (2019a) An extended mixed-effects framework for meta-analysis. *Statistics in medicine*, **38**(29), 5429–5444.

Sera, F., Armstrong, B., Tobias, A., Vicedo-Cabrera, A. M., Åström, C., Bell, M. L., Chen, B.-Y., de Sousa Zanotti Stagliorio Coelho, M., Matus Correa, P., Cruz, J. C. *et al.* (2019b) How urban characteristics affect vulnerability to heat and cold: a multi-country analysis. *International journal of epidemiology*, **48**(4), 1101–1112.

Sera, F., Hashizume, M., Honda, Y., Lavigne, E., Schwartz, J., Zanobetti, A., Tobias, A., Iñiguez, C., Vicedo-Cabrera, A. M., Blangiardo, M. *et al.* (2020) Air conditioning and heat-related mortality: a multi-country longitudinal study. *Epidemiology*, **31**(6), 779–787.

Sera, F., Armstrong, B., Abbott, S., Meakin, S., O'Reilly, K., von Borries, R., Schneider, R., Royé, D., Hashizume, M., Pascal, M. *et al.* (2021) A cross-sectional analysis of meteorological factors and sars-cov-2 transmission in 409 cities across 26 countries. *Nature communications*, **12**(1), 1–11.

Shadish, W. R. and Lecy, J. D. (2015) The meta-analytic big bang. *Research synthesis methods*, **6**(3), 246–264.

Shapiro, S. (1997) Is meta-analysis a valid approach to the evaluation of small effects in observational studies? *Journal of clinical epidemiology*, **50**(3), 223–229.

Sidik, K. and Jonkman, J. N. (2002) A simple confidence interval for meta-analysis. *Statistics in Medicine*, **21**(21), 3153–3159.

Sidik, K. and Jonkman, J. N. (2005) Simple heterogeneity variance estimation for meta-analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**(2), 367–384.

Sidik, K. and Jonkman, J. N. (2007) A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in medicine*, **26**(9), 1964–1981.

Sim, K., Kim, Y., Hashizume, M., Gasparrini, A., Armstrong, B., Sera, F., Ng, C., Honda, Y. and Chung, Y. (2020) Nonlinear temperature-suicide association in japan from 1972 to 2015: Its heterogeneity and the role of climate, demographic, and socioeconomic factors. *Environment International*, **142**. cited By 8.

Singer, J. D. (1998) Using sas proc mixed to fit multilevel models, hierarchical models, and individual growth models. *Journal of educational and behavioral statistics*, **23**(4), 323–355.

Srinivasjois, R. (2021) Fixed and random-effects models for meta-analysis. In *Principles and Practice of Systematic Reviews and Meta-Analysis*, pp. 73–78. Springer.

Stevens, J. R. and Taylor, A. M. (2009) Hierarchical dependence in meta-analysis. *Journal of Educational and Behavioral Statistics*, **34**(1), 46–73.

Stram, D. O. (1996) Meta-analysis of published data using a linear mixed-effects model. *Biometrics*, pp. 536–544.

Stroup, W. W. (2012) *Generalized linear mixed models: modern concepts, methods and applications*. CRC press.

Sutton, A. J. and Abrams, K. R. (2001) Bayesian methods in meta-analysis and evidence synthesis. *Statistical methods in medical research*, **10**(4), 277–303.

Sutton, A. J. and Higgins, J. (2008) Recent developments in meta-analysis. *Statistics in medicine*, **27**(5), 625–650.

Tarai, S., Kumar, R. and Bit, A. (2022) Neurocognitive signatures of prosocial and positive emotional behaviours: emerging research and social impact. *Neurocognitive Perspectives of Prosocial and Positive Emotional Behaviours*.

Thompson, S. G. and Higgins, J. (2002) How should meta-regression analyses be undertaken and interpreted? *Statistics in medicine*, **21**(11), 1559–1573.

Thompson, S. G., Turner, R. M. and Warn, D. E. (2001) Multilevel models for meta-analysis, and their application to absolute risk differences. *Statistical methods in medical research*, **10**(6), 375–392.

Tipton, E. (2015) Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, **20**(3), 375.

Tobías, A., Hashizume, M., Honda, Y., Sera, F., Ng, C. F. S., Kim, Y., Roye, D., Chung, Y., Dang, T. N., Kim, H. *et al.* (2021) Geographical variations of the minimum mortality temperature at a global scale: a multicountry study. *Environmental epidemiology*, **5**(5).

Trikalinos, T. A. and Olkin, I. (2012) Meta-analysis of effect sizes reported at multiple time points: a multivariate approach. *Clinical Trials*, **9**(5), 610–620.

Turner, R. M., Omar, R. Z., Yang, M., Goldstein, H. and Thompson, S. G. (2000) A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statistics in medicine*, **19**(24), 3417–3432.

Tuttle, L. J. and Donahue, M. J. (2022) Effects of sediment exposure on corals: a systematic review of experimental studies. *Environmental evidence*, **11**(1), 1–33.

Urban, A., Di Napoli, C., Cloke, H. L., Kyselỳ, J., Pappenberger, F., Sera, F., Schneider, R., Vicedo-Cabrera, A. M., Acquaotta, F., Ragettli, M. S. *et al.* (2021) Evaluation of the era5 reanalysis-based universal thermal climate index on mortality data in europe. *Environmental research*, **198**, 111227.

Van Houwelingen, H. C., Arends, L. R. and Stijnen, T. (2002) Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in medicine*, **21**(4), 589–624.

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J., Langan, D. and Salanti, G. (2016) Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research synthesis methods*, **7**(1), 55–79.

Vicedo-Cabrera, A., Sera, F., Liu, C., Armstrong, B., Milojevic, A., Guo, Y., Tong, S., Lavigne, E., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Pascal, M., Huber, V., Schneider, A., Katsouyanni, K., Samoli, E., Stafoggia, M., Scortichini, M., Hashizume, M., Honda, Y., Ng, C., Hurtado-Diaz, M., Cruz, J., Silva, S., Madureira, J., Scovronick, N., Garland, R., Kim, H., Tobias, A., Íñiguez, C., Forsberg, B., Åström, C., Ragettli, M., Röösli, M., Guo, Y.-L., Chen, B.-Y., Zanobetti, A., Schwartz, J., Bell, M., Kan, H. and Gasparrini, A. (2020) Short term association between ozone and mortality: global two stage time series study in 406 locations in 20 countries. *The BMJ*, **368**. cited By 44.

Vicedo-Cabrera, A., Scovronick, N., Sera, F., Royé, D., Schneider, R., Tobias, A., Astrom, C., Guo, Y., Honda, Y., Hondula, D., Abrutzky, R., Tong, S., Coelho, M., Saldiva, P., Lavigne, E., Correa, P., Ortega, N., Kan, H., Osorio, S., Kyselý, J., Urban, A., Orru, H., Indermitte, E., Jaakkola, J., Ryti, N., Pascal, M., Schneider, A., Katsouyanni, K., Samoli, E., Mayvaneh, F., Entezari, A., Goodman, P., Zeka, A., Michelozzi, P., de'Donato, F., Hashizume, M., Alahmad, B., Diaz, M., Valencia, C., Overcenco, A., Houthuijs, D., Ameling, C., Rao, S., Di Ruscio, F., Carrasco-Escobar, G., Seposo, X., Silva, S., Madureira, J., Holobaca, I., Fratianni, S., Acquaotta, F., Kim, H., Lee, W., Iniguez, C., Forsberg, B., Ragettli, M., Guo, Y., Chen, B., Li, S., Armstrong, B., Aleman, A., Zanobetti, A., Schwartz, J., Dang, T., Dung, D., Gillett, N., Haines, A., Mengel, M., Huber, V. and Gasparrini, A. (2021) The burden of heat-related mortality attributable to recent human-induced climate change. *Nature Climate Change*, **11**(6), 492–500. cited By 34.

Viechtbauer, W. (2007) Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in medicine*, **26**(1), 37–52.

Walton, C., Rice, S., Gao, C., Butterworth, M., Clements, M. and Purcell, R. (2021) Gender differences in mental health symptoms and risk factors in australian elite athletes. *BMJ Open Sport and Exercise Medicine*, **7**(1). cited By 10.

Wang, T. and Merkle, E. C. (2018) merderiv: Derivative computations for linear mixed effects models with application to robust standard errors. *Journal of Statistical Software*, **87**, 1–16.

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Beagley, J., Belesova, K., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D. *et al.* (2021) The 2020 report of the lancet countdown on health and climate change: responding to converging crises. *The Lancet*, **397**(10269), 129–170.

Wei, Y. and Higgins, J. (2013a) Bayesian multivariate meta-analysis with multiple outcomes. *Statistics in Medicine*, **32**(17), 2911–2934.

Wei, Y. and Higgins, J. (2013b) Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine*, **32**(7), 1191–1205.

Wen, B., Wu, Y., Xu, R., Guo, Y. and Li, S. (2022a) Excess emergency department visits for cardiovascular and respiratory diseases during the 2019–20 bushfire period in australia: A two-stage interrupted time-series analysis. *Science of The Total Environment*, **809**, 152226.

Wen, B., Wu, Y., Ye, T., Xu, R., Yu, W., Yu, P., Guo, Y. and Li, S. (2022b) Short-term exposure to ozone and economic burden of premature mortality in italy: A nationwide observation study. *Ecotoxicology and Environmental Safety*, **241**, 113781.

White, I. R. *et al.* (2011) Multivariate random-effects meta-regression: updates to mvmeta. *Stata journal*, **11**(2), 255.

White, I. R., Barrett, J. K., Jackson, D. and Higgins, J. (2012) Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research synthesis methods*, **3**(2), 111–125.

Whitehead, A., Omar, R. Z., Higgins, J., Savaluny, E., Turner, R. M. and Thompson, S. G. (2001) Meta-analysis of ordinal outcomes using individual patient data. *Statistics in Medicine*, **20**(15), 2243–2260.

Wieland, L., Hamel, C., Konstantinidis, M., Nourouzpour, S., Shipper, A. and Lipski, E. (2021) Zinc for prevention and treatment of the common cold. *Cochrane Database of Systematic Reviews*, **2021**(9). cited By 0.

Williams, J. R. and Crespi, C. M. (2020) Causal inference for multiple continuous exposures via the multivariate generalized propensity score. *arXiv preprint arXiv:2008.13767*.

Wohner, P. J., Duarte, A., Wikert, J., Cavallo, B., Zeug, S. C. and Peterson, J. T. (2022) Integrating monitoring and optimization modeling to inform flow decisions for chinook salmon smolts. *Ecological Modelling*, **471**, 110058.

Wong, C.-M., Vichit-Vadakan, N., Kan, H. and Qian, Z. (2008) Public health and air pollution in asia (papa): a multicity study of short-term effects of air pollution on mortality. *Environmental health perspectives*, **116**(9), 1195–1202.

Wu, Y., Langworthy, B. and Wang, M. (2022) Marginal structural models for multilevel clustered data. *Biostatistics*.

Yu, J., Park, J., Choi, T., Hashizume, M., Kim, Y., Honda, Y. and Chung, Y. (2021) Nonparametric bayesian functional meta-regression: Applications in environmental epidemiology. *Journal of Agricultural, Biological, and Environmental Statistics*, **26**(1), 45–70. cited By 0.

Zhao, S., van Dyk, D. A. and Imai, K. (2020) Propensity score-based methods for causal inference in observational studies with non-binary treatments. *Statistical methods in medical research*, **29**(3), 709–727.

Appendix A

R package mixmeta

Title: Package 'mixmeta'.
Author(s): Antonio Gasparrini, Francesco Sera.
Journal/Publisher: R Foundation for Statistical Computing.
Type of publication: Package documentation.
Stage of publication: Published online on 10 October 2021.
URL: `http://cran.r-project.org/web/packages/dlnm/index.html`.
Academic peer-reviewed: No.

Candidate's role: See Section 2.3.

Senior author: (Prof. Antonio Gasparrini)

Package 'mixmeta'

October 16, 2021

Type Package

Version 1.2.0

Date 2021-10-10

Title An Extended Mixed-Effects Framework for Meta-Analysis

Description A collection of functions to perform various meta-analytical models through a unified mixed-effects framework, including standard univariate fixed and random-effects meta-analysis and meta-regression, and non-standard extensions such as multivariate, multilevel, longitudinal, and dose-response models.

Author Antonio Gasparrini [aut, cre],
Francesco Sera [aut]

Maintainer Antonio Gasparrini <antonio.gasparrini@lshtm.ac.uk>

Imports stats, graphics, grDevices, utils

Depends R (>= 3.5.0)

Suggests metafor, meta, rmeta, dosresmeta, nlme, MASS, dlnm

URL https://github.com/gasparrini/mixmeta,
http://www.ag-myresearch.com/package-mixmeta

License GPL (>= 2)

LazyData yes

NeedsCompilation no

Repository CRAN

Date/Publication 2021-10-16 14:50:02 UTC

R topics documented:

`mixmeta-package` *An Extended Mixed-Effects Framework for Meta-Analysis*

Description

The package **mixmeta** consists of a collection of functions to perform various meta-analytical models in R through a unified mixed-effects framework, including standard univariate fixed and random-effects meta-analysis and meta-regression, and non-standard extensions such as multivariate, multilevel, longitudinal, and dose-response models.

Modelling framework

Standard applications of meta-analysis amount to the pooling of estimates of a single effect size, here defined generally as outcome, collected as unique observations in a set of independent studies, together with a measure of uncertainty (usually standard errors). Fixed-effects models do not assume heterogeneity across studies, and the estimates are conditional on the set of studies collected in the meta-analysis. Random-effects meta-analysis, instead, allows a degree of heterogeneity among studies, assuming the (true but unobserved) study-specific outcomes as randomly sampled from a (usually hypothetical) population of studies. Meta-regression extends both fixed and random-effects methods by allowing the pooled outcome to depend on study-level meta-predictors.

However, this traditional setting can be limited for many modern applications of meta-analysis. For instance, studies can provide estimates of different outcomes. Alternatively, studies can collect multiple estimates of the same outcome, either longitudinally or referring to different groups or levels of a continuous variable. Similarly, studies can be clustered, or being characterized by a hierarchical structure (i.e., by country). In all these instances, the key assumption of independence across estimates is not met, and basic models must be extended to consider potentially complex correlation structures within and between studies. This leads to extension to multivariate, multilevel, longitudinal, or dose-response models for meta-analysis, among others.

A unified modelling framework can be defined by casting the meta-analytical problem as a linear mixed model. In general terms, we assume that there is a set of n observations of k different outcomes, representing *units* of analysis aggregated in $i = 1, \ldots, m$ *groups* that are considered independent. An extended random-effects meta-regression model for the $\boldsymbol{y}_i$ outcomes in group i can be generally written as:

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b} + \boldsymbol{\epsilon}_i$$
$$\boldsymbol{b} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Psi}), \boldsymbol{\epsilon}_i \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{S}_i)$$

Here, $\boldsymbol{X}_i\boldsymbol{\beta}$ defines the fixed effects that represent the population-averaged outcomes in terms of p unit-level predictors in the design matrix $\boldsymbol{X}_i$ with fixed-effects coefficients $\boldsymbol{\beta}$. The random part of the model $\boldsymbol{Z}_i\boldsymbol{b}$ describes the deviation from the population averages in terms of q unit-level predictors in the design matrix $\boldsymbol{Z}_i$ and random-effects coefficients $\boldsymbol{b}$. The marginal (co)variance matrix $\boldsymbol{\Sigma}_i = \boldsymbol{Z}\boldsymbol{\Psi}\boldsymbol{Z}^t + \boldsymbol{S}_i$ is given by the sum of within (assumed known) and between-group contributions, defined by (co)variance matrices $\boldsymbol{S}_i$ and $\boldsymbol{\Psi}$, respectively.

All the models mentioned above, and other extensions, can be described as special cases of this unified framework. Specifically, in the standard random-effects univariate meta-analysis or meta-regression, each group represents a study with a single observation ($n = m$), where $\boldsymbol{Z_i} = \boldsymbol{1}$ ($q = 1$), and $\boldsymbol{y}_i$, $\boldsymbol{S}_i$ and $\boldsymbol{\Psi}$ are scalars. In fixed-effects models, $\boldsymbol{\Psi}$ and $\boldsymbol{Z_i}$ do not exist. In multivariate models, the k-dimensional $\boldsymbol{y}_i$ represents the different outcomes from study i, $\boldsymbol{X}_i$ is Kronecker-expanded to $k \times kp$, and $\boldsymbol{S}_i$ and $\boldsymbol{\Psi}$ are $k \times k$ matrices representing within and between-study correlations among outcomes, respectively. In multilevel models, where additional inner levels of grouping exist within each of the m outer-level groups, q is the sum of level-specific meta-predictors, while $\boldsymbol{\Psi}$ and $\boldsymbol{Z}_i$ have a block-diagonal and column-binded (and expanded) forms, respectively, with each part referring to a different level. In longitudinal and dose-response models, repeated measures are accomodated in a similar way through random-effects grouping.

Estimation methods

The aim is to estimate the kp coefficients $\boldsymbol{\beta}$ and, for random-effects models, the components of the between-group (co)variance matrix $\boldsymbol{\Psi}$. The parameters for the random part depend on the number

of random-effects levels, and for each of them, on the number of random-effects meta-predictors and the structure of the related part of the (co)variance matrix, with a maximum of $kq(kq+1)/2$ for single-level unstructured $\mathbf{\Psi}$.

Different estimators are implemented in the package **mixmeta**. The options available in the current version are:

- **Fixed-effects**
- **Maximum likelihood** (ML)
- **Restricted maximum likelihood** (REML)
- **Method of moments**
- **Variance components**

The fixed-effects model is fitted through generalized least squares (GLS), assuming the (co)variance structure, composed by the within-study errors only, as completely known. Likelihood-based random-effects estimators, ML and REML, represent the most comprehensive implementation of the modelling framework, and allow the specification all the various models described in the previous section through a flexible definition of the random-effects structure. They rely on two alternative iterative optimization procedures, based on Newton-type and (restricted) iterative generalized least squares (IGLS and RIGLS) algorithms, respectively. Estimators based on semiparametric alternatives such as the non-iterative method of moments or the iterative variance components are also included, although they are only available for models with a basic random-effects structure. Further details on estimation methods are given in the related help pages.

Functions included in the package

The main function in the package is `mixmeta`, which performs the various models illustrated above. This function resembles standard regression functions in R, and specifies the model through regression formulae for fixed and random-effects (see `mixmetaFormula`). The function returns a list object of class `"mixmeta"` (see `mixmetaObject`).

The estimation is carried out internally through `mixmeta.fit`, a wrapper which accepts data in a specific format, then prepares the various data components and calls ad hoc estimation functions for fitting the models. Specifically, `mixmeta.fixed` is applied for fixed-effects models, while estimators for random-effects models are implemented in the functions `mixmeta.ml` and `mixmeta.reml` for (restricted) maximum likelihood, `mixmeta.mm` for the method of moments, and `mixmeta.vc` for variance components. For likelihood-based methods, alternative iterative optimizations methods are provided in two sets of functions implementing `Newton-type` and `(R)IGLS` algorithms used for maximizing the (restricted) likelihood. The former method applies specific `likelihood functions`. Various types of likelihood-based models are defined by separate regression formulae for fixed and random-effects (see `mixmetaFormula`). Specific `(co)variance structures` for the between-group random effects at single or multiple levels are available. Fitting parameter options are set by `mixmeta.control`.

Method functions are available for objects of class `"mixmeta"` (see `mixmetaObject` for a complete list). The method `summary` produces a list of class `"summary.mixmeta"` for summarizing the fit of the model and providing additional results. The method function `predict` computes predicted values, optionally for a set of new values of the predictors. `blup` gives the (empirical) best linear unbiased predictions for the set of studies used for estimation. Other default or specific method

functions for regression can be used on objects of class "mixmeta", such as fitted and residuals, logLik, AIC and BIC, or drop1 and add1, among others.

Methods for model.frame, model.matrix, and terms are used to extract or construct the model frame, the design matrix, or the terms of the regression meta-analytical model, respectively. These specific methods for objects of class "mixmeta" are needed to appropriately deal with missing values and to account for model frames that include terms for both the fixed and random parts. In particular, methods for na.omit and na.exclude are used to handle correctly missing values.

Simulations can be produced using the function mixmetaSim and the method function simulate, which return one or multiple sets of simulated outcomes for a group of studies. The function inputna and inputcov are used internally to augment the missing data values and to input missing correlations, respectively.

The method function qtest.mixmeta (producing an object with class of the same name) performs the (multivariate) Cochran Q test for (residual) heterogeneity. For multivariate models, the function returns both an overall estimate and those for each single outcome. The generic method function is qtest.

Printing functions for the objects of classes defined above are also provided. Other functions are used internally in the source code, and not exported in the namespace. For users interested in getting into details of the package structure, these functions can be displayed using the triple colon (':::') operator. For instance, mixmeta:::glsfit displays the code of the function glsfit. Also, some comments are added in the original source code.

Datasets and applications

The package includes several datasets used for applications of the extended meta-analytical framework. The related help pages provide examples of specific models, and fully demonstrate the flexibility of the extended meta-analytical framework. In particular:

- **Standard meta-analysis** is illustrated using the dataset bcg, including examples of meta-regression.
- **Multivariate meta-analysis** is performed using various datasets, including bivariate models (berkey98, hyp, p53) and multivariate models with three or more outcomes (fibrinogen and hsls). The examples describe also how to deal with missing data or missing within-group correlations, and multivariate meta-regression.
- **Network meta-analysis** is shown in the dataset smoking. The examples illustrate an indirect mixed-treatment comparison including consistency and inconsistency models.
- **Multilevel meta-analysis** is displayed in the examples of the datasets school and thrombolytic, and include data with multiple nested levels of grouping and/or repeated measures within each group.
- **Dose-response meta-analysis** is illustrated in the dataset alcohol, using the recently proposed one-stage approach.
- **Longitudinal meta-analysis** is performed using the datasets dbs and gliomas. The two sets of examples present different cases using data in wide and long format, respectively.

Additional information

The **mixmeta** package is available on the Comprehensive R Archive Network (CRAN), with info at the related web page (CRAN.R-project.org/package=mixmeta). A development website is available

on GitHub (github.com/gasparrini/mixmeta). General information on the development and applications of this extended meta-analytical modelling framework, together with an updated version of the R scripts for running the examples in published papers, can be found in GitHub (github.com/gasparrini) or at the personal web page of the package maintainer (www.ag-myresearch.com).

The package **mixmeta** is an extension of the package **mvmeta**, previously developed to perform multivariate meta-analytical models. The latter now depends on the former, and while both are still maintained, users are encouraged to switch to **mixmeta** as it represents a more general and updated option. A list of changes included in the current and previous versions of **mixmeta** can be found by typing:

```
news(package="mixmeta")
```

Use `citation("mixmeta")` to cite this package.

Warnings

This release of the package **mixmeta** has been tested with different simulated and real datasets. The functions generally perform well under several scenarios, and comparisons with alternative software implementations show good agreement. However, bugs and bad performance under untested conditions may not be excluded. Please report any error or unexpected behaviour to the e-mail address below.

Note

The package **mixmeta** provides a unified modelling framework to perform standard and non-standard meta-analytical models. However, some of these can also be fitted using routines available in other R packages.

For instance, many packages such as **metafor**, **meta**, **rmeta** provide a more exhaustive and efficient set of methods for standard univariate meta-analysis and meta-regression, including a wide range or functions for specific plots and statistical tests.

Specific modelling extensions are also provided by other packages. For example, multivariate or multilevel models can be also be fitted using functions in **metafor** and **metaSEM**, while dose-response meta-analysis and meta-analysis of diagnostic measures can be performed using **dosresmeta** and **mada**, respectively.

See the CRAN Task View Meta-Analysis for a comprehensive illustration of methods available in various R packages.

Author(s)

Antonio Gasparrini and Francesco Sera

Maintainer: Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Pinheiro JC and Bates DM (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer Verlag.

Lindstrom MJ and Bates DM (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*. **83**(404):1014–1022.

Goldstein H (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*. **73**(1):43–56.

Goldstein H (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics \& Data Analysis*. **13**(1):63–71.

Stram DO (1996). Meta-analysis of published data using a linear mixed-effects model. *Biometrics*. **52**(2):536–544.

Stevens JR, Taylor AM. Hierarchical dependence in meta-analysis. *Journal of Educational and Behavioral Statistics*. **34**(1):46–73.

Jackson D, Riley R, White IR (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*. **30**(20);2481–2498.

Goldstein H, et al (2000). Meta-analysis using multilevel models with an application to the study of class size effects. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. **49**(3):399–412.

Crippa A, et al (2019). One-stage dose-response meta-analysis for aggregated data. *Statistical Methods in Medical Research*. **28**(5):1579–1596.

Ishak KJ, Platt RW, Joseph L, et al (2007). Meta-analysis of longitudinal studies. *Clinical Trials*. **4**(5):525–539.

`alcohol` *Alcohol Intake and Colorectal Cancer*

Description

The dataset contains the data on 8 cohort studies participating in the Pooling Project of Prospective Studies of Diet and Cancer. A total of 3,646 cases and 2,511,424 person-years were included in the analysis. Each study estimated the incidence relative rate in different categories of alcohol intake while controlling for a set of potential confounders, using non-drinkers as the reference. The categories where then converted in a dose by assigning to each the median value of individual consumptions, with studies reporting estimates at different levels in a continuous scale.

Usage

```
alcohol
```

Format

A data frame with 48 observations on the following 7 variables:

- `id`: label for each study, derived from the first author's name.
- `type`: code for study design (cohort estimating incidence rate).
- `dose`: assigned dose level (gr/day of alcohol intake).
- `cases`: number of cases for each dose category.
- `peryears`: amount of person-time for each dose category.
- `logrr`: estimated logarithm of the incidence relative rate.
- `se`: standard error of the estimates.

Details

The data are stored in a *long* format, with each record reporting the information for each dose categories and studies including multiple records. The reference category for each study included, although the log-RR is fixed to 0 with no standard error (comparing the category with itself). The information on these reference categories is needed to compute the approximate correlations between estimates in the same study.

Note

The data provide an example of application of dose-response meta-analysis, with repeated measurements of the effect size associated to different doses within each study. This requires a modelling structure that accounts for both within and between-study correlations of repeated measurements. The within-study correlations are usually reconstructed from published data using specific methods. Results can be compared with those reported by Crippa and Orsini (2016) and Orsini and colleagues (2012), although they are not identical: while the original analysis used a two-stage approach, the modelling framework applied here follows the more recent one-stage dose-response meta-analysis proposed by Crippa and collegues (2019).

The dataset is also available in the same format in the dataframe `alcohol_crc` of the package **dosresmeta**.

Source

Crippa A, et al (2019). One-stage dose-response meta-analysis for aggregated data. *Statistical Methods in Medical Research*. **28**(5):1579–1596.

Crippa A, Orsini N (2016). Multivariate dose-response meta-analysis: The dosresmeta R package. *Journal of Statistical Software*. **72**(1):1–15.

Orsini N, et al (2012). Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *American Journal of Epidemiology*. **175**(1):66–73.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN CRIPPA ET AL (2016) AND ORSINI ET AL (2012)

# LOAD THE PACKAGE dosresmeta AND splines
library(dosresmeta) ; library(splines)

# COMPUTE THE WITHIN-STUDY CORRELATIONS EXCLUDING THE REFERENCE
addS <- lapply(split(alcohol, alcohol$id), function(x)
  covar.logrr(y=logrr, v=se^2, cases=cases, n=peryears, type=type, data=x))
sub <- subset(alcohol, !is.na(se))

# NOT ACCOUNTING FOR WITHIN-STUDY CORRELATIONS
nocor <- mixmeta(logrr ~ 0 + dose, S=se^2, random= ~ 0 + dose|id, data=sub,
  method="ml")
summary(nocor)

# ACCOUNTING FOR WITHIN-STUDY CORRELATIONS
lin <- mixmeta(logrr ~ 0 + dose, random= ~ 0 + dose|id, data=sub, method="ml",
  control=list(addSlist=addS))
summary(lin)

# ALLOWING NON-LINEARITY IN BOTH FIXED AND RANDOM PARTS
nonlin <- mixmeta(logrr ~ 0 + ns(dose, knots=c(10,25)), data=sub,
  random= ~ 0 + ns(dose, knots=c(10,25))|id, method="ml",
  control=list(addSlist=addS))
summary(nonlin)

# SIMPLIFY THE MODEL BY ASSUMING LINEARITY IN THE RANDOM PART
nonlin2 <- update(nonlin, random= ~ 0 + dose|id)
summary(nonlin2)

# FIXED-EFFECTS MODEL (TRICK: random TO DEFINE THE GROUPING, THEN FIX IT TO 0)
nonlinfix <- mixmeta(logrr ~ 0 + ns(dose, knots=c(10,25)), random= ~ 1|id,
  data=sub, method="ml",bscov="fixed", control=list(addSlist=addS, Psifix=0))
summary(nonlinfix)

# COMPARE THE MODELS
AIC(nocor, lin, nonlin, nonlin2, nonlinfix)

# PREDICT THE RR FOR 12g/day FOM TWO MODELS
exp(predict(nocor, newdata=data.frame(dose=12), ci=TRUE))
exp(predict(lin, newdata=data.frame(dose=12), ci=TRUE))

# PREDICT (RECREATE SPLINES FOR EASY CODING)
predlin <- exp(predict(lin, newdata=data.frame(dose=0:60), ci=TRUE))
prednonlin <- exp(predict(nonlin, newdata=data.frame(dose=0:60), ci=TRUE))

# DISPLAY THE NON-LINEAR EFFECT
col1 <- do.call(rgb, c(as.list(col2rgb("blue") / 255), list(0.2)))
col2 <- do.call(rgb, c(as.list(col2rgb("green") / 255), list(0.2)))
plot(0:60, predlin[,1], type="l", ylim=c(0.85,1.9), ylab="RR",
  xlab="Alcohol intake (gr/day)", main="Dose-response")
```

```
polygon(c(0:60,60:0), c(predlin[,2], rev(predlin[,3])), col=col1, border=NA)
lines(0:60,prednonlin[,1], lty=5)
polygon(c(0:60,60:0), c(prednonlin[,2],rev(prednonlin[,3])), col=col2, border=NA)
```

---

bcg *Efficacy of BCG Vaccine in the Prevention of Tuberculosis*

Description

The dataset contains the data on 13 prospective clinical trials that compared the rates of tubercolosis in groups vaccinated with the Bacillus Calmette-Guerin (BCG) vaccine and non-vaccinated control populations. The outcome here is reported as both relative risk (RR) and odds ratio (OR), with associated uncertanty.

Usage

```
bcg
```

Format

A data frame with 13 observations on the following 13 variables:

- `trial`: sequence identifying the trial.
- `author`: label identifying the author(s).
- `year`: year of publication.
- `tpos,tneg`: number of positive and negative TB cases in the treated (vaccinated) group.
- `cpos,cneg`: number of positive and negative TB cases in the control (non-vaccinated) group.
- `ablat`: absolute latitude of the study location (in degrees).
- `alloc`: method of treatment allocation (random, alternate, or systematic assignment).

Note

The data provide an example of application of standard univariate meta-analysis and meta-regression, with independent studies providing a single estimate of a single effect size. Interestingly, the data can be analyzed also as a multivariate meta-analysis, using a bivariate outcome where risks or odds of TB can be measured separaterly in treatment and control groups. Results can be compared with those reported van Houwelingen, Arends, and Stijnen (2002).

The dataset is also available in the same format in the dataframe `dat.colditz1994` of the package **metafor**.

Source

van Houwelingen HC, Arends LR, Stijnen T (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine*. **21**(4):589–624.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN VAN HOUWELINGEN ET AL (2002)

# FIXED-EFFECTS META-ANALYSIS (SECTION 3.1.1)
unifix <- mixmeta(logor, logorvar, data=bcg, method="fixed")
print(summary(unifix), digits=3)

# RANDOM-EFFECTS META-ANALYSIS WITH MAXIMUM LIKELIHOOD (SECTION 3.1.2)
uniran <- mixmeta(logor, logorvar, data=bcg, method="ml")
print(summary(uniran), digits=3, report="var")

# ORIGINAL ESTIMATES AND BEST-LINEAR UNBIASED PREDICTIONS (FIGURE 3)
pred <- with(bcg, cbind(logor, logor-1.96*sqrt(logorvar),
  logor+1.96*sqrt(logorvar)))
blup <- blup(uniran, pi=TRUE)
plot(pred[,1], rev(bcg$trial)+0.2, xlim=c(-3,3), ylim=c(0,14), pch=18,
  axes=FALSE, xlab="Log odds ratio", ylab="Trial", main="Forest plot")
axis(1)
axis(2, at=bcg$trial, labels=rev(bcg$trial), lty=0, las=1)
abline(v=coef(uniran))
segments(pred[,2], rev(bcg$trial)+0.2, pred[,3], rev(bcg$trial)+0.2, lty=5)
points(blup[,1], rev(bcg$trial)-0.2, pch=19)
segments(blup[,2], rev(bcg$trial)-0.2, blup[,3], rev(bcg$trial)-0.2)

# COMPUTE THE OUTCOME SEPARATELY FOR TREATMENT AND CONTROL GROUPS
y <- with(bcg, log(cbind(tpos/tneg, cpos/cneg)))
S <- with(bcg, cbind(1/tpos+1/tneg, 1/cpos+1/cneg))

# BIVARIATE RANDOM-EFFECTS META-ANALYSIS (SECTION 4)
mvran <- mixmeta(y, S, method="ml")
print(summary(mvran), digits=3, report="var")

# META-REGRESSION (SECTION 5)
uniranlat <- update(uniran, .~. + ablat)
print(summary(uniranlat), digits=3, report="var")
drop1(uniranlat, test="Chisq")
```

---

bdiagMat *Block-Diagonal Expansion of a List of Matrices*

Description

The function `bdiagMat` builds a single matrix with block-diagonal from a list of matrices.

Usage

```
bdiagMat(x)
```

Arguments

x	a list of matrices, or a single matrix.

Value

A matrix with block-diagonal form if x is a list, or otherwise x itself if a matrix.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See functions bldiag in package **metafor**.

Examples

```
# GENERATE A LIST OF MATRICES, AND CREATE THE BLOCK-DIAGONAL MATRIX
(matlist <- list(matrix(1:4,2), matrix(1:8,2)))
bdiagMat(matlist)
```

---

berkey98 *Five Published Trials on Periodontal Disease*

Description

The dataset contains the results of 5 published trials comparing surgical and non-surgical treatments for medium-severity periodontal disease, one year after treatment. The 2 estimated outcomes are average improvements (surgical minus non-surgical, in mm) in probing depth (PD) and attachment level (AL).

Usage

```
berkey98
```

Format

A data frame with 5 observations on the following 7 variables:

pubyear publication year of the trial.

npat number of patients included in the trial.

PD estimated improvement of surgical versus non-surgical treatments in probing depth (mm).

AL estimated improvement of surgical versus non-surgical treatments in attachment level (mm).

var_PD variance of the estimated outcome for PD.

cov_PD_AL covariance of the estimated outcomes for PD and AL.

var_AL variance of the estimated outcome for AL.

Row names specify the author of the paper reporting the results of each trial.

Source

Berkey CS, et al. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*. **17**:2537–2550.

Berkey CS., et al. (1995). Multiple-outcomes meta-analysis of treatments for periodontal disease. *Journal of Dental Research*. **74**(4):1030–1039.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN BERKEY ET AL. (1998)

# INSPECT THE DATA
berkey98

# FIXED-EFFECTS
year <- berkey98$pubyear - 1983
mod1 <- mixmeta(cbind(PD,AL) ~ year, S=berkey98[5:7], data=berkey98,
  method="fixed")
print(summary(mod1), digits=3)

# GLS MODEL (VARIANCE COMPONENTS)
mod2 <- mixmeta(cbind(PD,AL) ~ year, S=berkey98[5:7], data=berkey98,
  method="vc",  control=list(vc.adj=FALSE))
print(summary(mod2), digits=3)
round(mod2$Psi, 3)

# ML MODEL
mod3 <- mixmeta(cbind(PD,AL) ~ year, S=berkey98[5:7], data=berkey98, method="ml")
print(summary(mod3), digits=3)
round(mod3$Psi, 3)
```

---

blup *Best Linear Unbiased Predictions*

Description

This is a generic function for generating best linear unbiased predictions (BLUPs) from the results of various fitting functions for meta-analytical models. The function invokes particular methods which depend on the `class` of the first argument. Currently, specific methods exist for several meta-analytical models in various packages: `blup.mixmeta`, `blup.rma.uni`, `blup.mvmeta`, and `blup.dosresmeta`.

Usage

```
blup(object, ...)
```

Arguments

`object`	a model object for which BLUPs are desired.
`...`	further arguments passed to or from other methods.

Details

The generic method function `blup` calls specific method functions to produces (empirical) best linear unbiased predictions (BLUPs) from model objects.

These predictions are a shrunk version of unit-specific realizations, where unit-specific estimates borrow strength from the assumption of an underlying (potentially multivariate) distribution in a (usually hypothetical) population. The amount of shrinkage depends from the relative size of the within and between-unit covariance matrices.

Value

The form of the value returned by `blup` depends on the class of its argument. See the documentation of the particular methods for details of what is produced by that method. Usually, the results consist of point estimates of BLUPs and optionally some measure of their uncertainty.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. Springer; 1997.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

Specific methods for various classes: `blup.mixmeta`, `blup.rma.uni`, `blup.mvmeta`, and `blup.dosresmeta`.

`blup.mixmeta` *Best Linear Unbiased Predictions from mixmeta Models*

Description

This method function computes (empirical) best linear unbiased predictions from fitted random-effects meta-analytical models represented in objects of class `"mixmeta"`. Quantities can represent prediction of outcomes given both fixed and random effects, or just random-effects residuals from the fixed-effects estimates. Predictions are optionally accompanied by standard errors, prediction intervals or the entire (co)variance matrix of the predicted outcomes.

Usage

```
## S3 method for class 'mixmeta'
blup(object, se=FALSE, pi=FALSE, vcov=FALSE, pi.level=0.95, type="outcome",
  level, format, aggregate="stat", ...)
```

Arguments

object	an object of class "mixmeta".
se	logical switch indicating if standard errors must be included.
pi	logical switch indicating if prediction intervals must be included.
vcov	logical switch indicating if the (co)variance matrix must be included.
pi.level	a numerical value between 0 and 1, specifying the confidence level for the computation of prediction intervals.
type	the type of prediction. This can be either outcome (default) or residual. See Details.
level	level of random-effects grouping for which predictions are to be computed. Default to the highest (inner) level, with 0 corresponding to fixed-effects predictions obtained through predict.
format	the format for the returned results. See Value.
aggregate	when format="matrix" and se or ci are required, the results may be aggregated by statistic or by outcome. See Value.
...	further arguments passed to or from other methods.

Details

The method function blup produces (empirical) best linear unbiased predictions from mixmeta objects. These can represent outcomes, given by the sum of fixed and random parts, or just random-effects residuals representing deviations from the fixed-effects estimated outcomes. In non-standard models with multiple hierarchies of random effects, the argument level can be used to determine the level of grouping for which predictions are to be computed.

These predictions are a shrunk version of unit-specific realizations, where unit-specific estimates borrow strength from the assumption of an underlying (potentially multivariate) distribution of outcomes or residuals in a (usually hypothetical) population. The amount of shrinkage depends from the relative size of the within and between-unit covariance matrices reported as components S and Psi in mixmeta objects (see mixmetaObject).

Fixed-effects models do not assume random effects, and the results of blup for these models are identical to predict (for type="oucome") or just 0's (for type="residuals").

How to handle predictions for units removed from estimation due to invalid missing pattern is determined by the na.action argument used in mixmeta to produce object. If na.action=na.omit, units excluded from estimation will not appear, whereas if na.action=na.exclude they will appear, with values set to NA for all the outcomes. This step is performed by napredict. See Note below.

In the presence of missing values in the outcomes y of the fitted model, correspondent values of point estimates and covariance terms are set to 0, while the variance terms are set to 1e+10. In this

case, in practice, the unit-specific estimates do not provide any information (their weight is virtually 0), and the prediction tends to the value returned by predict with interval="prediction", when applied to a new but identical set of predictors. See also Note below.

Value

(Empirical) best linear unbiased predictions of outcomes or random-effects residuals. The results may be aggregated in matrices (the default), or returned as lists, depending on the argument format. For multivariate models, the aggregation is ruled by the argument aggregate, and the results may be grouped by statistic or by outcome. If vcov=TRUE, lists are always returned.

Note

The definition of missing in model frames used for estimation in mixmeta is different than that commonly adopted in other regression models such as lm or glm. See info on missing values in mixmeta.

Differently from predict, this method function computes the predicted values in the presence of partially missing outcomes. Interestingly, BLUPs for missing outcomes may be slightly different than predictions returned by predict on a new but identical set of predictors, as the BLUP also depends on the random part of the model. Specifically, the function uses information from the random-effects (co)variance to predict missing outcomes given the observed ones.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. Springer; 1997.

See Also

See predict for standard predictions. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ 1, S=berkey98[5:7], data=berkey98)

# ONLY BLUP
blup(model)

# BLUP AND SE
blup(model, se=TRUE)

# SAME AS ABOVE, AGGREGATED BY OUTCOME, WITH PREDICTION INTERVALS
blup(model, se=TRUE, pi=TRUE, aggregate="outcome")
```

```
# WITH VCOV, FORCED TO A LIST
blup(model, se=TRUE, pi=TRUE, vcov=TRUE, aggregate="outcome")

# PREDICTING ONLY THE RANDOM-EFFECT RESIDUALS
blup(model, type="residual")
```

`coef.mixmeta` *Extract Coefficients and (Co)Variance Matrix from mixmeta Objects*

Description

These method functions return the estimated fixed-effects coefficients and their (co)variance matrix for fitted meta-analytical models represented in objects of class `"mixmeta"`.

Usage

```
## S3 method for class 'mixmeta'
coef(object, format=c("vector","matrix"), ...)

## S3 method for class 'mixmeta'
vcov(object, ...)
```

Arguments

`object`	an object of class `"mixmeta"`.
`format`	format of the returned object.
`...`	further arguments passed to or from other methods.

Value

For `coef`, by default a vector (default)with the estimated fixed-effects coefficients. For multivariate models, a matrix can also be returned.

For `vcov`, the (co)variance matrix of the estimated fixed-effects coefficients.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)

# COEFFICIENTS
model$coef
coef(model)
coef(model, format="matrix")
summary(model)$coef

# (CO)VARIANCE MATRIX
vcov(model)
```

---

dbs *Deep-Brain Stimulation for Patients with Parkinson's Disease*

Description

The dataset contains the data on 46 studies published between 1980 and 2004 that assessed the effect of deep-brain stimulation on the relief of symptoms of Parkinson's disease. The outcome is reported as a score motor function, defined with the Unified Parkinson's Disease Rating Scale (UPDRS-part III), with lower values indicating better prognosis. Changes in the score were measured at 3, 6, 12 months and long-term after the implantation of the stimulator.

Usage

```
dbs
```

Format

A data frame with 68 observations on the following 12 variables:

- `author`: label identifying the study.
- `year`: year of publication.
- `eff_month3,var_month3`: point estimate and variance of the change in the score at 3 months.
- `eff_month6,var_month6`: point estimate and variance of the change in the score at 6 months.
- `eff_month12,var_month12`: point estimate and variance of the change in the score at 12 months.
- `eff_long,var_long`: point estimate and variance of the change in the score in the long term.
- `duration`: average disease duration (years).
- `baseline`: average baseline score of the patients.

Details

The data are stored in a *wide* format, with each record belonging to a single study and different variables providing estimates of the outcome at different times. Each study report results at one or multiple times, with the remaining times set to missing. See the dataset gliomas for an example of similar dataset stored in *long* format.

Note

The data provide an example of application of longitudinal meta-analysis, with repeated measurements of the effect size taken at various time point within each study. This requires a modelling structure that accounts for both within and between-study correlations of repeated measurements. In this case, the analysis is performed in the wide-format dataset using a multivariate meta-analysis. However, a long format is better suited for longitudinal meta-analysis, as it is applicable even when estimates are reported at different times in each study (see the examples in the help page of the dataset gliomas). Results can be compared with those reported Ishak and colleagues (2007).

Source

Ishak KJ, et al (2007). Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *Clinical Trials*. **4**(5):525–39.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN ISHAK ET AL (2007), TABLES 1 AND 2

# CREATE THE OUTCOME AND WITHIN-STUDY MATRICES (THE LATTER WITHOUT CORRELATION)
y <- as.matrix(dbs[1:4*2+1])
S <- as.matrix(dbs[1:4*2+2])

# INDEPENDENT RANDOM EFFECTS (TABLE 1, FIRST MODEL)
mv1 <- mixmeta(y ~ 1, S, bscov="diag", data=dbs)
print(summary(mv1), digits=1, report="var")

# HETEROGENEOUS AR1 RANDOM-EFFECTS (TABLE 1, THIRD MODEL)
mv3 <- mixmeta(y ~ 1, S, bscov="har1", data=dbs)
print(summary(mv3), digits=1, report="var")

# BUILD THE LIST HETEROGENEOUS AR1 WITHIN-STUDY ERRORS (CORRELATION AT 0.97)
cormat <- 0.97^abs(col(matrix(1,4,4)) - row(col(matrix(1,4,4))))
addS <- lapply(seq(nrow(S)), function(i) inputcov(sqrt(S[i,]), cormat))
addS <- lapply(addS, function(x) x[apply(!is.na(x),1,any), apply(!is.na(x),2,any)])

# ADD HAR1 WITHIN-STUDY ERRORS (TABLE 1, FOURTH MODEL) USING addSlist
## Not run:
mv4 <- mixmeta(y ~ 1, bscov="har1", data=dbs, control=list(addSlist=addS))
print(summary(mv4), digits=1, report="var")
## End(Not run)

## Not run:
### USE A LONG FORMAT, AS MORE FLEXIBLE AND ALLOWS MORE COMPLEX MODELS

# RESHAPE THE DATASET
long <- reshape(dbs, direction="long", idvar="author", v.names=c("eff","var"),
  varying=list(1:4*2+1, 1:4*2+2))

# RE-RUN THE LAST (FOURTH) MODEL
```

```
mv4b <- mixmeta(eff ~ factor(time) - 1, random = ~ factor(time) -1 | author,
  bscov="har1", data=long, control=list(addSlist=addS))
print(summary(mv4b), digits=1, report="var")

# COMMON RANDOM EFFECTS (TABLE 1, SECOND MODEL)
mv2 <- mixmeta(eff ~ factor(time) - 1, var, random = ~ factor(time) -1 | author,
  bscov="id", data=long)
print(summary(mv2), digits=1, report="var")

# FOURTH MODEL WITH ADDITIONAL CENTERED META-PREDICTORS (TABLE 2)
mv4plus <- mixmeta(eff ~ factor(time) - 1 + I(duration-14) + I(baseline-52),
  random = ~ factor(time) -1 | author, bscov="har1", data=long,
  control=list(addSlist=addS))
print(summary(mv4plus), digits=1, report="var")
## End(Not run)


### SEE help(gliomas) FOR A COMPLEMENTARY EXAMPLE
```

`fibrinogen` *Fibrinogen Studies Collaboration*

Description

The Fibrinogen Studies Collaboration is a meta-analysis of individual data on 154,012 adults from 31 prospective cohort studies with information on plasma fibrinogen and major disease outcomes. The dataset reports a subset of the results of a first-stage analysis consisting of the log-hazard ratio of coronary heart disease for categories of levels of fibrinogen versus a baseline category.

Usage

```
fibrinogen
```

Format

A data frame with 31 observations on the following 15 variables:

- `cohort`: study ID.
- `b2,b3,b4,b5`: estimated log-hazard ratios for the second to fifth categories versus the baseline category.
- `V_2_2,V_3_3,V_4_4,V_5_5`: variances of the estimated log-hazard ratios.
- `V_2_3,V_2_4,V_2_5,V_3_4,V_3_5,V_4_5`: covariances of the estimated log-hazard ratios.

Details

The published analysis adopted a fixed-effects model on 10 categories of fibrinogen (Fibrinogen Studies Collaboration 2004, 2005). Here a subset of the results of the first-stage analysis is reported, namely the log-hazard ratio for 4 categories and associated (co)variance terms, ordered as the lower triangular elements of the (co)variance matrix taken by column. Details on the first-stage model and the second-stage meta-analysis are provided in White (2009) and Jackson and colleagues (2010).

Note

The data provide an example of application of multivariate meta-analysis for multi-parameter association, where a relationship is defined by functions specified by several coefficients. In this case, the coefficients refer to log-hazard ratio for strata of the original variable versus a baseline category. A general overview of the application of multivariate meta-analysis in this setting is provided by Gasparrini and colleagues (2012).

Source

Fibrinogen Studies Collaboration (2004). Collaborative meta-analysis of prospective studies of plasma fibrinogen and cardiovascular disease. *European Journal of Cardiovascular Prevention and Rehabilitation*. **11**:9–17.

Fibrinogen Studies Collaboration (2005). Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of the American Medical Association*. **294**:1799–1809.

White IR (2009). Multivariate random-effects meta-analysis. *Stata Journal*. **9**(1):40–56.

Jackson D, White IR, Thompson SG (2010). Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine*. **29**(12):1282–1297.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN WHITE (2009) AND JACKSON ET AL. (2010)

# INSPECT THE DATA
head(fibrinogen)

# REML MODEL
y <- as.matrix(fibrinogen[2:5])
S <- as.matrix(fibrinogen[6:15])
model <- mixmeta(y, S)

# SUMMARIZE THE RESULTS
print(summary(model), digits=3)
round(model$Psi, 3)
```

---

gliomas *Randomized Trials on Therapies for Malignant Gliomas*

Description

The dataset contains the data on 17 randomized controlled trials comparing post-operative radiation therapy plus chemotherapy versus radiation therapy alone in patients with malignant gliomas. The outcome of interest is the probability of survival along time, measured as odds ratio at 6, 12, 18, and 24 months.

Usage

```
gliomas
```

Format

A data frame with 68 observations on the following 8 variables:

- `study`: number identifying the trial.
- `time`: time (months) since the start of the treatment at which survival status is assessed.
- `ntreat,streat`: number of total patients at the beginning of the study and surviving patients at specific times, respectively, in the treatment group (radiation therapy plus chemotherapy).
- `dcontr,ncontr`: number of total patients at the beginning of the study and surviving patients at specific times, respectively, in the control group (radiation alone).
- `logOR,varOR`: log-odds ratio of survival between treatment and control groups.

Details

The data are stored in a *long* format, with each record providing the estimate at a single time and each study providing multiple records. There were missing data for study 17 at months 6 and 18. There were no survivors in the control group at month 24 for studies 3 and 10, although this still allows computation of the OR. See the dataset `dbs` for an example of similar dataset stored in *wide* format.

The log-odds ratio is computed empirically as $\log(s_t \times (n_t - s_t)/((n_c - s_c) \times s_c)$. Its variance is simply computed as $1/s_t + 1/(n_t - s_t) + 1/(n_c - s_c) + 1/s_c$.

Note

The data provide an example of application of longitudinal meta-analysis, with repeated measurements of the effect size taken at various time points within each study. This requires a modelling structure that accounts for both within and between-study correlations of repeated measurements. In this case, the same analysis can be performed in a wide-format dataset using a multivariate meta-analysis (see the examples in the help page of the dataset `dbs`). However, the long format is better suited for longitudinal meta-analysis, as it is applicable even when estimates are reported at different times in each study. Results can be compared with those reported Musekiwa and colleagues (2016). The same dataset was also used by Trikalinos and Olkin (2012), using a similar modelling scheme.

Source

Musekiwa A, et al (2012). Meta-analysis of effect sizes reported at multiple time points using general linear mixed model. *Plos One*. **11**(10):e0164898.

Trikalinos TA, Olkin I (2012). Meta-analysis of effect sizes reported at multiple time points: a multivariate approach. *Clinical Trials*. **9**(5):610–620.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN MUSEKIWA ET AL (2012), TABLES 3 AND 4

# INDEPENDENT RANDOM EFFECTS, NO WITHIN-STUDY CORRELATION (MODEL 1)
mod1 <- mixmeta(logOR~0+factor(time), S=logORvar, random=~0+factor(time)|study,
  bscov="diag", data=gliomas)
print(summary(mod1), digits=2, report="var")

# COMPOUND-SYMMETRY RANDOM EFFECTS, NO WITHIN-STUDY CORRELATION (MODEL 2)
# NB: THIS REQUIRES A TWO-LEVEL MODEL WITH THE INNER-LEVEL VARIANCE FIXED TO 0
unit <- factor(seq(nrow(gliomas)))
mod2 <- mixmeta(logOR~0+factor(time), S=logORvar, random=~1|study/unit,
  bscov=c("unstr","fixed"), data=gliomas, control=list(Psifix=list(unit=0)))
print(summary(mod2), digits=2, report="var")

# BUILD THE HETEROGENEOUS AR1 WITHIN-STUDY ERRORS (CORRELATION AT 0.61)
cormat <- 0.61^abs(col(matrix(1,4,4)) - row(col(matrix(1,4,4))))
addS <- lapply(split(sqrt(gliomas$logORvar), gliomas$study), inputcov, cormat)
addS <- lapply(addS, function(x) x[apply(!is.na(x),1,any), apply(!is.na(x),2,any)])

# INDEPENDENT RANDOM EFFECTS, HAR1 WITHIN-STUDY CORRELATION (MODEL 4)
mod4 <- mixmeta(logOR~0+factor(time), random=~0+factor(time)|study,
  bscov="diag", data=gliomas, control=list(addSlist=addS))
print(summary(mod4), digits=2, report="var")

# UNSTRUCTURED RANDOM EFFECTS, HAR1 WITHIN-STUDY CORRELATION (MODEL 6)
mod6 <- update(mod4, bscov="unstr")
print(summary(mod6), digits=2, report="var")

# COMPARE THE FIT
AIC(mod1, mod2, mod4, mod6)

## Not run:
### MORE FLEXIBLE MODELLING OF RANDOM EFFECTS

# RE-RUN BEST FITTING MODEL WITH ML (ALLOWS TESTING OF FIXED EFFECTS)
mod4ml <- update(mod4, method="ml")

# RANDOM-SLOPE MODEL WITH TIME AS CONTINUOUS (CENTERED IN random)
modnew <- mixmeta(logOR~time, random=~I(time-15)|study, bscov="diag",
  method="ml", data=gliomas, control=list(addSlist=addS, maxiter=200))
print(summary(modnew), digits=2, report="var")

# COMPARE
AIC(mod4ml, modnew)
## End(Not run)


### SEE help(dbs) FOR A COMPLEMENTARY EXAMPLE
```

hsls *High School Longitudinal Study*

Description

This is a nationally representative, longitudinal study of more than 21,000 9th graders in 944 schools who will be followed through their secondary and postsecondary years. The data are used for testing whether sex, socioeconomic status and sex by socio-economic status interaction are predictive of the mathematics standardized score in each of the eight race groups.

Usage

```
hsls
```

Format

A data frame with 8 observations on the following 10 variables:

- `race`: race group.
- `b1,b2,b3`: estimated regression coefficients for sex, socio-economic status and sex by socio-economic status interaction, respectively, on the mathematics standardized score.
- `V11,V22,V33`: variances of the estimated coefficients.
- `V12,V13,V23`: covariances of the estimated coefficients.

Source

Chen H, Manning AK, Dupuis J (2012). A method of moments estimator for random effect multivariate meta-analysis. *Biometrics*. **68**(4):1278–1284.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN CHEN ET AL. (2012)

# INSPECT THE DATA
hsls

# FIXED-EFFECTS MODEL
S <- as.matrix(hsls[5:10])
mod1 <- mixmeta(cbind(b1,b2,b3), S, data=hsls, method="fixed")
summary(mod1)

# MM MODEL
mod2 <- mixmeta(cbind(b1,b2,b3), S,data=hsls, method="mm")
summary(mod2)
mod2$Psi
```

hyp *Ten Studies Assessing an Hypertension Treatment*

Description

The dataset contains the results of ten studies that assess the effectiveness of hypertension treatment for lowering blood pressure. Each study provides complete data on two treatment effects, the difference in systolic blood pressure (SBP) and diastolic blood pressure (DBP) between the treatment and the control groups, where these differences are adjusted for the participants' baseline blood pressures. The within-study correlations of the two outcomes are known. Some trials are conducted on patients with isolated systolic hypertension (ISH).

Usage

```
hyp
```

Format

A data frame with 10 observations on the following 7 variables:

- `study`: study ID.
- `sbp,sbp_se`: estimated difference and its standard error in systolic blood pressure.
- `dbp,dbp_se`: estimated difference and its standard error in diastolic blood pressure.
- `rho`: within-study correlation between the estimated differences in systolic and diastolic blood pressure.
- `ish`: indicator for studies on patients with isolated systolic hypertension.

Note

The standard errors for the two outcomes are wrongly reported as variances in the original article by Jackson and colleagues (2013).

Source

Jackson D, White IR, Riley RD (2013). A matrix based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biometrical Journal*. **55**(2):231–45.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN JACKSON ET AL. (2013)

# INSPECT THE DATA
hyp
```

```
# INPUT THE CORRELATION (CAN ALSO BE INPUTTED DIRECTLY, SEE BELOW)
(S <- inputcov(hyp[c("sbp_se", "dbp_se")], cor=hyp$rho))
# CHECK WITH THE FIRST STUDY
cov2cor(xpndMat(S[1,]))

# META-ANALYSIS, REML MODEL
mod1 <- mixmeta(cbind(sbp,dbp), S=S, data=hyp)
print(summary(mod1), digits=2)
round(mod1$Psi,2)

# META-ANALYSIS, REML MODEL (INPUTTING THE CORRELATION DIRECTLY)
mod2 <- mixmeta(cbind(sbp,dbp), S=cbind(sbp_se,dbp_se)^2, data=hyp,
  control=list(Scor=hyp$rho))
print(summary(mod2), digits=2)

# META-ANALYSIS, MM MODEL
mod3 <- mixmeta(cbind(sbp,dbp), S=S, data=hyp, method="mm")
print(summary(mod3), digits=2)
round(mod3$Psi,2)

# META-REGRESSION, REML MODEL
mod4 <- mixmeta(cbind(sbp,dbp) ~ ish, S=S, data=hyp)
print(summary(mod4), digits=2)

# META-REGRESSION, MM MODEL
mod5 <- mixmeta(cbind(sbp,dbp) ~ ish, S=S, data=hyp, method="mm")
print(summary(mod5), digits=2)
```

`inputcov` *Input (Co)Variance Matrices*

Description

This function inputs (co)variance matrices of a set of outcomes given the corresponding standard deviation and correlation values.

Usage

```
inputcov(sd, cor=NULL)
```

Arguments

`sd`	a $m \times k$ matrix of standard deviations for k outcomes in m matrices, or a vector for k outcomes in a single matrix.
`cor`	either a vector of length 1, m or $k(k-1)/2$, or alternatively a $k \times k$ or $m \times k(k-1)/2$ matrix. See Details.

Details

Depending the number of outcomes k and matrices m, the argument `cor` is interpreted as:

- if a vector of length 1 (a scalar), the same correlation for all the k outcomes for all the m matrices;
- if a vector of length m, the same correlation for all the k outcomes for each of the m matrices;
- if a vector of length $k(k-1)/2$, the lower triangular elements (without diagonal, taken by column) of the correlation matrix of the k outcomes, the same for all the m matrices;
- if a $k \times k$ matrix, the correlation matrix for the single matrix (only when m=1);
- if a $m \times k(k-1)/2$ matrix, each row represents the lower triangular elements (without diagonal, taken by column) of the correlation matrix of the k outcomes for each of the m matrices.

Value

For a single matrix, the (co)variance matrix itself. For multiple matrices, a $m \times k(k+1)/2$ matrix, where each row represents the vectorized entries of the lower triangle (with diagonal, taken by column) of the related (co)variance matrix (see `vechMat`).

Note

This function is called internally by `mixmeta` for multivariate models to input the correlation(s) when only the within-unit variances are provided through the argument `S`. In this case, the correlation values are set through the argument `Scor` in the control list (see `mixmeta.control`).

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See `xpndMat`. See `mixmeta.control`.

Examples

```
# SOME RANDOM SD FOR A SINGLE MATRIX, WITH CONSTANT CORRELATION
(M <- inputcov(runif(4, 0.1, 3), 0.7))
# CHECK CORRELATION
cov2cor(M)

# NOW WITH A MORE COMPLEX CORRELATION STRUCTURE
(M <- inputcov(runif(3, 0.1, 3), c(0.7,0.2,0.4)))
cov2cor(M)

# MULTIPLE MATRICES
(V <- matrix(runif(5*3, 0.1, 3), 5, 3,
  dimnames=list(1:5, paste("V", 1:3, sep=""))))
inputcov(V, 0.6)

# WITH REAL DATA WHEN CORRELATIONS AVAILABLE
```

```
hyp
(S <- inputcov(hyp[c("sbp_se","dbp_se")], cor=hyp$rho))
# CHECK FIRST STUDY
cov2cor(xpndMat(S[1,]))

# USED INTERNALLY IN mixmeta
p53
inputcov(sqrt(p53[c("V1","V2")]), 0.5)
model <- mixmeta(cbind(y1,y2), S=cbind(V1,V2), data=p53, control=list(Scor=0.5))
model$S
```

`inputna` *Input Missing Values*

Description

This function augment data by replacing missing values. It can be used internally in `mixmeta` through the `control` list.

Usage

```
inputna(y, S, inputvar=10^4)
```

Arguments

`y`	Assuming a meta-analysis or meta-regression based on n units and k outcomes: a n-dimensional vector (for univariate models) or $m \times k$ matrix (for multivariate models) of outcomes.
`S`	series of within-unit variances (or (co)variance matrices for multivariate models) of the estimated outcome(s). For univariate models, this is usually a n-dimensional vector. For multivariate models, it can be provided as: a m-dimensional list of $k \times k$ matrices; a tri-dimensional $k \times k \times m$ array; a matrix or data frame with n rows and $k(k+1)/2$ or k columns, depending on the availability of the within-unit correlations.
`inputvar`	multiplier for inputting the missing variances in `S`.

Details

The function augments the data by replacing missing values in the outcomes and the associated (co)variances. Specifically, it replaces missing outcomes and missing covariances (if provided) with 0, and missing variances with the largest observed variance multiplied by `inputvar`. This value is expected to be very high, by default 10^4, so that the corresponding observation contributes only negligibly to the final estimate.

Value

A matrix with the first k column corresponding to the augmented outcomes, and the remaining $k(k+1)/2$ or k columns (depending on the availability of the within-study covariances) corresponding to vectorized entries of the lower triangle of the related (co)variance matrices.

Note

Data augmentation used to be the approach to deal with missing values in the first implementation of **mixmeta**. The current algorithms directly account for missing.

Inputting missing values can be useful when two or more outcomes are never observed jointly, and the estimation is entirely based on indirect comparison. This method can be applied in network meta-analysis, also called indirect treatment comparison.

This approach can produce different results than standard methods, especially when the occurrence of missing is substantial. Preliminary analyses indicate that likelihood-based estimation methods do not seem to be affected, while non-iterative estimators such as method of moments and variance components are more sensitive. The user should be careful on the application of missing augmentation.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Jackson D, Riley R, White IR (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*. **30**(20);2481–2498.

White IR (2009). Multivariate random-effects meta-analysis. *Stata Journal*. **9**(1):40–56.

White IR (2011). Multivariate random-effects meta-regression: updates to mvmeta. *Stata Journal*. **11**(2):255-270.

See Also

See `inputcov` for inputting (co)variance matrices.

Examples

```
# INSPECT THE DATA
head(smoking)

# STANDARD APPROACH TO MISSING DATA
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod1 <- mixmeta(y, S)
summary(mod1)

# WITH DATA AUGMENTATION
augdata <- inputna(y, S)
y <- augdata[,1:3]
S <- augdata[,-c(1:3)]
```

```
mod2 <- mixmeta(y, S)
summary(mod2)
# NB: SAME PARAMETER ESTIMATES, BUT WRONG NYUMBER OF OBS

# USED INTERNALLY IN mixmeta
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod3 <- mixmeta(y, S, control=list(inputna=TRUE))
summary(mod3)
# NOW RIGHT NUMBER OF OBS
```

`logLik.mixmeta` *Extract Log-Likelihood from mixmeta Objects*

Description

This method function returns the (restricted) log-likelihood for fitted meta-analytical models represented in objects of class `"mixmeta"`.

Usage

```
## S3 method for class 'mixmeta'
logLik(object, ...)
```

Arguments

`object`	an object of class `"mixmeta"`.
`...`	further arguments passed to or from other methods.

Value

A numeric scalar of class `"logLik"` with attributes, providing the (restricted) log likelihood of the model. Attributes correspond to the component `df` of `mixmeta` objects, namely the following scalars: `nall` (number of observations used for estimation, excluding missing values), `nobs` (equal to `nall`, minus the number of fixed-effects coefficients for REML models, `fixed` (number of estimated fixed-effects coefficients), `random` (number of estimated (co)variance terms).

Note

This functions is called by `AIC` and `BIC` for computing the Akaike and Bayesian information criteria.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See the default method logLik. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL)~pubyear,S=berkey98[5:7],data=berkey98)

# LOG-LIKELIHOOD
ll <- logLik(model)
ll
attributes(ll)

# AIC and BIC
AIC(model)
BIC(model)
```

---

mixmeta *Fitting Standard and Extended Meta-Analysis and Meta-Regression Models*

Description

The function mixmeta performs various meta-analytical models under a common mixed-effects framework, including standard univariate fixed and random-effects meta-analysis and meta-regression, and non-standard extensions such as multivariate, multilevel, longitudinal, and dose-response models. The function mixmeta.fit is a wrapper for actual fitting functions based on different estimation methods, usually called internally. See mixmeta-package for an overview.

Usage

```
mixmeta(formula, S, data, random, method="reml", bscov="unstr", offset, subset,
  contrasts=NULL, na.action, model=TRUE, control=list())

mixmeta.fit(X, Z, y, S, groups, method, bscov, control)
```

Arguments

Assuming a meta-analysis or meta-regression based on n units aggregated within m (outer-level) groups, k outcomes, p fixed-effects predictors, and q random-effects predictors:

an object of class "formula" (or one that can be coerced to that class) offering a symbolic description of the linear predictor for the fixed-effects part of the model. Alternatively, for meta-analysis with no fixed-effects predictors, a single vector (for univariate models) or matrix-type object (for multivariate models). Terms in formula must be vector or matrix-type objects, optionally provided in the data argument below. See mixmetaFormula.

S	
formula	series of within-unit variances (or (co)variance matrices for multivariate models) of the estimated outcome(s). For univariate models, this is usually a n-dimensional vector. For multivariate models, it can be provided as: a m-dimensional list of $k \times k$ matrices; a tri-dimensional $k \times k \times m$ array; a matrix or data frame with n rows and $k(k+1)/2$ or k columns, depending on the availability of the within-unit correlations. `mixmeta.fit` accepts only the last option. Optionally, for more complex error structures, this argument can be omitted and passed through `addSlist` in `control`. See Details below.
data	an optional data frame, list or environment (or object coercible by `as.data.frame` to a data frame) containing the variables in `formula` and `random`. If not found in `data`, the variables are taken from `environment(formula)`, typically the environment from which `mixmeta` is called.
random	a one-sided formula (or a list of formulae for multilevel models) offering a symbolic description of the linear predictor(s) and grouping structure for the random-effects part of the model. The usual form is `~ z1 + ... + zq | g`, with the grouping factor separated from the linear predictor by the symbol `'|'`. Multiple levels with the same linear predictor can be defined by separating multiple grouping factors using the symbol `'/'`. Alternatively, in a list form the grouping factors can also be provided as list names. In both cases, the levels are considered nested (from outer to inner following the order). See `mixmetaFormula` and Details below.
method	estimation method: `"fixed"` for fixed-effects models, `"ml"` or `"reml"` for random-effects models fitted through (restricted) maximum likelihood, `"mm"` for random-effects models fitted through method of moments, and `"vc"` for random-effects models fitted through variance components. See Details below. If `"model.frame"`, the model frame is returned, as in `lm` or `glm`.
bscov	a character vector defining the structure of the random-effects (co)variance matrices. Default to `"unstr"` (unstructured). Names corresponding to grouping factors (see `random` above) can be used to refer to specific random-effects levels for non-default values. If unnamed, the values can be recycled. Among various `(co)variance structures`, the user can select `"diag"` (diagonal), `"cs"` (compound symmetry), `"hcs"` (heterogeneous compound symmetry), `"ar1"` (autoregressive of first order), or `"fixed"` (fixed). See also Details.
offset	optionally, a n-dimensional numeric vector used to specify an a priori known component in the linear predictor. One or more `offset` terms can be included in the formula instead or as well. See `model.offset`.
subset	an optional vector specifying a subset of observations to be used in the fitting process.
contrasts	an optional list. See the `contrasts.arg` of `model.matrix`.
na.action	a function which indicates what should happen when the data contain `NA`s. Default to `na.action` setting of `options`, usually `na.omit`. `na.exclude` can be useful. See details on `missing values` in `mixmeta`.
model	a logical value indicating whether the model frame should be included as a component of the returned value. See the `model.frame` method function.
control	list of parameters for controlling the fitting process. These are passed to `mixmeta.control` to replace otherwise selected default values.

X	a $n \times p$ design matrix containing the p fixed-effects predictors, appropriately ordered by groups. Usually produced internally by mixmeta from formula above.
Z	a $n \times q$ design matrix (or a list of design matrices for multilevel models) containing the q random-effects predictors, appropriately ordered by groups. Usually produced internally by mixmeta from random above.
y	a n-dimensional vector (for univariate models) or $m \times k$ matrix (for multivariate models) of outcomes, appropriately ordered by groups. Usually produced internally by mixmeta from formula above.
groups	matrix with n rows, with each column identifying the groups for each level of random-effects, appropriately ordered. Usually produced internally by mixmeta from random above.

Details

The function mixmeta resembles standard regression functions in R. See lme in particular, or lm or glm, for information on most of the arguments. Internally, this function assembles the data components, defines the (potentially multiple) grouping levels and re-order the data accordingly, and then pass them to mixmeta.fit. This is a wrapper for actual fitting functions that are automatically selected. Functions other than mixmeta are not expected to be called directly for model fitting.

Fixed or random-effects models for meta-analysis are simply defined using y ~ 1 in formula, where y is a response vector optionally stored in data. In meta-regression models, other terms are added in the right-hand side of the formula as y ~ x1 + ... + xp, defining the linear meta-predictor. Factors, variable transformations and interactions are allowed, following the usual formula specification (see mixmetaFormula).

In this standard usage, each of the n rows is assumed to represent a single estimate of an outcome from a set of independent studies. In random-effects models, the grouping structure is automatically derived by assigning a group to each row of data (with $m = n$). Extensions to multivariate models ($k > 1$) are straightforward, and only require using a matrix in the left-hand side, where each of the k columns represents a different outcome, or the form cbind(y1,...,yk) ~ 1. See mixmetaFormula.

Non-standard random-effects models can be specified through the optional argument random. This is commonly represented by a one-sided formula, whose basic random-intercept form is ~ 1 | g, where g is a grouping factor. A more complex linear meta-predictor for the random-effects part can be also specified by ~ z1 + ... + zq | g. The argument random also accepts a list of one-sided formulae referring to multiple random-effects levels (see mixmetaFormula). The use of random extends the standard meta-analytical setting by relaxing the assumption of independence between units, allowing multiple estimates from the same group (with $m < n$) and multiple nested grouping levels. This provides the possibility to fit longitudinal, multilevel, and dose-response meta-analysis, among other extensions. See the examples below.

The argument bscov allows the definition of specific structures for the random-effects (co)variance matrices corresponding the each level. The default unstructured form requires kq(kq+1)/2 parameters for a single-level meta-analysis. The choice of other structures reduces the number of parameters, although requiring stronger assumptions. More information and complete list of options is available at a specific help page (see mixmetaCovStruct).

The within-unit (co)variances are provided through the argument S, usually as a matrix. If the correlations are available, each of the m row represents the $k(k+1)/2$ vectorized entries of the

lower triangle of the related (co)variance matrix, taken by column (see xpndMat). If correlations are not available, each row represents the k variances, and the correlations are inputted internally through the argument Scor of the control list (see inputcov). For more complex error structures that span multiple units, the argument S can be omitted and passed through addSlist in control, although requiring the observations to be re-ordered accordingly to groups (see mixmeta.control).

Different estimator are available in the package mixmeta and chosen through the argument method, with related fitting functions called internally. In the current version, the options are:

- method="fixed": Fixed-effects estimator
- method="ml": Maximum likelihood (ML) estimator
- method="reml": Restricted maximum likelihood (REML) estimator
- method="mm": Method of moments estimator
- method="vc": Variance components estimator

Note that non-standard random-effects models and the use of structured (co)variance matrices are only available for "ml" and "reml" methods. See their help pages for further details on the estimation procedures, following the links above.

Missing values are allowed in both sides of formula. In the case of missing predictors (right-hand side of formula), the related unit is entirely excluded from estimation. In contrast, a unit still contributes to estimation if at least outcome is non-missing. This behaviour is different from standard regression functions such as lm or glm. Before the call to mixmeta.fit, units matching such stricter missing definition are removed from the the model frame. The missing pattern in S must be consistent with that in y. See further details on handling missing values in mixmeta.

The fitting procedure can be controlled through the additional terms specified in control, which are passed to the function mixmeta.control.

Value

The mixmeta function typically returns a list object of class "mixmeta" representing the meta-analytical model fit, as described in mixmetaObject. When method="data.frame", the model is not fitted and the model frame is returned, namely a data frame with special attributes (see the default method model.frame) and, in this case, the additional class "data.frame.mixmeta".

The wrapper function mixmeta.fit is usually called internally in mixmeta, and returns an intermediate list object with some of the components expected in the "mixmeta" class.

Several method functions for regression objects are available, either default or specifically written for the "mixmeta" class. See mixmetaObject for a complete list.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Gasparrini A. (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

See Also

See additional info on the estimation procedures at the related page of the fitting functions See mixmetaFormula for the use of formulae to define the fixed and random parts of the model. See alternative (co)variance structures for likelihood-based estimation methods. See handling of missing values in mixmeta. See lme, lm or glm for standard regression functions. See mixmeta-package for an overview of this modelling framework.

Examples

```
### STANDARD MODELS

# RANDOM-EFFECTS META-ANALYSIS, ESTIMATED WITH REML
model <- mixmeta(logor, logorvar, data=bcg)
summary(model)

# RANDOM-EFFECTS META-REGRESSION, ESTIMATED WITH ML
model <- mixmeta(logor~ablat, logorvar, data=bcg, method="ml")
summary(model)


### MAIN METHOD FUNCTIONS

# COEFFICIENTS AND (CO)VARIANCE MATRIX
coef(model)
vcov(model)

# RESIDUALS AND FITTED VALUES
residuals(model)
fitted(model)

# MODEL FRAME AND MODEL MATRIX
model.frame(model)
model.matrix(model)

# LOG-LIKELIHOOD AND AIC VALUE
logLik(model)
AIC(model)

# COCHRAN Q TEST FOR RESIDUAL HETEROGENEITY
qtest(model)


### PREDICTIONS

# PREDICTED EFFECTS
predict(model)
predict(model, se=TRUE)
predict(model, newdata=data.frame(ablat=2:5*10), ci=TRUE)

# BEST LINEAR UNBIASED PREDICTION
blup(model)
blup(model, pi=TRUE)
```

```
# SEE help(predict.mixmeta) AND help(BLUP.mixmeta) FOR MORE INFO


### MULTIVARIATE MODELS

### BIVARIATE MODELS
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)
summary(model)
residuals(model)

### MULTIVARIATE META-ANALYSIS WITH MORE THAN 2 OUTCOMES
y <- as.matrix(fibrinogen[2:5])
S <- as.matrix(fibrinogen[6:15])
model <- mixmeta(y, S)
summary(model)
predict(model, se=TRUE)
predict(model, se=TRUE, aggregate="outcome")


### OTHER EXTENSIONS

# MULTILEVEL META-ANALYSIS
model <- mixmeta(effect, var, random= ~ 1|district/study, data=school)
summary(model)
# SEE help(school) AND help(thrombolytic) FOR MORE EXAMPLES

# DOSE-RESPONSE META-ANALYSIS (SIMPLIFIED)
model <- mixmeta(logrr ~ 0 + dose, S=se^2, random= ~ 0 + dose|id, data=alcohol,
 subset=!is.na(se))
summary(model)
# SEE help(alcohol) FOR MORE EXAMPLES

# LONGITUDINAL META-ANALYSIS
model <- mixmeta(logOR~time, S=logORvar, random=~I(time-15)|study, data=gliomas)
summary(model)
# SEE help(gliomas) AND help(dbs) FOR MORE EXAMPLES


### FIXED-EFFECTS MODELS AND ALTERNATIVE ESTIMATORS

# FIXED-EFFECTS MODEL
model <- mixmeta(sbp~ish, S=sbp_se^2, data=hyp, method="fixed")
summary(model)

# METHOD OF MOMENTS
S <- as.matrix(hsls[5:10])
model <- mixmeta(cbind(b1,b2,b3), S, data=hsls, method="mm")
summary(model)

# VARIANCE COMPONENTS ESTIMATOR
model <- mixmeta(cbind(PD,AL)~pubyear, S=berkey98[5:7], data=berkey98,
  method="vc")
```

```
summary(model)


### IN THE PRESENCE OF MISSING VALUES

# RUN THE MODEL
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
model <- mixmeta(y, S)
summary(model)
model.frame(model)

# SEE help(na.omit.data.frame.mixmeta) FOR MORE EXAMPLES


### WHEN WITHIN-STUDY COVIARIANCES ARE NOT AVAILABLE AND/OR NEED TO BE INPUTTED

# GENERATE S
(S <- inputcov(hyp[c("sbp_se","dbp_se")], cor=hyp$rho))

# RUN THE MODEL
model <- mixmeta(cbind(sbp,dbp), S=S, data=hyp)

# INPUTTING THE CORRELATION DIRECTLY IN THE MODEL
model <- mixmeta(cbind(y1,y2), cbind(V1,V2), data=p53, control=list(Scor=0.95))
summary(model)

# SEE help(hyp) AND help(p53) FOR MORE EXAMPLES


### STRUCTURING THE BETWEEN-STUDY (CO)VARIANCE

# DIAGONAL
S <- as.matrix(hsls[5:10])
model <- mixmeta(cbind(b1,b2,b3), S, data=hsls, bscov="diag")
summary(model)
model$Psi

# COMPOUND SYMMETRY
model <- mixmeta(cbind(b1,b2,b3), S, data=hsls, bscov="cs")
summary(model)
model$Psi

# SEE help(mixmetaCovStruct) FOR DETAILS AND ADDITIONAL EXAMPLES


### USE OF THE CONTROL LIST

# PRINT THE ITERATIONS AND CHANGE THE DEFAULT FOR STARTING VALUES
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
model <- mixmeta(y, S, control=list(showiter=TRUE, igls.inititer=20))
```

```
# SEE help(mixmeta.control) FOR FURTHER DETAILS
```

`mixmeta.control` *Ancillary Parameters for Controlling the Fit in mixmeta Models*

Description

This internal function sets the parameter options used for fitting meta-analytical models, commonly to pre-specified default values. It is usually internally called by `mixmeta`.

Usage

```
mixmeta.control(optim=list(), showiter=FALSE, maxiter=100, initPsi=NULL, Psifix=NULL,
  Scor=NULL, addSlist=NULL, inputna=FALSE, inputvar=10^4, loglik.iter="hybrid",
 igls.inititer=10, hessian=FALSE, vc.adj=TRUE, reltol=sqrt(.Machine$double.eps),
  checkPD=NULL, set.negeigen=sqrt(.Machine$double.eps))
```

Arguments

`optim`	list of parameters passed to the `control` argument of the function `optim`, which performs the `quasi-Newton optimization` in likelihood-based random-effects models. See `optim` for the list of arguments. See Details for additional info.
`showiter`	logical. If `TRUE`, the progress of iterative optimization is shown.
`maxiter`	positive interger value. Maximum number of iterations in methods involving optimization procedures.
`initPsi`	either a matrix or a vector of its lower triangular elements (with diagonal, taken by column), or optionally a named list with one or more of such objects. Used as starting values of random-effects parameters in likelihood-based optimization routines. See Details.
`Psifix`	either a matrix or a vector of its lower triangular elements (with diagonal, taken by column), or optionally a named list with one or more of such objects. Used to define fixed parts of the random-effects `(co)variance structures`. See Details.
`Scor`	either a scalar, vector or matrix representing the within-unit correlation(s) to be inputted when the covariances are not provided in multivariate models, and ignored if they are. See `inputcov`.
`addSlist`	a list of m matrices for the (outer-level) groups of units defining the (known) error (co)variance structure, when this cannot be passed through the argument `S` of `mixmeta`. See Details.
`inputna`	logical. If missing values must be internally inputted. To be used with caution. See `inputna`.
`inputvar`	multiplier for inputting the missing variances, to be passed as an argument to `inputna`.
`loglik.iter`	iterative scheme used in in likelihood-based optimization routines. Options are `"hybrid"`, `"newton"`, and `"igls"` or `"RIGLS"`. See `mixmeta.ml`.

igls.inititer	number of iterations of the (restricted) iterative generalized least square algorithm when used in the initial phase of hybrid optimization procedure of likelihood-based estimators. See mixmeta.ml.
hessian	logical. If TRUE, the Hessian matrix of the parameters estimated in the optimization process is computed and returned. Only applicable to likelihood-based estimation methods. For details, see the info provided in the help pages of the optimizations algorithms and (co)variance structures.
vc.adj	logical. If TRUE, an adjustement to the way the marginal variance part is computed in the (co)variance components estimator is applied in the variance components estimator. See mixmeta.vc.
reltol	relative convergence tolerance in methods involving optimization procedures. The algorithm stops if it is unable to reduce the value by a factor of reltol * (abs(val) + reltol) at a step.
checkPD	logical. Determines if the semi-positiveness of within-unit error or random-effects (co)variance matrices must be checked.
set.negeigen	positive value. Value to which negative eigenvalues are to be set in estimators where such method is used to force semi-positive definiteness of the estimated between-study (co)variance matrix.

Details

This function has default values for most of the arguments, some of them set internally. Non-default values are passed through the control argument of mixmeta. Many arguments refer to specific fitting procedures. See the help page of the related estimator for details.

The function automatically sets non-default values for some control arguments for optim, unless explicitly set in the list passed to it. Specifically, the function selects fnscale=-1, maxit=maxiter and reltol=reltol, where the latter two are specified by other arguments of this function.

The arguments initPsi and Psifix are used to provide information for estimation procedures of the random-effects parameters in likelihood-based methods. Specifically, the former is used to choose non-default starting values (see mixmeta.ml), and the latter for defining the fixed (known) part of specific (co)variance structures. In multilevel models, these arguments must be lists with named components referring to one or more levels of grouping defined by the argument random of mixmeta.

The argument addSlist can be used to define more complex (known) error structures of the outcome(s) that are usually provided through the argument S of mixmeta as within-unit variances (or (co)variance matrices for multivariate models). This can be useful when these error structures spans multiple units (rows), and the between-unit correlation cannot be defined through S, for instance in dose-response meta-analysis (see examples in mixmeta). Note that this information is passed internally after the data have be re-ordered following the grouping defined by random in mixmeta, and this should be consistent in addSlist. Specifically, the grouping variables are assumed as factors and therefore the groups are taken in alphabetical/numeric order. It is suggested to re-order the data according to this order of the groups before fitting the model, so to ensure consistency between the grouped data and addSlist.

Value

A list with components named as the arguments.

Note

The function is expected to be extended and/or modified at every release of the package **mixmeta**. It is strongly suggested to check the arguments of this function at every release.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See `mixmeta`. See also `glm.control`. See the help pages of the related fitting functions for details on each parameter. See `mixmeta-package` for an overview of this modelling framework.

Examples

```
# PRINT THE ITERATIONS (SEE ?optim) AND CHANGE THE DEFAULT FOR STARTING VALUES
mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98,
  control=list(showiter=TRUE, igls.inititer=20))

# INPUT THE CORRELATION
mixmeta(cbind(y1,y2), S=cbind(V1,V2), data=p53, control=list(Scor=0.5))

# FIX (PARTS OF) THE RANDOM-EFFECTS (CO)VARIANCE MATRIX
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mixmeta(y, S, bscov="prop", control=list(Psifix=diag(3)+1))
```

---

mixmeta.fixed *Fixed-Effects Estimator for mixmeta Models*

Description

This function implements a generalized least square estimator for fixed-effects meta-analysis and meta-regression, including standard univariate models and non-standard multivariate extensions. It is meant to be used internally and not directly run by the users.

Usage

```
mixmeta.fixed(Xlist, ylist, Slist, nall, control, ...)
```

Arguments

	Assuming a meta-analysis or meta-regression based on m independent groups (usually studies) providing a single estimate (unit of analysis), k outcomes and p fixed-effects predictors:
	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.
`Slist`	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
`nall`	numeric scalar with the total number of observations (excluding missing).
`control`	list of parameters for controlling the fitting process, usually internally set to default values by `mixmeta.control`.
`...`	further arguments passed to or from other methods. Currently not used.

Details

The estimation involves only the kp fixed-effects coefficients. Note that, in this fixed-effects estimator, each unit is assumed independent from the others, and therefore the number of groups (the length of the lists) is identical to the number of units (`m=n`). However, this is not important in fixed-effects models, where no random (and therefore grouping) structure is used.

The routine is based on a standard generalized least square (GLS) algorithm implemented in the internal function `glsfit`. The between-study (co)variance matrix is set to zero, so the marginal (co)variance matrix, composed only by elements of the within-unit component, is assumed as completely known. Similarly to the likelihood-based estimators implemented in `mixmeta.ml` and `mixmeta.reml`, the computation involves Cholesky and and QR decompositions for computational stability and efficiency. The method is described in details in Gasparrini and collaborators (2012) (see references below).

Value

These functions return an intermediate list object, with some components then processed and some others added later within `mixmeta.fit` and `mixmeta` to finalize an object of class `"mixmeta"`. See `mixmetaObject`.

Note

As stated earlier, this function is called internally by `mixmeta.fit`, and is not meant to be used directly. In particular, its code does not contain any check on the arguments provided, which are expected in specific formats. The function is however exported in the namespace and documented for completeness.

The arguments above are prepared by `mixmeta.fit` from its arguments `X`, `y` and `S`. The list structure, although requiring more elaborate coding, is computationally more efficient, as it avoids the specification of sparse block-diagonal matrices, especially for meta-analysis involving a large number of studies.

Some parameters of the fitting procedures are determined by the `control` argument, with default set by `mixmeta.control`. No missing values are accepted in the fitting functions. See details on `missing values`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Berkey CS, Anderson JJ, Hoaglin DC (1996). Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine*. **15**(5):537–547.

Berkey CS, et al. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*. **17**(22):2537–2550.

See Also

See `mixmeta` for the general usage of the functions. See `mixmeta.control` to determine specific parameters of the fitting procedures. Use the triple colon operator (':::') to access the code of the internal functions, such as `glsfit`. See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# UNIVARIATE FIXED-EFFECTS MODEL
mod1 <- mixmeta(yC, S=SCC, data=smoking, method="fixed")
summary(mod1)

# MULTIVARIATE FIXED-EFFECTS MODEL
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod2 <- mixmeta(y, S, method="fixed")
summary(mod2)

# MULTIVARIATE FIXED-EFFECTS MODEL: REPLICATE THE RESULTS IN BERKEY ET AL. 1998
mod3 <- mixmeta(cbind(PD,AL) ~ I(pubyear-1983), S=berkey98[5:7], data=berkey98,
  method="fixed")
summary(mod3)
```

`mixmeta.ml` *ML and REML Estimators for mixmeta Models*

Description

These functions implement maximum likelihood (ML) and restricted maximum likelihood (REML) estimators for random-effects meta-analysis and meta-regression, including standard univariate models, and non-standard extensions such as multivariate, multilevel, longitudinal, and dose-response models. These functions are meant to be used internally and not directly run by the users.

Usage

```
mixmeta.ml(Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, bscov, control, ...)

mixmeta.reml(Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, bscov, control, ...)
```

Arguments

	Assuming a meta-analysis or meta-regression based on n units aggregated within m (outer-level) groups, k outcomes, p fixed-effects predictors, and q random-effects predictors:
	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`Xlist`	a m-dimensional list of group-specific design matrices for the random-effects part of the model. Each element of this list represents a list of matrices corresponding to the (optionally multiple) grouping levels of random effects. In each matrix, rows corresponding to missing outcomes have been excluded.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.
`Slist`	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
`nalist`	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
`rep`	matrix with m rows where each column identifies the number of repetitions (number of groups) for each grouping level. The first column (outer level) is by definition a vector of 1's.
`k, q, nall`	number of outcomes, number of random-effects predictors (including the intercept), total number of observations (excluding missing), respectively. While usually all are scalars, in the case of multilevel models q can be a numeric vector representing the number of predictors for each level.
`bscov`	a character vector defining the structure of the (co)variance matrix for each level or random effects. See `mixmeta`.
`control`	list of parameters for controlling the fitting process, usually internally set to default values by `mixmeta.control`.
`...`	further arguments passed to or from other methods. Currently not used.

Details

The estimation involves kp fixed-effects coefficients and random-effects parameters, whose number depends on the number of grouping levels and, for each of them, on the chosen `(co)variance structure` for the between-study (co)variance matrices. A maximum of $kq(kq+1)/2$ parameters are needed in the case of or single-level models with unstructured form for the random-effects (co)variance matrix.

(Restricted) maximum likelihood estimators implemented in **mixmeta** rely on two iterative algorithms: `(R)IGLS` and `quasi-Newton` iterative methods. The former implements a (restricted) iterative generalized least squares method, while the latter is based on a Newton-type maximization routine using specific `likelihood functions`. The default estimation method is based on a hybrid procedure, with few runs of of the (R)IGLS algorithm and then quasi-Newton iterations until convergence. This approach is optimal in exploiting the properties of both algorithms, with (R)IGLS being robust to the choice of initial values and quick in getting near the maximum, while the quasi-Newton is fast to converge from that point. Full (R)IGLS or quasi-Newton methods can be alternatively selected using the `control` argument of `mixmeta` (see `mixmeta.control`). Follow the links above for details on each iterative algorithm.

Both estimation algorithms adopt a profiled (or concentrated) approach, where the optimization is expressed only in terms of the random-effects parameters. Cholesky and and QR decompositions are used for computational stability and efficiency, and for assuring the positive-definiteness of the estimated between-study (co)variance matrix. The method is described in details in Gasparrini and collaborators (2012) (see references below).

Value

These functions return an intermediate list object, with some components then processed and some others added later within `mixmeta.fit` and `mixmeta` to finalize an object of class `"mixmeta"`. See `mixmetaObject`.

Note

As stated earlier, these functions are called internally by `mixmeta.fit`, and are not meant to be used directly. In particular, their code does not contain any check on the arguments provided, which are expected in specific formats. The functions are not exported in the namespace, and only documented for completeness.

The arguments above are prepared by `mixmeta.fit` from its arguments `X`, `Z`, `y`, `S`, `groups`, and `bscov`. The list structure, although requiring more elaborate coding, is computationally more efficient, as it avoids the specification of sparse block-diagonal matrices, especially for meta-analysis involving a large number of studies.

Some parameters of the fitting procedures are determined by the `control` argument, with default set by `mixmeta.control`. No missing values are accepted in the fitting functions. See details on `missing values`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Pinheiro JC and Bates DM (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer Verlag.

Lindstrom MJ and Bates DM (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*. **83**(404):1014–1022.

Goldstein H (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*. **73**(1):43.

Goldstein H (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics \& Data Analysis*. **13**(1):63–71.

See Also

See `mixmeta` for the general usage of the functions. See `mixmeta.control` to determine specific parameters of the fitting procedures. Use the triple colon operator (':::') to access the code of the internal functions, such as `glsfit`. See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# REML ESTIMATOR: UNIVARIATE MODEL
mod1 <- mixmeta(yC, S=SCC, data=smoking)
summary(mod1)

# ML ESTIMATOR: MULTIVARIATE MODEL
year <- berkey98$pubyear - 1983
mod2 <- mixmeta(cbind(PD,AL) ~ year, S=berkey98[5:7], data=berkey98,method="ml")
print(summary(mod2), digits=3)
round(mod2$Psi,3)

# STRUCTURED BETWEEN-STUDY (CO)VARIANCE
y <- as.matrix(fibrinogen[2:5])
S <- as.matrix(fibrinogen[6:15])
mod3 <- mixmeta(y, S, bscov="hcs")
summary(mod3)

# MULTILEVEL MODEL
mod4 <- mixmeta(effect, var, random= ~ 1|district/study, data=school)
summary(mod4)

# LONGITUDINAL MODEL
mod5 <- mixmeta(logOR~time, S=logORvar, random=~I(time-15)|study, bscov="diag",
  method="ml", data=gliomas)
```

```
summary(mod5)
```

`mixmeta.mm` *Method of Moments Estimator for mixmeta Models*

Description

This function implements a method of moments estimator for multivariate and univariate random-effects meta-analysis and meta-regression. It is meant to be used internally and not directly run by the users.

Usage

```
mixmeta.mm(Xlist, ylist, Slist, nalist, k, m, p, nall, control, ...)
```

Arguments

	Assuming a meta-analysis or meta-regression based on m independent groups (usually studies) providing a single estimate (unit of analysis), k outcomes and p fixed-effects predictors:
	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.
`Slist`	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
`nalist`	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
`k, m, p, nall`	numeric scalars: number of outcomes, number of groups included in estimation (equal to the length of lists above), number of predictors (including the intercept), total number of observations (excluding missing).
`control`	list of parameters for controlling the fitting process, usually internally set to default values by `mixmeta.control`.
`...`	further arguments passed to or from other methods. Currently not used.

Details

In this method of moments estimator, only a basic random-effects structure is allowed, where each group (usually corresponding to an independent study) provides a single estimate (unit of analysis) for one or multiple outcomes. This implies that the number of groups (*i.e.*, the length of the lists) is identical to the number of units (`m=n`). In addition, only an unstructured form for the(co)variance matrix of the single level of random effects is permitted. Therefore, the estimation involves kp fixed-effects coefficients and $k(k+1)/2$ random-effects parameters, corresponding to the lower triangular entries of the between-study (co)variance matrix.

The method of moment estimator implemented here represents a multivariate extension of the traditional estimator proposed by DerSimonian and Laird (1986), and simplifies to the standard method in the univariate case. The estimator used here is described in Jackson and collaborators (2013) as a generalization of that developed by Chen and collaborators (2012). However, this general version is computationally more intensive, and may turn out to be slow when applied to meta-analysis of a relatively high number of studies. An alternative and computationally faster method of moment estimator was previously proposed by Jackson and collaborators (2010), although it is not invariant to reparameterization. This latter estimator is currently not implemented in **mixmeta**. See references below.

This method of moments estimator is not bounded to provide a positive semi-definite random-effects (co)variance matrix, as shown in the simulation study by Liu and colleagues (2009). Here positive semi-definiteness is forced by setting the negative eigenvalues of the estimated matrix to a positive value close to zero at each iteration (see `control`). Little is known about the impact of such constraint.

Value

This function returns an intermediate list object, with some components then processed and some others added later within `mixmeta.fit` and `mixmeta` to finalize an object of class `"mixmeta"`. See `mixmetaObject`.

Note

As stated earlier, this function is called internally by `mixmeta.fit`, and is not meant to be used directly. In particular, its code does not contain any check on the arguments provided, which are expected in specific formats. The function is however exported in the namespace and documented for completeness.

The arguments above are prepared by `mixmeta.fit` from its arguments `X`, `y` and `S`. The list structure, although requiring more elaborate coding, is computationally more efficient, as it avoids the specification of sparse block-diagonal matrices, especially for meta-analysis involving a large number of studies.

Some parameters of the fitting procedures are determined by the `control` argument, with default set by `mixmeta.control`. No missing values are accepted in the fitting functions. See details on `missing values`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Jackson D, White IR, Riley RD (2013). A matrix based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biometrical Journal*. **55**(2):231–45.

See Also

See `mixmeta` for the general usage of the functions. See `mixmeta.control` to determine specific parameters of the fitting procedures. Use the triple colon operator (':::') to access the code of the internal functions, such as `fbtr`. See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# MM ESTIMATOR: UNIVARIATE MODEL
mod1 <- mixmeta(PD ~ pubyear, S=berkey98[,5], data=berkey98, method="mm")
summary(mod1)

# MULTIVARIATE MODEL: REPRODUCE THE RESULTS IN CHEN ET AL. (2012)
S <- as.matrix(hsls[5:10])
mod2 <- mixmeta(cbind(b1,b2,b3), S, data=hsls, method="mm")
summary(mod2)

# MULTIVARIATE MODEL: REPRODUCE THE RESULTS IN JACKSON ET AL. (2013)
S <- inputcov(hyp[c("sbp_se","dbp_se")], cor=hyp$rho)
mod3 <- mixmeta(cbind(sbp,dbp), S=S, data=hyp, method="mm")
summary(mod3)
```

---

mixmeta.vc *Variance Components Estimator for mixmeta Models*

Description

This function implements a variance components estimator for multivariate and univariate random-effects meta-analysis and meta-regression. It is meant to be used internally and not directly run by the users.

Usage

```
mixmeta.vc(Xlist, ylist, Slist, nalist, k, m, p, nall, control, ...)
```

Arguments

	Assuming a meta-analysis or meta-regression based on m independent groups (usually studies) providing a single estimate (unit of analysis), k outcomes and p fixed-effects predictors:
	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.

`Slist`	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
`nalist`	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
`k, m, p, nall`	numeric scalars: number of outcomes, number of groups included in estimation (equal to the length of lists above), number of predictors (including the intercept), total number of observations (excluding missing).
`control`	list of parameters for controlling the fitting process, usually internally set to default values by `mixmeta.control`.
`...`	further arguments passed to or from other methods. Currently not used.

Details

In this variance components estimator, only a basic random-effects structure is allowed, where each group (usually corresponding to an independent study) provides a single estimate (unit of analysis) for one or multiple outcomes. This implies that the number of groups (*i.e.*, the length of the lists) is identical to the number of units (`m=n`). In addition, only an unstructured form for the(co)variance matrix of the single level of random effects is permitted. Therefore, the estimation involves kp fixed-effects coefficients and $k(k+1)/2$ random-effects parameters, corresponding to the lower triangular entries of the between-study (co)variance matrix.

The procedure is based on the estimate of the between-group (co)variance as the difference between the marginal (co)variance and the average within-group (co)variance. This in turn requires the estimate of the marginal (co)variance, obtained by the residuals of the fitted model. The procedure is iterative, with the current estimate of the between-group (co)variance plugged into a generalized least square (GLS) routine. Starting values are provided by a fixed-effects estimator (see `mixmeta.fixed`). The algorithm is fast and generally converges with few iterations.

Similar versions of this estimator has been previously proposed. Berkey and collaborators (1998) simply called it GLS method, and a non-iterative approach was proposed by Ritz and collaborators (2008), referred to as MVEE3 in their article. A non-iterative version for univariate models is discussed in Sidik and Jonkman (2007). The results from Berkey and collaborators (1998) are reproduced in the example below.

In the original approach, the estimate of the marginal (co)variance is obtained from the sum of the residual components using a denominator equal to `m-p`. Following the development proposed by Kauermann and Carroll (2001) and Fay and Graubard (2001) in the context of sandwich (co)variance estimators, then discussed by Lu and collaborators (2007), an adjusted denominator can be computed as a quantity derived from the hat matrix. This method is expected to perform better in the presence of missing values and small data sets. This alternative adjustment is chosen by default by setting `vc.adj=TRUE` in the `control` argument.

This variance component estimator is not bounded to provide a positive semi-definite between-study (co)variance matrix, as shown in the simulation study by Liu and colleagues (2009). Here positive semi-definiteness is forced by setting the negative eigenvalues of the estimated matrix to a value close to zero at each iteration (see `control`). Little is known about the impact of such constraint.

Value

This function returns an intermediate list object, with some components then processed and some others added later within `mixmeta.fit` and `mixmeta` to finalize an object of class `"mixmeta"`. See `mixmetaObject`.

Note

As stated earlier, this function is called internally by `mixmeta.fit`, and is not meant to be used directly. In particular, its code does not contain any check on the arguments provided, which are expected in specific formats. The function is however exported in the namespace and documented for completeness.

The arguments above are prepared by `mixmeta.fit` from its arguments `X`, `y` and `S`. The list structure, although requiring more elaborate coding, is computationally more efficient, as it avoids the specification of sparse block-diagonal matrices, especially for meta-analysis involving a large number of studies.

Some parameters of the fitting procedures are determined by the `control` argument, with default set by `mixmeta.control`. No missing values are accepted in the fitting functions. See details on `missing values`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Gasparrini A, Armstrong B, Kenward MG (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. **31**(29):3821–3839. [Freely available **here**].

Ritz J, Demidenko E, Spiegelman G (2008). Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference*. **139**(7):1919–1933.

Berkey CS, et al. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*. **17**(22):2537–2550.

Liu Q, et al (2009). A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose-response data. *Computational Statistics and Data Analysis*. **53**(12):4157–4167

Sidik K, Jonkman JN (2007). A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in Medicine*. **26**(9):1964–81.

See Also

See `mixmeta` for the general usage of the functions. See `mixmeta.control` to determine specific parameters of the fitting procedures. Use the triple colon operator (':::') to access the code of the internal functions, such as `sumlist`. See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# VC ESTIMATOR: UNIVARIATE MODEL
mod1 <- mixmeta(PD ~ pubyear, S=berkey98[,5], data=berkey98, method="vc")
summary(mod1)

# VC ESTIMATOR: MULTIVARIATE MODEL
mod2 <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98,
  method="vc")
summary(mod2)

# VC ESTIMATOR: NON-ITERATIVE VERSION
mod3 <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98,
  method="vc", control=list(maxiter=1))
summary(mod3)

# VARIANCE COMPONENTS ESTIMATOR: REPLICATE THE RESULTS IN BERKEY ET AL. (1998)
mod4 <- mixmeta(cbind(PD,AL) ~ I(pubyear-1983), S=berkey98[5:7], data=berkey98,
  method="vc", control=list(vc.adj=FALSE))
summary(mod4)
```

---

mixmetaCovStruct *(Co)variance Structures for mixmeta Models*

Description

Alternative options for the (co)variance structure of the random effects random effects in meta-analytical models, usually defined through the argument `bscov` of the function `mixmeta`.

Options

Assuming a meta-analysis or meta-regression based on k outcomes, for each grouping level with q random-effects predictors the matrix can be specified in various forms listed below. For multivariate models with multiple predictors, the order implies a sequence of q parameters for each k outcomes. These are the options:

- `unstr`: an unstructured form for a general positive-definite matrix. The matrix is represented by $kq(kq+1)/2$ unrestricted parameters defined as the upper triangular entries of its Cholesky decomposition.
- `diag`: a diagonal positive-definite matrix. The matrix is represented by kq unrestricted parameters defined as the logarithm of the diagonal values.
- `id`: a multiple of the identity positive-definite matrix. The matrix is represented by a single unrestricted parameter defined as the logarithm of the diagonal value.
- `cs`: a positive-definite matrix with compound symmetry structure. The matrix is represented by 2 unrestricted parameters defined as the logarithm of the identical diagonal value and the transformed correlation. The latter is parameterized so to obtain a correlation value between $-1/(kq-1)$ and 1, in order to ensure positive-definiteness.

- `hcs`: a positive-definite matrix with heterogeneous compound symmetry structure. The matrix is represented by $kq+1$ unrestricted parameters defined as the logarithm of the diagonal values and the transformed correlation. The latter is parameterized so to obtain a correlation value between $-1/(kq-1)$ and 1, in order to ensure positive-definiteness.
- `ar1`: a positive-definite matrix with autoregressive structure of first order. The matrix is represented by 2 unrestricted parameters defined as the logarithm of the identical diagonal value and the logistic transformed correlation. The latter is parameterized so to obtain a correlation value between -1 and 1.
- `har1`: a positive-definite matrix with heterogeneous autoregressive structure of first order. The matrix is represented by $kq+1$ unrestricted parameters defined as the logarithm of the diagonal value and the logistic transformed correlation. The latter is parameterized so to obtain a correlation value between -1 and 1.
- `prop`: a positive-definite matrix proportional to that provided by the user through the argument `Psifix` in the control list (see `mixmeta.control`). The matrix is represented by 1 unrestricted parameter defined as the logarithm of the multiplier.
- `cor`: a positive-definite matrix with correlation structure provided by the user through the argument `Psifix` (with `cov2cor`) in the control list (see `mixmeta.control`). The matrix is represented by k unrestricted parameters defined as the logarithm of the diagonal values.
- `fixed`: a known matrix provided by the user through the argument `Psifix` in the control list (see `mixmeta.control`). The matrix is known and no parameters are needed to represent it.

Details

Structures other than `unstr` are only available for models estimated through (restricted) maximum likelihood.

The unrestricted parameters defining the random-effects (co)variance matrix (or matrices for multilevel models) are estimated in the iterative optimization algorithm (see `mixmeta.ml`). Although rarely needed and not recommeded, the user can provided a starting value of the (co)variance matrix, from which the parameters are derived (see `mixmeta.control`).

Note

The choice of structures can affect the performance of the optimization procedure, determining forms of likelihood surfaces which induce convergence to local maxima. In particular, structures such as multiple of identity or proportional to a fixed matrix are based on strong assumptions and should be used with caution.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Pinheiro JC and Bates DM (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer Verlag.

See Also

See mixmeta. See lm or glm for standard regression functions. See mixmeta-package for an overview of this modelling framework.

Examples

```
# UNSTRUCTURED AND STRUCTURED BETWEEN-STUDY (CO)VARIANCE
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod1 <- mixmeta(y, S)
summary(mod1)
mod1$Psi

# DIAGONAL
mod2 <- mixmeta(y, S, bscov="diag")
summary(mod2)
mod2$Psi

# HETEROGENEOUS COMPOUND SYMMETRY
mod3 <- mixmeta(y, S, bscov="hcs")
summary(mod3)
mod3$Psi

# PROPORTIONAL
mod4 <- mixmeta(y, S, bscov="prop", control=list(Psifix=diag(3)+1))
summary(mod4)
mod4$Psi

# CORRELATION
Psicor <- matrix(0.2, 3, 3) ; diag(Psicor) <- 1
mod5 <- mixmeta(y, S, bscov="cor", control=list(Psifix=Psicor))
summary(mod5)
mod5$Psi
```

---

mixmetaFormula *Formulae in mixmeta Models*

Description

Meta-analytical models fitted with mixmeta are defined by specific formulae in its arguments formula and random. The formulae offer compact symbolic expressions with form y ~ x + z, where the response y in the left-hand side of the operator ~ is modelled in terms of meta-predictors x and z in the right-hand side. Terms are separated by +, and additional syntactic operators and existing functions can be used within a formula to specify transformations such as categorization and interactions, among others, as in standard formula expressions (see formula for details). The usage of formulae in **mixmeta** for the random-effects part follows closely the definition in the the **nlme** package.

Fixed-effects formula

The argument `formula` of `mixmeta` defines the fixed-effects part. Models for meta-analysis with no meta-predictors can be specified using the form `y ~ 1`, or alternatively including only the term `y` (in this case, the formula is reconstructed internally). Multivariate models can be defined by using a matrix-type `y`, with columns as multiple outcomes, or directly in the formula with form `cbind(y1 + y2) ~ 1`. In meta-regression models, other terms are added in the right-hand side of the formula as `y ~ x1 + ... + xp`, defining the linear meta-predictor. In multivariate meta-regression, the same linear predictor is specified for each outcome.

Random-effects formula or formulae

The argument `random` of `mixmeta` defines the random-effects part. When this is not specified, it is assumed that each row of data is from an independent study and assigned to a different group, as in standard meta-analytical models. If provided, this is usually represented by a one-sided formula whose basic random-intercept form is `~ 1 | g`. The term `g` at the right-hand side of the special operator `|` is the grouping factor, always required in a single random-effects formula. A more complex random-effects part can be also specified by `~ z1 + ... + zq | g`, where the terms in the left-hand side defines a linear meta-predictor, with syntax identical to the usual formulae.

The argument `random` also accepts a list of one-sided formulae referring to multiple random-effects levels in multilevel meta-analytical models. In this case, levels are assumed to be nested in the order of the list, from the lowest (outer) to the highest (inner), consistently with the grouping factors. These are usually defined by different terms in the right-hand side of the `|` operator, although in the list form can also be provided as names of the list components. This information is used internally to reconstruct the grouping structure and the random-effects design matrices. Each level can have different linear predictors, but if these are identical across levels the random-effects part can be defined by a single equation `~ z | g1 / g2`, where the special operator `/` separates the grouping factors `g2` nested in `g1`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See `mixmeta`. See `formula` for standard regression formulae. See `mixmeta-package` for an overview of this modelling framework.

Examples

```
# STANDARD RANDOM-EFFECTS META-ANALYSIS (WITH DIFFERENT SYNTAXES)
mixmeta(logor, logorvar, data=bcg)
mixmeta(logor ~ 1, logorvar, data=bcg)

# META-REGRESSION
```

```
mixmeta(logor ~ ablat, logorvar, data=bcg)

# MULTIVARIATE MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)

# NON-STANDARD MODEL: REPEATED MEASURED WITHING THE SAME GROUPS
mixmeta(effect, var, random= ~ 1|district, data=school)
mixmeta(absrisk, var, random= ~ 1|trial, data=thrombolytic)

# NON-STANDARD MODEL: MORE COMPLEX RANDOM-EFFECTS PREDICTOR
mixmeta(logOR~time, logORvar, random= ~ I(time-15)|study, data=gliomas)

# MULTILEVEL MODEL (WITH DIFFERENT SYNTAXES)
mixmeta(effect, var, random= ~ 1|district/study, data=school)
mixmeta(effect, var, random=list(~ 1|district, ~ 1|study), data=school)
mixmeta(effect, var, random=list(district = ~ 1, study = ~ 1), data=school)
```

`mixmetaObject` *mixmeta Objects*

Description

An object returned by the `mixmeta` function, inheriting from class `"mixmeta"`, and representing a fitted univariate or multivariate meta-analytical model.

Value

Objects of class `"mixmeta"` are lists with defined components. Dimensions of such components may refer to k outcome parameters, p fixed-effects and q random-effects predictors, m groups and n units used for fitting the model (the latter can be different from those originally selected due to missing). Depending on the type of meta-analytical model, the following components can bu included in a legitimate `mixmeta` object:

`coefficients`	a kp-dimensional vector of the fixed-effects coefficients.
`vcov`	estimated $kp \times kp$ (co)variance matrix of the fixed-effects coefficients.
`Psi`	the estimated $kq \times kq$ random-effects (co)variance matrix, or a list of matrices for multilevel models. Only for random-effects models.
`residuals`	a n-dimensional vector (for univariate models) or $n \times k$ matrix (for multivariate models) of residuals, that is observed minus fitted values.
`fitted.values`	a n-dimensional vector (for univariate models) or $n \times k$ matrix (for multivariate models) of fitted mean values.
`df.residual`	the residual degrees of freedom.
`rank`	the numeric rank of the fixed-effects part of the fitted model.
`logLik`	the (restricted) log-likelihood of the fitted model. Set to `NA` for non-likelihood models.

`converged, niter`
: for models with iterative estimation methods, logical scalar indicating if the algorithm eventually converged and number or iterations, respectively.

`par`
: parameters estimated in the optimization process when using likelihood-based estimators. These correspond to trasformations of entries of the random-effects (co)variance matrix, dependent on chosen `(co)variance structure`. For multilevel models, the vector includes the parameters of multiple matrices. Returned also for full (R)IGLS optimization, even if not directly used. See also the `mixmeta.ml` for details.

`hessian`
: Hessian matrix of the estimated parameters in `par` above, only returned if `hessian=TRUE` in `mixmeta.control`. See the related `optimizations algorithms` for details.

`dim`
: list with the following components: `k` (scalar, number of outcome parameters), `n` (scalar, number of units included in estimation, which could be lower than the total number in the presence of missing values), `m` (scalar, number of outer-level groups), `p` (scalar, number of fixed-effects predictors), `q` (scalar or vector,number of random-effects predictors).

`df`
: list with the following scalar components: `nall` (number of observations used for estimation, excluding missing values), `nobs` (equal to `nall`, minus the number of fixed-effects coefficients in REML models), `fixed` (number of estimated fixed-effects coefficients), `random` (number of estimated random-effects (co)variance terms).

`lab`
: list with the following label vectors: `k` for the outcome parameters, and `p` and `q` for the fixed and random-effects predictors, respectively (including intercept). The last one can be a list for multilevel models.

`S`
: a $n \times k(k+1)/2$ matrix, where each row represents the vectorized entries of the lower triangle of the related within-unit (co)variance error matrix, taken by column. See `mixmeta`.

`call`
: the function call.

`formula`
: the formula for the fixed-effects part of the model. See `mixmetaFormula`.

`model`
: the model frame used for fitting. Reported if `model=TRUE` in `mixmeta`. See `model.frame`.

`terms`
: the `terms` object representing the fixed-effects part of the fitted model.

`contrasts`
: (where relevant) the contrasts used.

`xlevels`
: (where relevant) a record of the levels of the factors used in fitting.

`na.action`
: (where relevant) information returned by `model.frame` on the special handling of NAs. See info on `missing values`.

`method`
: the estimation method.

`random`
: the formula (or list of formulae for multilevel models) for the random-effects part of the model. See `mixmetaFormula`.

`bscov`
: a string defining the random-effects (co)variance structure in likelihood based models. See `model.frame.mixmeta`.

`control`
: a list with the values of the control arguments used, as returned by `mixmeta.control`.

Methods

A number of methods functions are available for `mixmeta` objects, most of them common to other regression functions.

Specifically-written method functions are defined for `predict` (standard predictions) and `blup` (best linear unbiased predictions). The method function `simulate` produces simulated outcomes from a fitted model, while `qtest` performs the Cochran Q test for heterogeneity. Other methods have been produced for `summary`, `logLik`, `coef`, and `vcov`.

Specific methods are also available for `model.frame` and `model.matrix`. In particular, the former produces the model frame (a data frame with special attributes storing the variables used for fitting) with the additional class `"data.frame.mixmeta"`. A method `terms` is also available for extracting the terms object (only for fixed-effects or full). Methods `na.omit` and `na.exclude` for this class are useful for the handling of missing values in `mixmeta` objects.

Printing functions for the objects of classes defined above are also provided. `anova` methods for performing tests in `mixmeta` objects are in development.

All the methods above are visible (exported from the namespace) and documented. In additions, several default method functions for regression are also applicable to objects of class `"mixmeta"`, such as `fitted`, `residuals`, `AIC` and `BIC`, `drop1` and `add1`, or `update`, among others.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See `mixmeta`. See `lm` or `glm` for standard regression functions. See `mixmeta-package` for an overview of this modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL)~pubyear, S=berkey98[5:7], data=berkey98)

# INSPECT THE OBJECT
names(model)

# LABELS
model$lab

# FORMULA
model$formula

# CONVERGED?
model$converged
```

`mixmetaSim` *Simulating Responses for mixmeta Models*

Description

These functions simulate sets of responses (either univariate or multivariate) for a group of units, in terms of their mean (expected) values and within and between-group (co)variances. These sets of outcomes can be used in meta-analytical models for simulation purposes.

Usage

```
mixmetaSim(y, S, Psi, random, data, nsim=1, seed=NULL, ...)

## S3 method for class 'mixmeta'
simulate(object, nsim=1, seed=NULL, ...)
```

Arguments

	In order to simulate k outcomes for n units: a n-dimensional vector (for simulating univariate responses) or $m \times k$ matrix (for simulating multivariate responses) of mean (expected) outcomes.
`S`, `y`	series of within-unit variances (or (co)variance matrices for multivariate models) of the estimated outcome(s). For the list of accepted format, see the argument with the same name in `mixmeta`. Covariances or more complex error structures can be passed through additional arguments. See Details below.
`Psi`	the random-effects (co)variance matrix (or a list of matrices for multilevel models) of the outcomes. Dimension must be consistent with the specification of the random-effects structure in `random`.
`random`	a one-sided formula (or a list of formulae for multilevel models) offering a symbolic description of the linear predictor(s) and grouping structure for the random-effects part of the model. See the argument with the same name in `mixmeta`.
`data`	an optional data frame, list or environment (or object coercible by `as.data.frame` to a data frame), optionally containing the variables in `y`, `S`, and `random`.
`nsim`	number of simulation sets.
`seed`	an object specifying if and how the random number generator should be initialized.
`object`	an object of class `"mixmeta"`.
`...`	further optional arguments.

Details

The set(s) of responses can be simulated either from a fitted model, using the method function `simulate` for objects of class `"mixmeta"`, or directly through the function `mixmetaSim`. In the former case, the fitted values from the model are used as mean (expected) outcomes, together with the within-unit and estimated random-effects (co)variance structure. In the latter option, this information need to be provided by the user in the correct dimensions and forms.

Additional arguments can be passed in `'...'`. Specifically, arguments `Scor` and `addSlist` can be added to input missing within-unit error covariances, or to specify more complex within-unit error structures, respectively. Another argument can be `checkPD` (logical), that checks the semi-positive definiteness of the matrices). See `mixmeta.control` for details.

The functions simulate the responses for each study separately from a marginal multivariate normal distribution with mean equal to the expected values and (co)variance equal to the sum of the within-unit errors and random-effects components. The computation is identical to that implemented in the function `mvrnorm` of the package **MASS**, involving a eigen decomposition of the marginal (co)variance matrix. Numerical negative definiteness is checked, and positive semi-definiteness then forced by truncating the eigenvalues at a value close to zero (see `control`).

Value

If `nsim=1`, a matrix or vector of simulated k outcomes for the n units. If more simulation sets are required (`nsim` higher than 1), a list of matrices or vectors.

Note

Studies with missing values in the fitted values or in the components of the within (co)variances are excluded by `simulate`. Missing values are instead not accepted in `metaSim`.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See `simulate` for the general method function. See `inputcov` for inputting correlations. See `mixmeta-package` for an overview of the package and modelling framework.

Examples

```
# RUN A MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)

# SIMULATE A NEW SET OF OUTCOMES
simulate(model)
```

```
# SIMULATE FROM SCRATCH: 3 OUTCOMES, 8 STUDIES
(y <- matrix(0, 8, 3))
(S <- inputcov(matrix(runif(8*3, 0.1, 2), 8, 3, dimnames=list(NULL,
  c("V1","V2","V3"))), cor=c(0,0.5,0.7)))
(Psi <- inputcov(1:3, cor=0.3))
mixmetaSim(y, S, Psi)

# 2 SIMULATION SETS
mixmetaSim(y, S, Psi, nsim=2)
```

`ml.igls` *IGLS and RIGLS Iterative Algorithms for mixmeta Models*

Description

These functions implements (restricted) iterative generalized least squares (IGLS and RIGLS) algorithms for (restricted) maximum likelihood estimators for random-effects meta-analytical models. They are meant to be used internally and not directly run by the users.

Usage

```
ml.igls(Psi, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const, bscov,
  fix, control)

reml.rigls(Psi, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const, bscov,
  fix, control)

igls.iter(Psi, Qlist, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, bscov,
  fix, control)

rigls.iter(Psi, Qlist, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, bscov,
  fix,  control)
```

Arguments

	Assuming a meta-analysis or meta-regression based on n units aggregated within m (outer-level) groups, k outcomes, p fixed-effects predictors, and q random-effects predictors:
	a matrix (or a list of matrices for multilevel models) representing the initial estimate of the random-effects (co)variance matrix.
`Xlist`	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`Zlist`	a m-dimensional list of group-specific design matrices for the random-effects part of the model. Each element of this list represents a list of matrices corresponding to the (optionally multiple) grouping levels of random effects. In each matrix, rows corresponding to missing outcomes have been excluded.

`Qlist`	a m-dimensional list of group-specific design matrices mapping the random-effects parameters to be estimated in `Psi`. See references below for details.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.
`Slist`	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
`nalist`	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
`rep`	matrix with m rows where each column identifies the number of repetitions (number of groups) for each grouping level. The first column (outer level) is by definition a vector of 1's.
`k, q, nall`	number of outcomes, number of random-effects predictors (including the intercept), total number of observations (excluding missing), respectively. While usually all are scalars, in the case of multilevel models q can be a numeric vector representing the number of predictors for each level.
`const`	value of the constant to be included in the (restricted) likelihood, therefore not computed in the iterative algorithms.
`bscov`	a character vector defining the structure of the (co)variance matrix for each level or random effects. See `mixmeta`.
`fix`	a matrix (or optionally a list of matrices for multilevel models) defining the fixed components of the random-effects part of the model. See `mixmeta.control` for details.
`control`	list of parameters for controlling the fitting process, usually internally set to default values by `mixmeta.control`.

Details

These functions are called internally by the fitting functions `mixmeta.ml` and `mixmeta.reml` to perform (R)IGLS optimization algorithms for estimating random-effects meta-analytical models.

These estimators are not sensitive to the choice of the starting values, and quickly converge to the vicinity of the (restricted) maximum likelihood. The starting values in `Psi` are therefore defined by default as a matrix (or matrices) with a diagonal form and 0.001 variances, or otherwise selected by the user in the `control` argument of `mixmeta` (see `mixmeta.control`).

The functions `ml.igls` and `reml.rigls` first produce a design matrix that maps the entries of `Psi`, and then call `iter.igls` and `iter.rigls`, respectively, to obtain updated results at each iteration following a (R)IGLS procedure described in Goldstein and colleagues (1992). Convergence is assessed as (lack of) changes in `Psi`. Positive semi-definiteness is forced by setting the negative eigenvalues of the estimated matrix to a value close to 0 at each iteration (see `control`).

Value

The functions `ml.igls` and `reml.rigls` return an intermediate list object, with components corresponding to the estimated random-effects (co)variance matrix (or list of matrices), its parameters, the maximum (restricted) log-likelihood value, an indicator of convergence, and the number

of iterations. These are then re-processed, with other components added later within other functions to finalize an object of class "mixmeta" (see mixmetaObject). The functions iter.igls and iter.rigls return an updated version of Psi.

Note

As stated earlier, these functions are called internally by mixmeta.ml and mixmeta.reml, and are not meant to be used directly. In particular, their code does not contain any check on the arguments provided, which are expected in specific formats. They are however exported in the namespace and documented for completeness.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Goldstein H (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics \& Data Analysis*. **13**(1):63–71.

Goldstein H (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*. **73**(1):43–56.

Goldstein H (1989). Restricted unbiased iterative generalized least-squares estimation. *Biometrika*. **76**(3):622–623.

See Also

See mixmeta.fit and mixmeta.ml for additional info on the fitting procedures. See mixmeta.control to determine specific parameters of the fitting procedures. See mixmetaCovStruct for (co)variance structures. See mixmeta-package for an overview of the package and modelling framework.

---

ml.loglik.fn *Likelihood Functions for mixmeta Models*

Description

These functions compute the value of the log-likelihood and the related vectors of first partial derivatives for random-effects meta-analytical models, in terms of model parameters. They are meant to be used internally and not directly run by the users.

Usage

```
ml.loglik.fn(par, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix)
ml.loglik.gr(par, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix)

reml.loglik.fn(par, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix)
reml.loglik.gr(par, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix)
```

Arguments

	Assuming a meta-analysis or meta-regression based on n units aggregated within m (outer-level) groups, k outcomes, p fixed-effects predictors, and q random-effects predictors:
	a vector representing the random-effects parameters defining the random-effects (co)variance matrix (or multiple matrices for multilevel models).
Xlist	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
Zlist	a m-dimensional list of group-specific design matrices for the random-effects part of the model. Each element of this list represents a list of matrices corresponding to the (optionally multiple) grouping levels of random effects. In each matrix, rows corresponding to missing outcomes have been excluded.
ylist	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.
Slist	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
nalist	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
rep	matrix with m rows where each column identifies the number of repetitions (number of groups) for each grouping level. The first column (outer level) is by definition a vector of 1's.
k, q, nall	number of outcomes, number of random-effects predictors (including the intercept), total number of observations (excluding missing), respectively. While usually all are scalars, in the case of multilevel models q can be a numeric vector representing the number of predictors for each level.
const	value of the constant to be included in the (restricted) likelihood, therefore not computed in the iterative algorithms.
bscov	a character vector defining the structure of the (co)variance matrix for each level or random effects. See mixmeta.
fix	a matrix (or optionally a list of matrices for multilevel models) defining the fixed components of the random-effects part of the model. See mixmeta.control for details.

Details

These functions are called internally by fitting functions, in particular `ml.newton` and `reml.newton`, to compute the (restricted) log-likelihood and its first partial derivatives in terms of random-effects parameters for meta-analytical models.

These functions actually specify the *profiled* version of the (restricted) likelihood, expressed only in terms of random-effects parameters, while the estimate of the fixed-effects coefficients is computed at each iteration using a generalized least squares estimator, based on the current value of the between-study (co)variance matrix. At convergence, the value of this profiled version is identical to the full (restricted) likelihood. This approach is computationally efficient, as it reduces the number of parameters in the optimization routine, especially for meta-regression models.

The random-effects parameters in `par` depends on the chosen `structure(s)` for the random-effects (co)variance matrix (or multiple matrices for multilevel models). The parameterization ensures positive-definiteness. A Cholesky decomposition is then performed on the marginal (co)variance matrix in order to re-express the problem as standard least square equations, an approach which speeds up the computation of matrix inverses and determinants. These equations are finally solved through a QR decomposition, which guarantees stability. More details are provided in the references below.

Some parameters of the fitting procedures are determined through `mixmeta.control`. Specifically, the user can obtain the Hessian matrix of the estimated parameters (appropriately transformed, see `mixmetaCovStruct`) in the optimization function by setting `hessian=TRUE`, and specific settings of the optimization process can be defined by the control list argument `optim`. These values are passed to the optimization function `optim`.

Value

`ml.loglik.fn` and `reml.loglik.fn` return the value of the (restricted) log-likelihood for a given set of parameters in `par`. `ml.loglik.gr` and `reml.loglik.gr` return instead the related vector of first partial derivatives.

Note

As stated earlier, these functions are called internally by `mixmeta.ml` and `mixmeta.reml`, and are not meant to be used directly. In particular, their code does not contain any check on the arguments provided, which are expected in specific formats. They are however exported in the namespace and documented for completeness.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Lindstrom MJ and Bates DM (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*. **83**(404):1014–1022.

Harville DA (1977) Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*. **72**(358):320–338.

Pinheiro JC and Bates DM (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer Verlag.

See Also

See `mixmeta.fit` and `mixmeta.ml` for additional info on the fitting procedures. See `mixmeta.control` to determine specific parameters of the fitting procedures. See `mixmetaCovStruct` for (co)variance structures. See `chol` and `qr` for info on the Cholesky and QR decomposition. See `mixmeta-package` for an overview of the package and modelling framework.

`ml.newton` *Quasi-Newton Iterative Algorithms for mixmeta Models*

Description

These functions implement quasi-Newton iterative algorithms for (restricted) maximum likelihood estimators for random-effects meta-analytical models. They are meant to be used internally and not directly run by the users.

Usage

```
ml.newton(Psi, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix, control)

reml.newton(Psi, Xlist, Zlist, ylist, Slist, nalist, rep, k, q, nall, const,
  bscov, fix, control)
```

Arguments

	Assuming a meta-analysis or meta-regression based on n units aggregated within m (outer-level) groups, k outcomes, p fixed-effects predictors, and q random-effects predictors:
	a matrix (or a list of matrices for multilevel models) representing the initial estimate of the random-effects (co)variance matrix.
`Xlist`	a m-dimensional list of group-specific design matrices for the fixed-effects part of the model. Rows corresponding to missing outcomes have been excluded.
`Zlist`	a m-dimensional list of group-specific design matrices for the random-effects part of the model. Each element of this list represents a list of matrices corresponding to the (optionally multiple) grouping levels of random effects. In each matrix, rows corresponding to missing outcomes have been excluded.
`ylist`	a m-dimensional list of group-specific vectors of estimated outcomes. Entries corresponding to missing outcomes have been excluded.

Slist	a m-dimensional list of within-group (co)variance matrices of estimated outcomes. Rows and columns corresponding to missing outcomes have been excluded.
nalist	a m-dimensional list of group-specific logical vectors, identifying missing outcomes.
rep	matrix with m rows where each column identifies the number of repetitions (number of groups) for each grouping level. The first column (outer level) is by definition a vector of 1's.
k, q, nall	number of outcomes, number of random-effects predictors (including the intercept), total number of observations (excluding missing), respectively. While usually all are scalars, in the case of multilevel models q can be a numeric vector representing the number of predictors for each level.
const	value of the constant to be included in the (restricted) likelihood, therefore not computed in the iterative algorithms.
bscov	a character vector defining the structure of the (co)variance matrix for each level or random effects. See mixmeta.
fix	a matrix (or optionally a list of matrices for multilevel models) defining the fixed components of the random-effects part of the model. See mixmeta.control for details.
control	list of parameters for controlling the fitting process, usually internally set to default values by mixmeta.control.

Details

These functions are called internally by the fitting functions mixmeta.ml and mixmeta.reml to perform quasi-Newton iterative optimization algorithms for estimating random-effects meta-analytical models.

Starting values for the iterations are defined by Psi, representing a random-effects (co)variance matrix (or a list of matrices for multilevel models). In the default hybrid procedure (see mixmeta.ml), these are provided using few iterations of a (R)IGLS algorithm. If a full quasi-Newton method is used, the starting values are instead defined by default as a matrix (or matrices) with a diagonal form and 0.001 variances, or otherwise selected by the user in the control argument of mixmeta (see mixmeta.control).

The functions first re-define Psi as a set of random-effects parameters, depending on the chosen structure(s), using parameterizations that ensure the positive-definiteness of the estimated matrix (or matrices). Then, the function optim with method="BFGS" is called internally to perform the quasi-Newton optimization, using specific likelihood functions that compute the value of the (restricted) likelihood and (optionally) the vector of its first partial derivatives. The latter are used only in the case of basic random-effects structures, or otherwise the derivatives are computed numerically.

Some parameters of the optimization procedures are determined through mixmeta.control. Specifically, the user can obtain the Hessian matrix of the estimated parameters (appropriately transformed, see mixmetaCovStruct) in the optimization function by setting hessian=TRUE, and specific settings of the optimization process can be defined by the control list argument optim. These values are passed to the optimization function optim.

Value

These functions return an intermediate list object, with components corresponding to the estimated random-effects (co)variance matrix (or list of matrices), the maximum (restricted) log-likelihood value, an indicator of convergence, the number of iterations, and optionally the Hessian matrix. These are then re-processed, with other components added later within other functions to finalize an object of class `"mixmeta"`. See `mixmetaObject`.

Note

As stated earlier, these functions are called internally by `mixmeta.ml` and `mixmeta.reml`, and are not meant to be used directly. In particular, their code does not contain any check on the arguments provided, which are expected in specific formats. They are however exported in the namespace and documented for completeness.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Lindstrom MJ and Bates DM (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*. **83**(404):1014–1022.

Harville DA (1977) Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*. **72**(358):320–338.

Pinheiro JC and Bates DM (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer Verlag.

See Also

See `mixmeta.fit` and `mixmeta.ml` for additional info on the fitting procedures. See `mixmeta.control` to determine specific parameters of the fitting procedures. See `mixmetaCovStruct` for (co)variance structures. See `chol` and `qr` for info on the Cholesky and QR decomposition. See `mixmeta-package` for an overview of the package and modelling framework.

`model.frame.mixmeta` *Extract Model Frame and Design Matrix from mixmeta Objects*

Description

These method functions return the model frame and design matrix for meta-analytical models represented in objects of class `"mixmeta"`.

Usage

```
## S3 method for class 'mixmeta'
model.frame(formula, ...)

## S3 method for class 'mixmeta'
model.matrix(object, ...)
```

Arguments

`object, formula`	an object of class `"mixmeta"`.
`...`	further arguments passed to or from other methods.

Details

The model frame is produced by `mixmeta` when fitting the meta-analytical model, and stored in the `mixmeta` object if argument `model=TRUE`. Alternatively, the model frame is directly returned from a call to `mixmeta` with argument `method="model.frame"`. The method function `model.frame` simply extracts the saved model frame if available, or otherwise evaluates a call to `mixmeta` when `method="model.frame"`.

The method function `model.matrix` extracts the design matrix for the fixed-effects part of a fitted meta-analytical model. It first extract the model frame by calling `model.frame`, and then passes the call to the default method.

Note that the model frame of `mixmeta` models consist of terms for both the fixed and random-effects parts, the latter including also the grouping factors. This information can be used to reconstruct the proper model frame or matrix for each part.

These methods functions are similar to those provided for regression objects `lm` and `lm`.

Value

For `model.frame`, a data.frame with special attributes (see the default method `model.frame`) and the additional class `"data.frame.mixmeta"`.

For `model.matrix`, the design matrix used to fit the model.

Note

The reason why these specific method functions are made available for class `mixmeta`, and in particular why a new class `"data.frame.mixmeta"` has been defined for model frames, lies in the special handling of missing values in multivariate meta-analysis models fitted with `mixmeta`. Methods `na.omit` and `na.exclude` for class `"data.frame.mixmeta"` are useful for properly accounting for missing values when fitting these models.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See the default methods model.frame and model.matrix. See na.omit and na.exclude on the handling of missing values. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL AND SUMMARIZE THE RESULTS
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98,
  method="ml")

# MODEL FRAME
model$model
model.frame(model)
update(model, method="model.frame")
class(model.frame(model))

# MODEL MATRIX
model.matrix(model)
```

---

na.omit.data.frame.mixmeta

Handling Missing Values in mixmeta Models

Description

These method functions exclude rows corresponding to units with invalid missing pattern from model frames of class "data.frame.mixmeta". This guarantees the correct handling of missing values while fitting meta-analytical models.

Usage

```
## S3 method for class 'data.frame.mixmeta'
na.omit(object, ...)

## S3 method for class 'data.frame.mixmeta'
na.exclude(object, ...)
```

Arguments

object	an object of class "data.frame.mixmeta".
...	further arguments passed to or from other methods.

Details

A model frame of class "data.frame.mixmeta" is produced by mixmeta. A call to na.omit or na.exclude removes from the model frame the rows corresponding to studies with invalid missing pattern. In addition, a na.action attribute is added to the model frame, namely a numeric vector corresponding to the removed rows and class "omit" or "exclude", respectively. This information is used by naresid and napredict to deal with missing values in functions such as fitted, residuals, predict and blup, among others.

The definition of missing, identifying an invalid missing pattern, is different in meta-analytical models performed through mixmeta if compared to other regression functions such as lm or glm, in particular for the multivariate case. Specifically, while a unit is removed if at least an observation for one predictor is missing, partially missing outcomes do not prevent the unit to contribute to estimation (see mixmeta). Specific methods na.omit and na.exclude for class "data.frame.mixmeta" allow this different definition.

Value

These functions returns the model frame object with rows corresponding to units with invalid missing pattern being removed. They also add the related na.action attribute as explained above.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See na.action, naresid and napredict. See model.frame. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# INPUT MISSING VALUES IN PREDICTOR AND ONE RESPONSE
data <- berkey98
data[2,1] <- data[4,3] <- NA
data

# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=data[5:7], data=data, method="ml")

# SUMMARIZE: NOTE THE NUMBER OF STUDIES AND OBSERVATIONS
summary(model)
df.residual(model)

# EXTRACT THE MODEL FRAME WITH na.pass
model.frame(model, na.action="na.pass")
# EXTRACT THE MODEL FRAME WITH na.omit (DEFAULT)
model.frame(model, na.action="na.omit")

# COMPARE WITH DEFAULT METHOD FOR na.omit
frame <- model.frame(model, na.action="na.pass")
na.omit(frame)
```

```
class(frame)
class(frame) <- "data.frame"
na.omit(frame)

# WITH na.exclude
residuals(model)
residuals(update(model, na.action="na.exclude"))
```

---

p53 *Mutant p53 Gene and Squamous Cell Carcinoma*

Description

The dataset includes studies providing evidence about whether the presence of mutant p53 tumour suppressor gene is a prognostic factor for patients presenting with squamous cell carcinoma arising from the oropharynx cavity. Unadjusted estimates of log hazard ratios of mutant p53 to normal p53 for disease-free and overall survival, together with the associated variances, are collected from 6 observational studies.

Usage

```
p53
```

Format

A data frame with 6 observations on the following 5 variables:

- `study`: study ID.
- `y1,V1`: estimate and associated variance of the log hazard ratio for disease-free survival.
- `y2,V2`: estimate and associated variance of the log hazard ratio for overall survival.

Details

Only 3 studies provide estimates for disease-free survival. The within-study correlations are not reported in the original studies but are expected to be highly positively correlated. The original data are described in Tandon and colleagues (2010) and used as an example by Jackson and colleagues (2011).

Note

The data provide an example of application of multivariate meta-analysis when the within-study correlations are not known. These correlations can be inputted directly in the `mixmeta` function through the `control` argument. See `mixmeta.control` for details.

Source

Jackson D, Riley R, White IR (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*. **30**(20);2481–2498.

Tandon S, Tudur-Smith C, Riley RD, et al. (2010). A systematic review of p53 as a prognostic factor of survival in squamous cell carcinoma of the four main anatomical subsites of the head and neck. *Cancer Epidemiology, Biomarkers and Prevention*. **19**(2):574–587.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS OF EXAMPLE 3 IN JACKSON ET AL. (2011)

# INSPECT THE DATA
p53

# REML MODEL WITH INPUTTED CORRELATION EQUAL TO 0.95
model <- mixmeta(cbind(y1,y2), cbind(V1,V2), data=p53, control=list(Scor=0.95))
print(summary(model), digits=2)
```

`predict.mixmeta` *Predicted Values from mixmeta Models*

Description

This method function computes predictions from fitted univariate or multivariate meta-analytical models represented in objects of class `"mixmeta"`, optionally for a new set of predictor values in meta-regression models. Predictions are optionally accompanied by standard errors, confidence intervals or the entire (co)variance matrix of the predicted outcomes.

Usage

```
## S3 method for class 'mixmeta'
predict(object, newdata, se=FALSE, ci=FALSE, vcov=FALSE, ci.level=0.95,
  format, aggregate="stat", na.action=na.pass, ...)
```

Arguments

`object`	an object of class `"mixmeta"`.
`newdata`	An optional data frame in which to look for variables values with which to predict from meta-regression models.
`se`	logical switch indicating if standard errors must be included.
`ci`	logical switch indicating if confidence intervals must be included.
`vcov`	logical switch indicating if the (co)variance matrix must be included.
`ci.level`	a numerical value between 0 and 1, specifying the confidence level for the computation of confidence intervals.

format	the format for the returned results. See Value.
aggregate	when format="matrix" and se or ci are required, the results may be aggregated by statistic or by outcome. See Value
na.action	a function which indicates what should happen when the data contain NAs. The default to the value saved in object. See Note.
...	further arguments passed to or from other methods.

Details

The method function predict produces predicted values from mixmeta objects, obtained by evaluating the original call to mixmeta in the frame newdata. For both fixed and random-effects models, estimated predictions are only based on the fixed part of the model, ignoring study-specific deviations, differently from blup.

If newdata is omitted, the predictions are based on the data used for the fit. In that case how to handle predictions for units removed from estimation due to invalid missing pattern is determined by the na.action argument used in mixmeta to produce object. If na.action=na.omit, units excluded from estimation will not appear, whereas if na.action=na.exclude they will appear, with values set to NA for all the outcomes. This step is performed by napredict. See Notes.

Value

The results may be aggregated in matrices (the default), or returned as lists, depending on the argument format. For multivariate models, the aggregation is ruled by the argument aggregate, and the results may be grouped by statistic or by outcome. If vcov=TRUE, lists are always returned.

Note

The definition of missing in model frames used for estimation in mixmeta is different than that commonly adopted in other regression models such as lm or glm. See info on missing values in mixmeta.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See blup for best linear unbiased predictions. See the default method predict. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)

# PREDICTED FROM YEAR 1985 TO 1987, WITH LABELS
newdata <- data.frame(pubyear=1985:1987, row.names=1985:1987)

# AVERAGED OUTOCOMES AND SE
predict(model, newdata, se=TRUE)

# SAME AS ABOVE, AGGREGATED BY OUTCOME
predict(model, newdata, se=TRUE, aggregate="outcome")

# WITH VCOV, FORCED TO A LIST
predict(model, newdata, se=TRUE, vcov=TRUE, aggregate="outcome")
```

---

qtest *Cochran Q Test of Heterogeneity*

Description

This is a generic function to perform a Cochran Q test of (residual) heterogeneity. The function invokes particular `methods` which depend on the `class` of the first argument. Currently, specific methods exist for several meta-analytical models in various packages: `qtest.mixmeta`, `qtest.mvmeta`, and `qtest.dosresmeta`.

Usage

```
qtest(object, ...)
```

Arguments

`object`	an object for which the test is desired
`...`	further arguments passed to specific methods.

Details

The test assesses the null hypothesis that the variability in the distribution of the outcomes is explained only in terms of within-unit estimation errors. This corresponds to a test on the hypotesis that there is no variation attributable to random-effects terms.

Value

Returned values depend on the specific class. Usually, the results of the test.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

References

Cochran WG (1950). The comparison of percentages in matched samples". *Biometrika*. **37**(3/4):256–266.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

Specific methods for various classes: `qtest.mixmeta`, `qtest.mvmeta`, and `qtest.dosresmeta`.

`qtest.mixmeta` *Cochran Q Test of Heterogeneity for mixmeta Models*

Description

This method function performs a Cochran Q test of (residual) heterogeneity on fitted meta-analytical models represented in objects of class `"mixmeta"`.

Usage

```
## S3 method for class 'mixmeta'
qtest(object, ...)

## S3 method for class 'qtest.mixmeta'
print(x, digits=3, ...)
```

Arguments

`object, x`	objects of classes `"mixmeta"` and `"qtest.mixmeta"`, respectively.
`digits`	an integer specifying the number of digits to which printed results must be rounded.
`...`	further arguments passed to or from other methods.

Details

The test assesses the null hypothesis that the variability in the distribution of the outcomes is explained only in terms of estimation error in each unit, measured by the within-unit (co)variance matrices stored in the component `S` of `mixmeta` objects. This is equal to test the hypothesis that the random-effects (co)variance matrix (or all matrices in multilevel models) is a zero matrix, and there is no random deviation in unit-specific estimates. For multivariate models, tests for single outcome parameters, comparable to estimates from multiple univariate meta-analysis, are also reported. This test reduces to the standard Q test in univariate single-level models.

The function compute the statistics by actually fitting the related fixed-effects model, re-evaluating the call of the model with `method` changed to `"fixed"`.

Value

A list object of class "qtest.mixmeta" with the following components:

Q	the vector of test statistics for overall and outcome-specific tests, distributed under the null hypothesis as a Chi-square with degrees of freedom df.
df	the vector of degrees of freedom of the null distribution for overall and outcome-specific tests. For the overall test, equal to the number of observations used for estimation minus the number of coefficients in the fixed part of the model. For outcome-specific test, equal to number of observed values minus the number of coefficients.
pvalue	the vector of p-values for overall and outcome-specific tests.
residual	logical switch indicating if a meta-regression model is assessed, meaning that the tested heterogeneity is residual.
k	dimensionality of the overall test, that is the number of outcome parameters in the model.

As usual, the print method function for class "qtest.mixmeta" does not return any value.

Note

In multivariate models, tests on single outcome parameters are performed by extracting the related estimates and variances, but they do not account for the correlation between them, which nevertheless has been considered in estimation. These tests are not therefore comparable with those performed by running a univariate model on each outcome parameter.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Cochran WG (1950). The comparison of percentages in matched samples". *Biometrika*. **37**(3/4):256–266.

See Also

See qtest for the generic method function. See mixmeta-package for an overview of the package and modelling framework.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ 1, S=berkey98[5:7], data=berkey98)

# MULTIVARIATE COCHRAN Q TEST FOR HETEROGENEITY
test <- qtest(model)
```

```
print(test, digits=2)
unclass(test)
```

`school` *Studies on Modified School Calendar and Student Achievement*

Description

The dataset contains the results of 56 studies that evaluate the effect of a modified school calendar on student achievement. The studies assessed students from grade 1 to 9 and reported standardized reading achievement differences between schools that follow a year-round versus the traditional nine-month calendar. The studies were performed in separate school districts, with at least three studies in each district.

Usage

```
school
```

Format

A data frame with 56 observations on the following 5 variables:

- `district,study`: numbers identifying the school district and study, respectively.
- `effect`: estimated standardized effect, reported as difference in reading achievement expressed in standard deviation units.
- `var`: within-study variance of the estimated effects.
- `year`: year when the study was performed.

Note

The data provide an example of application of multilevel meta-analysis with multiple nested random-effects levels, where effect sizes are correlated between studies within school district. This more complex correlation structure is modelled by two levels of random effects. Results can be compared with the so-called three-level model in Kostantopoulos (2011), that is defined as a two-level meta-analysis here.

Source

Kostantopoulos S (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods*. **2**(1):61–76.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN KOSTANTOPOULOS (2011), TABLES 4 AND 5

# STANDARD META-ANALYSIS (NB: random NOT STRICTLY NEEDED HERE)
mod1 <- mixmeta(effect, var, random= ~ 1|study, data=school, method="ml")
print(summary(mod1), digits=3, report="var")

# STANDARD META-REGRESSION
yearcen <- school$year - mean(school$year)
mod2 <- mixmeta(effect ~ yearcen, var, random= ~ 1|study, data=school,
  method="ml")
print(summary(mod2), digits=3, report="var")

# TWO-LEVEL META-ANALYSIS
mod3 <- mixmeta(effect, var, random= ~ 1|district/study, data=school,
  method="ml")
print(summary(mod3), digits=3, report="var")

# TWO-LEVEL META-REGRESSION
yearcen2 <- with(school, year - mean(tapply(year, district, mean)))
mod4 <- mixmeta(effect ~ yearcen2, var, random= ~ 1|district/study, data=school,
  method="ml")
print(summary(mod4), digits=3, report="var")

### SEE help(thrombolytic) FOR A COMPLEMENTARY EXAMPLE
```

`smoking` *Meta-Analysis of Interventions to Promote Smoking Cessation*

Description

The dataset contains the results of 24 trials comparing four alternative interventions to promote smoking cessation. The trials have different designs, comparing two or three different interventions. The data consist of the number of successes out of the total participants, and the estimated log-odds ratio for arms B (self-help), C (individual counselling), and D (group counselling) relative to arm A (no contact), as well as the (co)variance matrix of these three estimates.

Usage

```
smoking
```

Format

A data frame with 24 observations on the following 19 variables:

- `study`: study ID.
- `design`: design of the trial, reporting the interventions being compared.
- `dA,dB,dC,dD`: number of successes for each intervention.

- nA,nB,nC,nD: number of participants for each intervention.
- yB,yC,yD: estimated log-odds ratios for interventions B, C and D versus intervention A.
- SBB,SBC,SBD,SCC,SCD,SDD: variances and co-variances of the estimated log-odds ratios for interventions B, C and D versus intervention A. The order corresponds to the lower triangular elements of the (co)variance matrix taken by column.

Details

Intervention A is chosen as the reference category. Trials without an arm A (trials 2 and 21-24) are augmented with an arm A with 0.01 individuals and 0.001 successes. Trials containing zero cells (trials 9 and 20) have 1 individual with 0.5 successes added to each intervention. Details on the data augmentation and estimation of (co)variances of the log-odds ratios are provided by White (2011).

Note

The data provide an example of application of network meta-analysis, also referred to as indirect mixed-treatment comparison. Additional information using examples based on these data are provided by Lu and Ades (2006), White (2011) and Higgins and colleagues (2012).

Source

Lu G and Ades AE (2006). Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*. **101**:447–459.

Higgins JPT, et al. (2012). Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods*. **3**(2):98–110.

White IR (2011). Multivariate random-effects meta-regression. *The Stata Journal*. **11**:255–270.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN WHITE (2011)

# INSPECT THE DATA
head(smoking)
names(smoking)

# CONSISTENCY MODEL, UNSTRUCTURED BETWEEN-STUDY (CO)VARIANCE
y <- as.matrix(smoking[11:13])
S <- as.matrix(smoking[14:19])
mod1 <- mixmeta(y, S)
summary(mod1)

# CONSISTENCY MODEL, STRUCTURED BETWEEN-STUDY (CO)VARIANCE (PROPORTIONAL)
mod2 <- mixmeta(y, S, bscov="prop", control=list(Psifix=diag(3)+1))
summary(mod2)

# TRANSFORM IN LONG FORMAT, WITH S AS LIST (EXCLUDING MISSING)
long <- na.omit(reshape(smoking[,c(1,2,11:13)], varying=list(3:5), idvar="study",
```

```
  v.names="y", timevar="outcome", times=colnames(y), direction="long"))
Slist <- lapply(lapply(seq(nrow(S)), function(i) xpndMat(S[i,])), function(x)
  x[!is.na(diag(x)), !is.na(diag(x)), drop=FALSE])

# THE MODELS ABOVE CAN BE REPLICATED IN THE LONG FORMAT
mod2b <- mixmeta(y ~ 0 + factor(outcome), random= ~ 0 + factor(outcome)|study,
  data=long, bscov="prop", control=list(addS=Slist, Psifix=diag(3)+1))
summary(mod2b)

# DEFINE AND ADD INDICATORS FOR OUTCOME AND DESIGN
dummy <- cbind(model.matrix(~0+outcome, long), model.matrix(~0+design, long))
colnames(dummy) <- c(levels(factor(long$outcome)), levels(long$design))
long <- cbind(long, data.frame(dummy))

# INCONSISTENCY MODEL (SPECIAL PARAMETERIZATION OF OUTCOME-BY-DESIGN INTERACTION)
formula <- y ~ 0 + yB + yC + yC:acd + yC:bc + yC:bcd + yD + yD:acd + yD:bcd +
  yD:bd + yD:cd
mod3 <- update(mod2b, formula=formula)
summary(mod3)
```

`summary.mixmeta` *Summarizing mixmeta Models*

Description

Print and summary method functions for fitted meta-analytical models represented in objects of class "mixmeta".

Usage

```
## S3 method for class 'mixmeta'
summary(object, ci.level=0.95, ...)

## S3 method for class 'summary.mixmeta'
print(x, digits=4, report=c("sd","var"), ...)

## S3 method for class 'mixmeta'
print(x, digits=4, ...)
```

Arguments

`object`	an object of class "mixmeta" produced by a call to mixmeta.
`x`	an object of class "mixmeta" or "summary.mixmeta", produced by calls to mixmeta or summary.mixmeta, respectively.
`ci.level`	a numerical value between 0 and 1, specifying the confidence level for the computation of confidence intervals.
`digits`	an integer specifying the number of digits to which printed results must be rounded.

report	if standard deviations (sd) or variances (var) must be reported for summarizing the random-effects (co)variance structure.
...	further arguments passed to or from other methods.

Details

The print method function for class "mixmeta" only returns basic information on the fitted model, namely the call, estimated fixed-effects coefficients, dimensions and fit statistics (log-likelihood, AIC, BIC).

The summary method function computes additional statistics and tests, and produces a list object of class "summary.mixmeta". The print method function for this class shows additional information, such as tables reporting the estimates for the fixed and random-effects parts of the model, Cochran Q test for heterogeneity and I-squared.

Value

The summary method function for mixmeta objects produces a list of class "summary.mixmeta". The components of the lists are some of those stored in the related mixmeta object, plus the following:

coefficients	a matrix reporting point estimates, standard errors, z statistics and related p-values of the test, and confidence intervals for the kp fixed-effects coefficients. Note this is different than the component with the same name stored in mixmeta objects, simply reporting the point estimates (see mixmetaObject).
AIC	the value of the Akaike information criterion for the fitted mixmeta model, obtained through a call to AIC.
BIC	the value of the Bayesian information criterion for the fitted mixmeta model, obtained through a call to BIC.
corFixed	the $kp \times kp$ correlation matrix of the fixed-effects coefficients, obtained from the (co)variance matrix vcov (see mixmetaObject and vcov).
corRandom	the $kq \times kq$ correlation matrix of the random effects, obtained from the random-effects (co)variance matrix Psi, or a list of multiple matrices for multilevel models. See mixmetaObject.
qstat	results from the Cochran Q test for heterogeneity, namely a list corresponding to a qtest.mixmeta object without its class, obtained through qtest.
i2stat	I-squared statistic for the meta-analytical model.
ci.level	the confidence level used for defining the confidence intervals for the estimates of the fixed-effects coefficients.

As usual, the print method functions for classes "mixmeta" and "summary.mixmeta" do not return any value.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>> and Francesco Sera <<francesco.sera@lshtm.ac.uk>>

References

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

See Also

See `mixmeta` and `mixmetaObject`.

Examples

```
# RUN THE MODEL
model <- mixmeta(cbind(PD,AL) ~ pubyear, S=berkey98[5:7], data=berkey98)

# SIMPLE PRINT
model

# DEFINE DIGITS
print(model, digit=2)

# SUMMARY WITH 80TH CONFIDENCE INTERVALS
summary(model, ci.level=0.80)

# REPORT RANDOM EFFECTS IN TERMS OF VARIANCES (USE print)
print(summary(model), report="var")
```

---

terms.mixmeta *Extract Model Terms from mixmeta Objects*

Description

These method function returns the terms object that defines meta-analytical models represented in objects of class `"mixmeta"`.

Usage

```
## S3 method for class 'mixmeta'
terms(x, type="fixed", ...)
```

Arguments

`x`	an object of class `"mixmeta"`.
`type`	the type of terms. Either `"fixed"` or `"full"`. See Details.
`...`	further arguments passed to or from other methods.

Details

The terms object is produced by mixmeta when fitting the meta-analytical model, and stored as an attribute of the model.frame. Note that this object consists of terms for both the fixed and random-effects parts, the latter including also the grouping factors.

By using the default type="fixed", this method function removes the random-effects terms. This can then be used, for instance, for creating the model.matrix for the fixed effects. Otherwise with type="full", the full set of terms is returned.

Value

An object of class c("terms","formula") which contains the terms representation of a symbolic meta-analytical model. See terms.object for its structure.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See the methods model.frame and model.matrix.

Examples

```
# RUN A MODEL
model <- mixmeta(effect, var, random= ~ 1|district/study, data=school)

# TERMS (FIXED AND FULL)
terms(model)
terms(model, "full")
attr(model.frame(model), "terms")
```

---

thrombolytic *Randomized Trials of Thrombolytic Therapy*

Description

The dataset contains the data on 20 randomized trials of thrombolytic therapy, which evaluated effect on short-term mortality after a myocardial infarction (up to 35 days) in 50,246 patients in relation to treatment delay. The hypothesis is that the thrombolytic therapy reduces the mortality risk following the myocardial infarction, and that the benefit is particularly substantial for very early treatment. Some of the trials report separate results according to treatment delay, generating 38 observations from full trials or subgroups of trials. Effect sizes were reported as absolute risk reduction computed as the difference between treated and control groups in each trial or subgroup.

Usage

```
thrombolytic
```

Format

A data frame with 38 observations on the following 10 variables:

- `trial`: label identifying the trial.
- `time2treat`: treatment delay after the onset of the symptoms of a myocardial infaction, reported in hours.
- `dtreat,ntreat`: number of deaths and total patients in the treated group, respectively.
- `dcontr,ncontr`: number of deaths and total patients in the control group, respectively.
- `risktreat,riskcontr`: risk of death in the treatment and control groups, respectively.
- `absrisk`: absolute risk difference of death between the treatment and control groups. See Details.
- `var`: variance of the absolute risk difference. See Details.

Details

The absolute risk is simply the difference in risk, which is computed empirically as ratio of the number of deaths and the number of total patients in treated and control groups ($p_1 = d_1/N_1$ and $p_0 = d_0/N_0$, respectively). The variance of the absolute risk difference is computed as $p_0(1 - p_0)/N_0 + p_1(1 - p_1)/N_1$. See Thompson and colleagues (2001) for details.

Note

The data provide an example of application of multilevel meta-analysis with repeated observations in an inner level within an outer level, corresponding here to treatment subgroups within each trial. This more complex correlation structure is modelled by two levels of random effects, including meta-predictors that can explain part of the heterogeneity at each level. Results can be compared with those reported by Thompson and colleagues (2001).

Source

Thompson SG, Turner RM, Warn DE (2001). Multilevel models for meta-analysis, and their application to absolute risk differences. *Statistical Methods in Medical Research*. **10**(6):375–392.

Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). An extended mixed-effects framework for meta-analysis.*Statistics in Medicine*. 2019;38(29):5429-5444. [Freely available **here**].

Examples

```
### REPRODUCE THE RESULTS IN THOMPSON ET AL (2001), TABLES 2, 3, AND 4

# STANDARD FIXED-EFFECTS META-ANALYSIS
mod1 <- mixmeta(absrisk, var, data=thrombolytic, method="fixed")
print(summary(mod1), digits=5)

# STANDARD RANDOM-EFFECTS META-ANALYSIS
subtrial <- seq(nrow(thrombolytic))
mod2 <- mixmeta(absrisk, var, random= ~ 1|subtrial, data=thrombolytic)
print(summary(mod2), digits=5)
```

```
# TWO-LEVEL RANDOM-EFFECTS META-ANALYSIS
mod3 <- mixmeta(absrisk, var, random= ~ 1|trial/subtrial, data=thrombolytic)
print(summary(mod3), digits=5)

# TWO-LEVEL RANDOM-EFFECTS META-REGRESSION
mod4 <- mixmeta(absrisk~time2treat, var, random= ~ 1|trial/subtrial,
  data=thrombolytic)
print(summary(mod4), digits=5)

# TWO-LEVEL RANDOM-EFFECTS META-REGRESSION WITH NON-LINEAR TERM
mod5 <- mixmeta(absrisk ~ time2treat + I(1/time2treat), var,
  random= ~ 1|trial/subtrial, data=thrombolytic)
print(summary(mod5), digits=5)

### SEE help(school) FOR A COMPLEMENTARY EXAMPLE
```

---

vechMat *Vectorization and Expansion of Symmetric Matrices*

Description

The function vechMat transforms a symmetric matrix in a vector containing its lower triangular elements, taken by column. The function xpndMat reverses this transformation.

Usage

```
vechMat(mat, diag=TRUE)

xpndMat(vech)
```

Arguments

mat	a square matrix.
vech	a vector.
diag	a logical switch indicating if the diagonal entries must be included.

Value

A vector for vechMat, a symmetric matrix for xnpdMat.

Author(s)

Antonio Gasparrini <<antonio.gasparrini@lshtm.ac.uk>>

See Also

See functions vech and xpnd in package **MCMCpack**.

Examples

```
# GENERATE A POSITIVE-DEFINITE MATRIX, VECTORIZE IT AND THEN RE-EXPAND
(M <- crossprod(matrix(rnorm(9),3)))
(v <- vechMat(M))
xpndMat(v)

# EXTRACT VECTORIZED S, EXPAND TO A LIST, AND RE-VECTORIZE
(S <- as.matrix(berkey98[5:7]))
(Slist <- lapply(seq(nrow(S)), function(i) xpndMat(S[i,])))
t(sapply(Slist,vechMat))
```

Index