# BMJ Open

# Protocol of an individual participant data meta-analysis to quantify the impact of high ambient temperatures on maternal and child health in Africa (HE$^2$AT IPD)

Darshnika Pemi Lakhoo ![ORCID],[1] Matthew Francis Chersich ![ORCID],[1] Chris Jack,[2] Gloria Maimela ![ORCID],[1] Guéladio Cissé ![ORCID],[3] Ijeoma Solarin ![ORCID],[1] Kristie L Ebi,[4] Kshama S Chande,[1] Cherlynn Dumbura,[5] Prestige Tatenda Makanga ![ORCID],[5,6] Lisa van Aardenne ![ORCID],[2] Bonnie R Joubert,[7] Kimberly A McAllister,[7] Maliha Ilias,[8] Sibusisiwe Makhanya,[9] Stanley Luchters ![ORCID],[5,10,11] HE2AT Center IPD Study Group

For numbered affiliations see end of article.

**Correspondence to**
Professor Stanley Luchters;
stanley.luchters@ceshhar.co.zw

## ABSTRACT

**Introduction** Globally, recognition is growing of the harmful impacts of high ambient temperatures (heat) on health in pregnant women and children. There remain, however, major evidence gaps on the extent to which heat increases the risks for adverse health outcomes, and how this varies between settings. Evidence gaps are especially large in Africa. We will conduct an individual participant data (IPD) meta-analysis to quantify the impacts of heat on maternal and child health in sub-Saharan Africa. A detailed understanding and quantification of linkages between heat, and maternal and child health is essential for developing solutions to this critical research and policy area.

**Methods and analysis** We will use IPD from existing, large, longitudinal trial and cohort studies, on pregnant women and children from sub-Saharan Africa. We will systematically identify eligible studies through a mapping review, searching data repositories, and suggestions from experts. IPD will be acquired from data repositories, or through collaboration with data providers. Existing satellite imagery, climate reanalysis data, and station-based weather observations will be used to quantify weather and environmental exposures. IPD will be recoded and harmonised before being linked with climate, environmental, and socioeconomic data by location and time. Adopting a one-stage and two-stage meta-analysis method, analytical models such as time-to-event analysis, generalised additive models, and machine learning approaches will be employed to quantify associations between exposure to heat and adverse maternal and child health outcomes.

**Ethics and dissemination** The study has been approved by ethics committees. There is minimal risk to study participants. Participant privacy is protected through the anonymisation of data for analysis, secure data transfer and restricted access. Findings will be disseminated through conferences, journal publications, related policy and research fora, and data may be shared in accordance with data sharing policies of the National Institutes of Health.

## STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Prospectively collected data from cohorts and trials provide high-quality longitudinal data and contain a large number of variables.
⇒ Longitudinal data enable analysis of temporal relationships between repeated exposure-outcomes data, and more detailed causal analysis modelling than data collected at a single time point.
⇒ The large individual participant dataset will have statistical power to assess rare exposures and outcomes, explore high risk subgroups, and make risk comparison across different areas, regions, and countries, enhancing the study's external power and generalisability.
⇒ Our dataset may not be representative of the entire continent as data availability reflects research capacity across countries and not necessarily size of population, or urban–rural make-up, for example.
⇒ Missing data may occur due to reliance on the willingness of investigators to share data; however, this bias may be non-differential as the included studies had assessed different exposure outcome relationships than in our study.

## INTRODUCTION
### Background
Climate change is one of the greatest global health threats ever faced by humanity.[1][2] Increasing anthropogenic greenhouse emissions have caused the mean temperature of the world to rise by more than 1°C, and by as much as 2°C in many parts of Africa.[3–5] Projected temperature increases in both average temperatures and extreme events, such as heat-waves, are especially concerning.
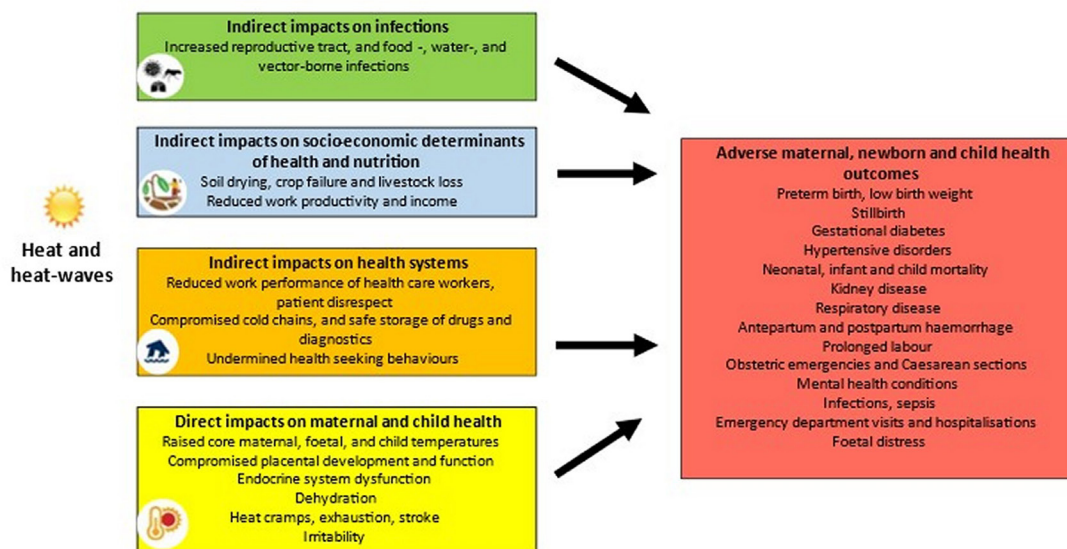
**Figure 1** Indirect and direct heat and heat-wave effects on maternal and child health.

Some estimates indicate that half of the global population will be exposed to more than 20 days of deadly heat per year by 2100,[6] but recent heat extremes suggest these figures may be an underestimate.

The climate change crisis, and heat in particular, has a wide range of deleterious effects on health. The indirect impacts of rising temperatures are well documented, such as an expanding geographical range of malaria vectors[7–9] and increased soil drying leading to food insecurity and malnutrition. Heat also indirectly affects health by fomenting wildfires, which destroy ecosystems and infrastructure.[10 11]

The direct impacts on human health due to exposure to high ambient temperatures (referred hereafter as heat) are increasingly recognised and affect a range of vulnerable populations.[2] Heat-waves cause increased rates of emergency room visits and hospitalisations, with an accompanying escalation in healthcare costs,[12] and result in substantial excess mortality. Moreover, the mental health sequelae of heat exposure are considerable, including generalised anxiety, depression, and eco-anxiety.[3 13]

### Heat exposure impacts on maternal and child health

Heat is hazardous for high-risk populations, including pregnant women and children (figure 1). The physiological and anatomical changes in pregnancy, pregnancy-related weight gain, heat generated by fetal metabolism, and exertion during labour, makes it challenging for pregnant women to maintain a normal temperature range when exposed to heat.[14 15] Manifestations of heat exposure include adverse pregnancy and birth outcomes, such as preterm birth, low birth weight, stillbirths,[13] gestational diabetes,[16 17] and hypertension in pregnancy.[18] Proposed biological mechanisms underlying the impact of heat on preterm birth include a reduction in placental blood flow, dehydration, and inflammatory responses.

However, further research is required to describe these biological mechanisms.[15]

Children, and particularly infants, have physiological, anatomic, and social factors that increase their vulnerability to heat, such as increased body surface to volume ratio, higher metabolic rate, and reliance on a caregiver.[19–21] Multiple studies have demonstrated a detrimental effect of heat on mortality,[19 22–26] kidney disease,[27] asthma and other respiratory disease,[28] and infectious diseases.[29 30] A modelling paper reported that under a high-emission scenario, heat-related child mortality in Africa may exceed 38 000 deaths per year in 2049.[31]

Several studies have shown that exposure to heat in utero negatively affects health throughout the life course, such as increased risks of stunting.[32] The consequences also extend to the larger health systems by increasing the burden on already stretched health resources due to increased rates of caesarean sections,[33] hospitalisation,[34] emergency department visits,[35 36] and outpatient and inpatient health facility visits.[37–39]

### Research gaps

Research on heat and health has been mostly restricted to stand-alone individual studies with relatively small sample sizes, poor-quality data from household surveys or healthcare facilities, considerable variation in research methodology, and limited geographical and temporal coverage.[13 19] Most studies have insufficient power to answer questions about which specific aspects of heat exposures (eg, timing and duration), which temperature patterns/thresholds (eg, night-time or day-time, or averages) are most harmful for different clinical conditions, and in which climate zones, settings, and subgroups.

Although some of the world's largest clinical trials have been conducted in Africa,[40] very little work has focused on heat impacts in key African population groups, such as pregnant women and children. Given the unique
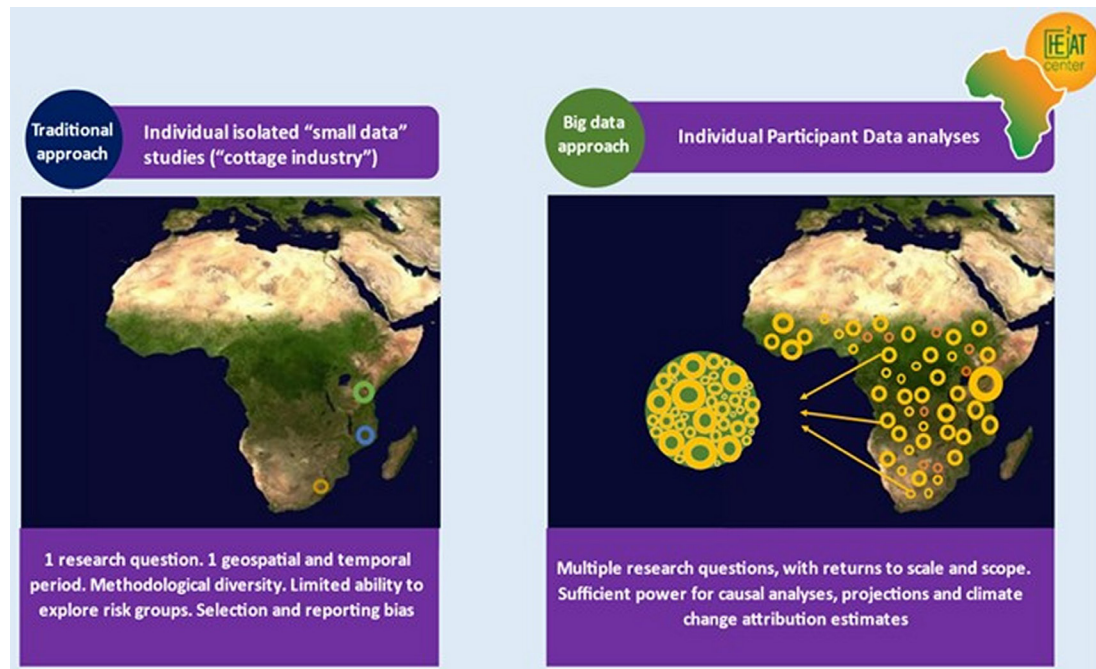
**Figure 2** The differences between traditional and individual participant data analysis approach to heat-health research in sub-Saharan Africa.

demographic profile, disease spectrum, built environment, and resource constraints in Africa, the most at-risk groups will likely differ from those in the Global North.

### Rationale for the individual participant data meta-analysis

We will conduct an individual participant data (IPD) meta-analysis using data collected from longitudinal cohorts and clinical trials on maternal and child health across sub-Saharan Africa. Utilising the IPD, we will quantify the current and future impacts of heat on maternal and child health in sub-Saharan Africa. Individual-level information enables more flexible and robust analyses than is possible in systematic reviews using aggregate study results from published data (figure 2).[41 42] The advantages of the IPD methodology are especially apparent in heat-health research, where larger sample sizes are required to detect relatively small exposure effects, and effects on rare outcomes.

### Public health relevance of the study findings

The study aims to better understand heat-health associations among pregnant women and children and results will inform monitoring of the heat-health burden, such as through indicators that could be used in a District Health Information System. Understanding the historical patterns of heat-health impacts is an important step towards monitoring changes in disease burden over time and projecting future burdens under different climate change scenarios and adaptation responses. By performing adequately powered and high-quality 'impact' studies, we will generate the information required to calculate the burden of disease from climate change. This, in turn, strengthens arguments for allocating sufficient resources to address climate-related impacts, and

for hastening societal changes required to avert further climate breakdown.

The study forms part of the Data Science Initiative Africa (DS-I Africa)[43] which aims to make optimum use of existing data resources across Africa to address the most pressing health concerns on the continent. The study constitutes one of two research projects within the HEat and HEalth African Transdisciplinary Center (HE$^2$AT Center)[44] project funded through the DS-I Africa Program.

### Study objectives

The study's overall objective is to use innovative data science approaches to quantify the current and future impacts of heat exposure on maternal and child health in sub-Saharan Africa.

The specific objectives are:

1. To systematically identify, acquire, collate, and integrate prospectively collected data from cohort studies and clinical trials on maternal and child health in sub-Saharan Africa.
2. To link maternal and child health outcome data spatially and temporally with weather and other environmental data, as well as socioeconomic and other data.
3. To use classic statistical and novel machine learning approaches to understand and quantify the impact of heat exposure on maternal and child health.
4. To document variations in the relationship between heat exposure and maternal and child health outcomes across different climate zones, settings, and population subgroups.
5. To develop innovative data science solutions for district-level surveillance of the impacts of heat on health.

## METHODS AND ANALYSES

### Study design and protocol registration

The IPD-MA will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) IPD extension guidelines.[45] The study began in June 2022, and we plan to conclude in June 2026. We completed the mapping review, started contacting data providers, and received 18 datasets. No analyses have been conducted. The protocol has been registered in PROSPERO (registration number: CRD42022346068).

### Study population

The study population are women in sub-Saharan Africa in pregnancy, childbirth and up to 2 years post partum, and their children up to 2 years of age, exposed to heat.

### Eligibility criteria

Eligibility is determined at the study and individual levels. Study-level inclusion criteria are:

1. Enrolment of at least 1000 pregnant women in 1 or more study sites, in 1 or more countries in sub-Saharan Africa.
2. Identified through published literature (published between January 2012 and June 2022) from the systematic mapping, clinical trial registry, data repository, or from study investigators and experts.
3. Randomised or non-randomised clinical trial, or an observational or interventional cohort with prospectively collected data.
4. At least two of the 'key' maternal and/or child health outcome variables (key maternal and child health outcomes have been selected based on evidence of heat-health impacts, an in alignment with the top causes of maternal and child mortality in sub-Saharan Africa.) have been collected as part of the study (online supplemental file 1).
5. Relevant local ethics approvals received, and documented.

At the individual level, the following inclusion criteria apply:

1. Enrolment into an eligible study, during pregnancy, or intrapartum.
2. IPD is available on the newborn's date of birth, date of diagnosis/occurrence of an adverse health outcome, or date of the end of pregnancy in cases of maternal deaths or abortion.
3. IPD is available on location of birth, or study follow-up.

### Rationale for eligibility criteria

Longitudinal data from clinical trials and cohort studies allow for the assessment of temporal trends and may avoid exposure biases as women are followed up over time, whereas in birth registries, for example, the women may have given birth in a place that is some distance away from where she spent much of her pregnancy.

The study only includes cohorts/trials that enrolled more than 1000 participants given that the large amount of time and resources required for data acquisition, preparation, harmonisation and analysis of each individual study make it difficult to justify the inclusion of smaller studies. Additionally, a large sample size is required for adequately powered studies for heat-health effect estimates that can be small. We selected recent studies published between 2012 and 2022 to ensure data availability, quality and relevance. Earlier studies may have used outdated clinical definitions and diagnostic criteria for adverse outcomes, which could complicate data harmonisation. Limiting the time frame improves our ability to identify data providers and their datasets, while also enhancing the quality of available environmental exposure data.

We are including studies where women are enrolled during pregnancy and intrapartum and including child data to the age of 2 years, if they are followed up as part of the study. Our primary focus is on heat exposure during pregnancy and intrapartum, and how that affects the pregnant mother and their child. Additionally, enrolling women in pregnancy may increase the likelihood of acquiring more accurate gestational age data, to explore windows of susceptibility.

### Data sources

First, we draw on studies identified through a systematic mapping. The search was conducted in 2020 in Medline (PubMed) and updated in 2022, using controlled vocabulary and free-text terms. Search terms for maternal health, for World Bank defined sub-Saharan African countries, and for filters to locate cohorts and clinical trials were included (online supplemental file 2).[46] The search strategy replicates those used in the study titled Multilateral Association for Studying health inequalities and enhancing north-south and south-south COoperaTions (MASCOT-1), which mapped global maternal health literature from 2000 to 2012.[40 47–50]

Using EPPI-Reviewer software,[51] screening of titles and abstracts was done independently, in duplicate, with differences between reviewers reconciled through discussion, or by a third reviewer. The full text was screened if eligibility could not be ascertained from the title or abstract. We extracted the following variables:

► Population: country, number enrolled.
► Methods: study design, topic.
► Identifiers: name, acronym, clinical registration number, authors, funders.

The second way we identify studies is through data repositories, such as the Bill and Melinda Gates Foundation Knowledge Integration platform, National Institutes of Health repositories, and ClinicalTrials.gov. Lastly, we will seek additional studies, published and unpublished, through direct contact with data providers and other experts.

### Risk-of-bias assessment

The quality of the studies will be evaluated using the Cochrane risk-of-bias tool for randomised trials (RoB 2), and the risk of bias in non-randomised studies of intervention for cohorts or non-randomised trials. Each study

and outcome will have an overall grading, which will be considered in meta-analyses and sensitivity analyses.

## Data collection

### Acquisition of IPD

Data will be acquired either through data access platforms such as Worldwide Antimalarial Resistance Network,[52] or directly from data providers. We will make at least five attempts to contact study investigators, including through contacting first, last, and other authors, and funders, through multiple communication platforms such as email, phone calls, and LinkedIn. Reasons for unattainable IPD will be included in a flow chart, such as the inability to contact data providers, unwillingness to share data, or the destruction of data. Data providers who agree to join the collaboration will sign a data sharing agreement that sets out the terms of data sharing, data security, and authorship. Opportunities for authorship, networking, and collaboration in study activities will be outlined and continually communicated. Collaborators will supply meta-data and key documentation, which will be used to confirm eligibility and for data management. Key documentation will include the study protocol, informed consent forms, codebook, and ethics approval confirmation.

### Acquisition of environmental exposure data

Weather data include observational-based datasets (weather station, or satellite remote sensing which is data from satellite sensors, mainly optical imagery (eg, satellite images of urban centres)), that provides information about physical attributes such as land surface temperature, vegetation characteristics, and land use) and processed or gridded observations. Climate-related data will mostly involve accessing open data repositories such as Copernicus Climate Data Store or Earth System Grid Federation data systems.

Air pollution data will be obtained from proxy satellite-derived air quality data such as Aerosol Optical Depth, in combination with land cover use, to overcome the challenge of gaps in the coverage of ground-based stations (figure 3). Where available, we will use data on pollutant concentrations, namely $PM_{10}$, $PM_{2.5}$, $NO_2$, $SO_2$, and CO for developing indices of air quality as well as leveraging global databases of air quality indicators such as the World Air Quality Index[53] and OpenAQ.[54]

## Data management and analysis

### Database development

In figure 4, we outline the steps that will result in the comprehensive database formation. In the first phase, the focus is on collecting IPD and metadata, and data quality and integrity. Once a data transfer agreement has been signed, IPD will be transferred to a password-protected platform using a secure data transfer.

Data from multiple studies will be characterised by differences in quality and this will be addressed before the synthesis phase of the meta-analysis, using Processing, Replication, Imputation, Merging, and Evaluation: PRIME-IPD[55] (table 1).

### Data harmonisation

We expect that the number of covariates for each study will be large and variable across studies. The harmonisation step will identify these discrepancies and formulate a strategy, mainly the creation of proxy covariates. The health covariates will be defined using published, reputable sources such as WHO, the Global Alignment of Immunization safety Assessment in pregnancy
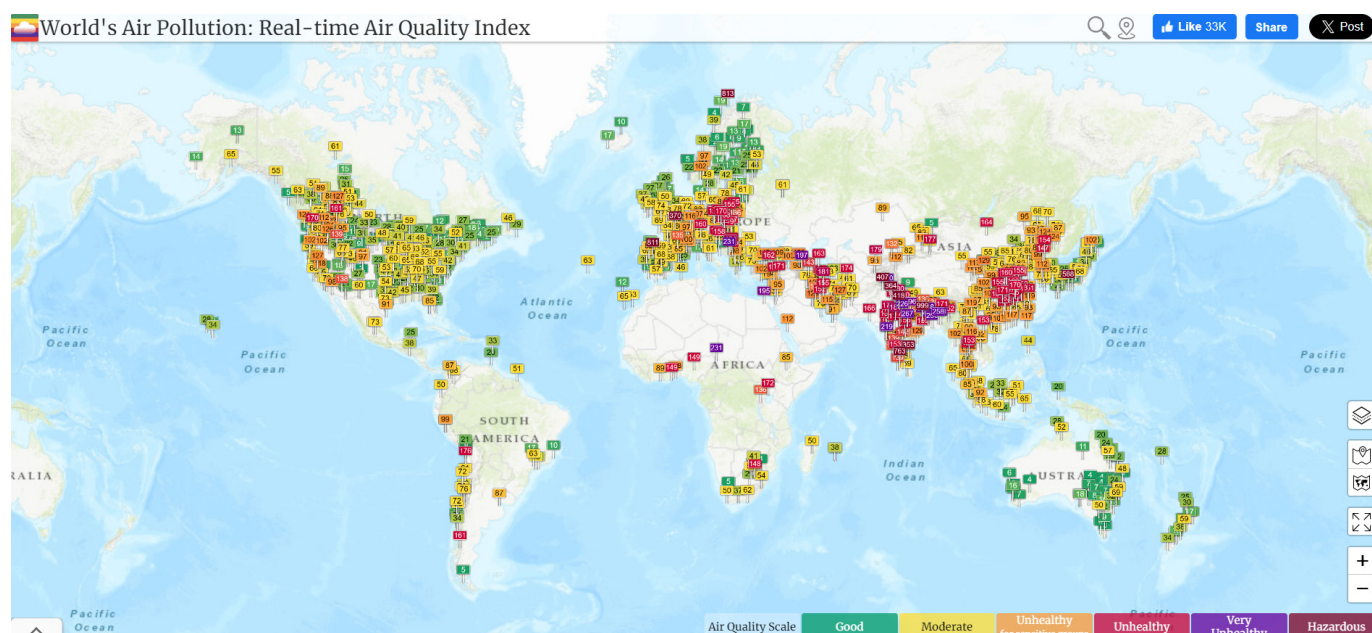


**Figure 3** Real-time air quality index for PM2.5 globally. The map shows the coverage of the monitoring network in Africa.[53]
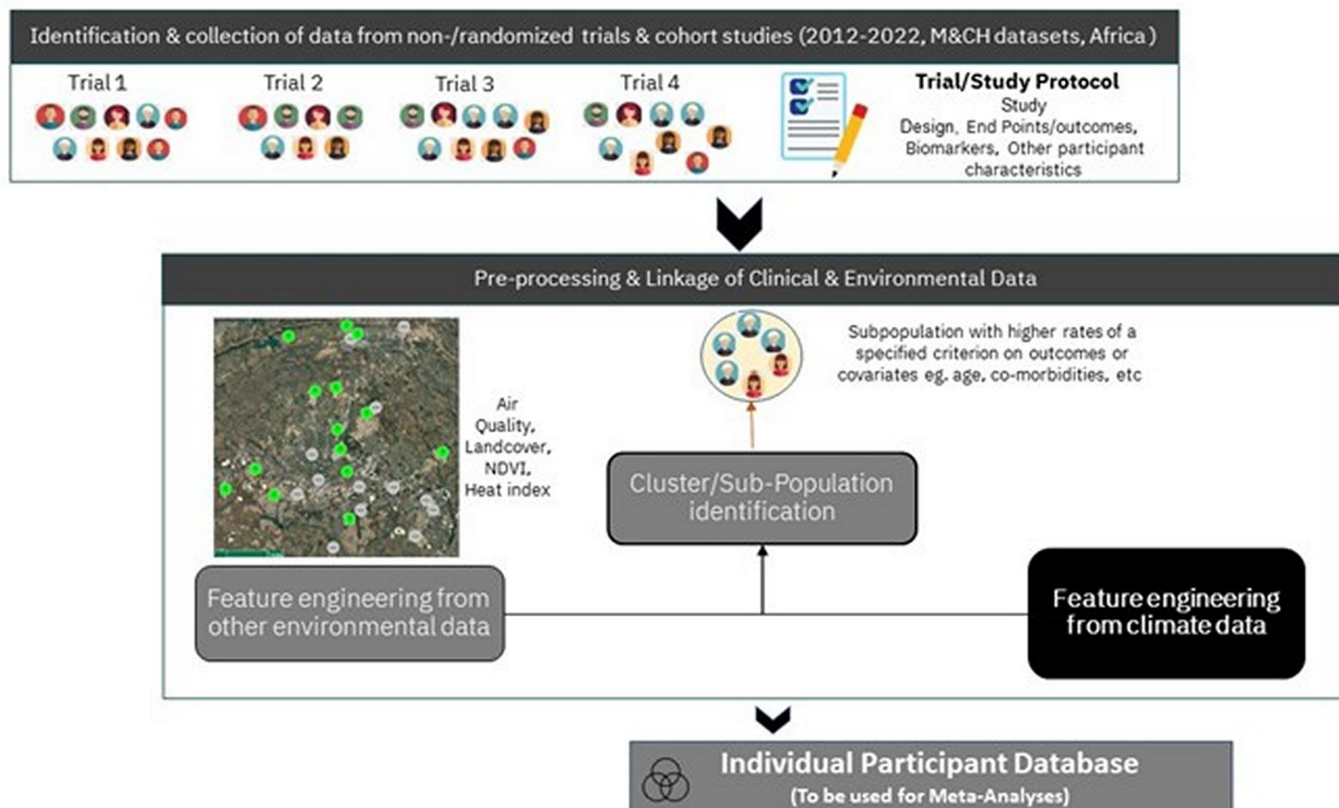
**Figure 4** Two phases of development of the database.

terminology,[56] standard ontologies, and local obstetric and paediatric guidelines.

The second phase of database formation, which runs parallel to the first, will consist of preprocessing climate, socioeconomic, and environmental data, to derive variables that will be included in the harmonised health database. The climate data are linked to time (eg, date of birth, date of health event) and location. Where we have high-resolution location information (GPS coordinates, home addresses), we will aggregate up to the appropriate administrative level that will minimise exposure mischaracterisation while protecting the privacy of individual participants. Once climate data are merged with health data, indirectly identifiable information such as date of birth and location will be removed from the integrated dataset to further protect participant confidentiality. To produce higher resolution daily temperature estimates, we will additionally combine satellite data, ERA5 land daily temperatures, and weather station data. In this context, further advances based on the exploitation of recent developments in geospatial artificial intelligence will be used. Some of the approaches that the team will implement include natural gradient boosting algorithms,[57] and quantile random forest spatial interpolations,[58] with implementation of maximum covariance analysis[59] to detect the structure of the covariance between these various forms of spatiotemporal datasets.

These two phases will result in an individual participant database consisting of health outcome variables, demographic covariates, and climate and environmental covariates. The integrated datasets will be made available to HE²AT Center partners for analysis through an access controlled, Jupyter Hub platform that is managed by the HE²AT Center data management and analysis core.

## Statistical analysis

The baseline characteristics of participants from each of the cohorts or trials will be described using R or Python.[60 61] A two-stage analysis approach will be used primarily, whereby, in the first stage, each study is analysed individually. In the second stage, the data from the individual studies will be aggregated to provide an overall pooled estimate of effect. We will explore analysing each study independently, and in combination with other studies through pooled analyses. We will evaluate heterogeneity of effects and precision of effect estimates to inform our approach.

### Statistical method for the first stage of the meta-analysis

The core modelling method for this study is linear and non-linear distributed lag models.[62] These models are specifically relevant where the outcome variable is a time series, typical with counts of adverse events such as preterm births. The most valuable characteristics of advanced forms of these models, such as the semiparametric generalised additive model (GAM) following a quasi-Poisson distribution with a distributed lag non-linear model, is the ability to account for non-linear, short-term, and lagged

**Table 1** Checklist for PRIME-IPD tool[58]

| PRIME | Items |
|---|---|
| Processing | Convert data into single format for statistical programme of choice |
| | Compare the total number of participants in the acquired datasets to those reported in published articles |
| | Verify the presence of the variables of interest in the acquired dataset |
| | Standardise variable names across datasets |
| | Identify and standardise the measurement scales used to report the variables of interest |
| | Identify and standardise coding for missing values |
| | Identify and correct any implausible values that may result from data conversion |
| Replication | Recalculate reported descriptive and summary statistics using the acquired datasets |
| | Calculate the standardised difference to quantitively assess the difference between the replicated and published results |
| | If the standardised difference is >10%, investigate and address potential causes |
| Imputation | Assess the appropriateness of conducting imputation of missing data using missing data theory |
| | If multiple imputation is conducted, carefully consider the number of imputations to be run |
| Merging | Ensure in processing step that variable order and codes are correct |
| | Merge the imputed dataset into a single, pooled dataset, taking into consideration the number of imputed datasets, if appropriate |
| Evaluation | Assess continuous variables for normality by residual analysis either visually or by statistical tests |
| | If required, calculate new variables for standardised comparison of effects |

effects of environmental exposures on health outcomes.[57] In general, the GAM framework provides the flexibility to account for non-linearity and overdispersions in the temporal dimension and clustering in the spatial dimension. The additive modelling framework may be expanded further to account for multiple health outcomes, associated uncertainties,[58] spatial effects, and interactions.[59 63] Therefore, large geospatial (and spatiotemporal) climate data from satellites and sensor networks can be leveraged. Further, depending on the type of outcome and duration of exposure, we will use additional statistical methodologies such as case-crossover, time-to-event, and longitudinal random forest methodologies. The case-crossover study design, commonly used to assess short-term environmental exposures and health outcomes, adjusts for all observed and unobserved individual level confounders as each case serves as its own control. Time-to-event analyses increases statistical power as all participants at risk are included, there is control of temporal trends (eg,

gestational age), and it can be used to investigate windows of susceptibility.[64] Longitudinal random forests are a machine learning approach that can be used to identify longitudinal exposure-related predictors of health.[65] In addition, the attributable risk of heat to adverse outcomes will be calculated as per the described methodologies by the Intergovernmental Panel on Climate Change.

### Machine learning informed covariate selection
We will consider traditional variable selection approaches for large data, as well as tree-based ensemble learning approaches, namely extreme gradient boosted trees or random forest algorithms. For both approaches, there is a split of the dataset into a training and testing set, the implementation of the tree algorithm, the evaluation of the variable importance ranking, and the possibility to use partial dependence plots to identify the functional form between pairwise or multiple covariates (including interactions) and the response variable. Both forms of tree algorithms can be used for binary, continuous, and time-to-event response variables. While autoencoding algorithms may be implemented for feature engineering from the geospatial or spatiotemporal climate datasets, ensemble tree algorithms may be used for automatic feature (covariate) selection, maintaining a level of explainability required when trying to understand the health effects of environmental exposures.[66]

### Statistical methods for the second stage of meta-analysis
Using the statistical methods described above, the association between health outcomes of interest and heat exposure for each study will be performed and a summary statistic presented to describe the estimated effects. The summary statistic will differ depending on the outcome and the analysis method used.

In the second stage, a weighted average of the effects of heat on maternal and child outcomes will be calculated, if levels of statistical heterogeneity are acceptable, and illustrated in a forest plot.

### Exploration of variation in effects across studies and subgroups
Exploration of variation in effects will be done involving stratified analyses within the following strata: study, geographical area, climate zone, time period, and income group of the country. Data will also be stratified on individual characteristics, such as maternal age, socioeconomic status, sex and health conditions such as HIV status. In these analyses, we generate estimates of impact (aggregate data) for each stratum separately and then combine these summary statistics using standard meta-analysis methods, if appropriate.

### Risk of bias across the IPD sources
Using the PRISMA-IPD flow chart, we will report the numbers of studies screened and included in the IPD, giving reasons for exclusions. We will describe the distribution of studies and the characteristics of participants for variables like location and age. We will compare study-level variables between the studies that we collected data

from, to those we could not. Drawing on this and factors such as the overall rate of participation in eligible studies, we will assess the potential risk of bias associated with non-availability of IPD from some studies.

### Additional analyses

We will perform sensitivity analyses to assess the robustness of results according to risk of bias, missing data, and quality of individual variables. For example, gestational age is prone to measurement bias and studies that had a poor methodological approach to measuring it may be excluded.

The expected outcomes of the study based on our primary and secondary hypotheses are summarised in online supplemental file 3.

### DISCUSSION

This is the first IPD-MA to investigate the impacts of heat exposure on maternal and child health. The IPD-MA will allow us to explore powered and flexible analyses on different aspects of heat exposure, in many maternal and child health outcomes, across diverse settings, climate zones, and subgroups in sub-Saharan Africa. The study results will inform monitoring efforts focused on the effects of heat on maternal and child health, that could be used to track changes in burden of disease over time and for assessing adaptation responses.

We acknowledge the potential limitations in the study design. Our IPD-MA may not be geographically representative due to differing research capacity across sub-Saharan African countries. We may not encounter the typical publication bias which occur with meta-analyses of published data as we draw on databases regardless of whether the exposure outcome of interest has been reported or if information is available on the presence or size of the association.[41] Nonetheless, we recognise the potential for published studies to be impacted by publication bias in the outcomes that the study had evaluated.

Further, we are limited to IPD shared by willing investigators from historical studies. We cannot avoid potential biases of the study characteristics (eg, selection bias by age of study) and of the quality of data collected, which may potentially vary by country, and thus climate zones. Lastly, our study may be at risk of exposure misclassification, common in heat-health research. Individual heat exposure will not have been collected and we assume women remain in one location throughout the study. To mitigate this risk, we employ longitudinal studies ensuring prolonged participant follow-up, leverage appropriate spatiotemporal scales for environmental data, use heat indices to represent heat strain, and potentially include housing type in analysis, where information is available.

### ETHICS AND DISSEMINATION
#### Ethical consideration and protection of human subjects

The study has been approved by the Wits Human Research Ethics Committee, Johannesburg (220605) and the National Ethics Committee for Life and Health Sciences, Cote d'Ivoire (176-22/MSHPCMU/CNESVS-kp). This study follows key guidelines such as the Declaration of Helsinki, South Africa Protection of Personal Information Act, South African Department of Health's Ethics in Health Research, US Department of Health and Human Services regulations 45 CFR 46, and other country-specific data protection legislation and ethics guidelines. The key ethical and legal considerations are (1) use of secondary data for research purposes, (2) risks associated with potential indirectly identifiable information and (3) cross-border data sharing in accordance with country-specific data protection legislation.

Firstly, the use of anonymised secondary data for research purposes in the HE$^2$AT Center, without the requirement of further informed consent procedures, meets the standards outlined in the guidelines described above.

Secondly, data may contain indirectly identifiable information like date of birth and location. We will take steps to minimise the risk of a privacy breach. We will not collect names of participants or other directly identifiable information, and no identifiable data will be published. The data will be safeguarded in a password-protected server with limited access. Lastly, where relevant, we will further anonymise data through geographical aggregation, jittering of home addresses, and removal of date of birth once climate variables have been linked.

Lastly, the use of health data requires consideration of country-specific ethical guidelines, legislation on the use of personal data, and the cross-border transfer of such datasets. Data providers will be required to provide contractual assurance in a data sharing agreement that consent for sharing is provided, that the required ethical procedures were followed, and that sharing of the data follows applicable data protection legislation.

#### Dissemination

We will promote the project and its findings, guided by good participatory practice guidelines, among communities where the research was conducted, and among maternal and child healthcare practitioners to promote awareness of heat-health risks. Dissemination tools such as newsletters, project posters, community advisory board discussions, and media will be used.

Project results will be disseminated to local, provincial and national authorities to provide technical support, and potentially inform policies. Many HE$^2$AT IPD investigators are active members of the Climate-Health Africa Network for Collaboration and Engagement, which facilitates communication among policy makers in Africa, and aims to enhance coherence in climate change and health policy across countries in Africa, which will be used for engagement. Our engagement plan includes

publications in open-access journals and presentations at conferences/meetings.

Lastly, anonymised data collected from this study may be made available through open-source platforms, with the permission of data providers, and approval from a HE²AT Center data access committee, to promote future research activities.

**Author affiliations**
[1]Wits RHI, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa
[2]Climate System Analysis Group, University of Cape Town, Rondebosch, South Africa
[3]University Peleforo Gon Coulibaly, Korhogo, Côte d'Ivoire
[4]University of Washington, Seattle, Washington, USA
[5]Centre for Sexual Health and HIV/AIDS Research, Harare, Zimbabwe
[6]Place Alert Labs, Department of Surveying and Geomatics, Faculty of the Built Environment, Midlands State University, Gweru, Zimbabwe
[7]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Durham, North Carolina, USA
[8]National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA
[9]IBM Research Africa, Johannesburg, South Africa
[10]Liverpool School of Tropical Medicine, Liverpool, UK
[11]Department of Public Health and Primary Care, Ghent Unviersity, Ghent, Belgium

**ORCID iDs**
Darshnika Pemi Lakhoo http://orcid.org/0000-0001-6114-7477
Matthew Francis Chersich http://orcid.org/0000-0002-4320-9168
Gloria Maimela http://orcid.org/0000-0002-7704-6803
Guéladio Cissé http://orcid.org/0000-0002-6286-976X
Ijeoma Solarin http://orcid.org/0000-0002-2045-3053
Prestige Tatenda Makanga http://orcid.org/0000-0002-9025-2564
Lisa van Aardenne http://orcid.org/0000-0002-9848-0940
Stanley Luchters http://orcid.org/0000-0001-5235-5629

## REFERENCES

1 Costello A, Abbas M, Allen A, et al. Managing the health effects of climate change. *Lancet* 2009;373:1693–733.
2 Romanello M, Di Napoli C, Drummond P, et al. The 2022 report of the lancet Countdown on health and climate change: health at the mercy of fossil fuels. *Lancet* 2022;400:1619–54.
3 Watts N, Amann M, Arnell N, et al. The 2020 report of the lancet Countdown on health and climate change: responding to converging crises. *Lancet* 2021;397:129–70.
4 Department of Environmental Affairs Republic of South Africa. South Africa's third national communication under the United Nations framework convention on climate change. 2018.
5 Engelbrecht F, Adegoke J, Bopape M-J, et al. Projections of rapidly rising surface temperatures over Africa under low mitigation. *Environ Res Lett* 2015;10:085004.
6 Global Commission on Adaptation. adapt now: a global call for leadership on climate Resilience. 2019.
7 Attaway DF, Jacobsen KH, Falconer A, et al. Risk analysis for Dengue suitability in Africa using the Arcgis predictive analysis tools (PA tools). *Acta Trop* 2016;158:248–57.
8 Ebi KL. Adaptation costs for climate change-related cases of Diarrhoeal disease, malnutrition, and malaria in 2030. *Global Health* 2008;4:9.
9 Ebi KL, Nealon J. Dengue in a changing climate. *Environ Res* 2016;151:115–23.
10 Xu R, Yu P, Abramson MJ, et al. Global climate change, and human health. *N Engl J Med* 2020;383:2173–81.
11 Zhang Y, Beggs PJ, McGushin A, et al. The 2020 special report of the MJA-lancet Countdown on health and climate change: lessons learnt from Australia's "black summer. *Med J Aust* 2020;213:490–2.
12 Ebi K, Hess J, Watkiss P, et al. Health risks and costs of climate variability and change. In: *Injury Prevention and Environmental Health*. 3rd edn. Washington, DC: The World Bank, 2017.
13 Chersich MF, Pham MD, Areal A, et al. Associations between high temperatures in pregnancy and risk of Preterm birth, low birth weight, and stillbirths: systematic review and meta-analysis. *BMJ* 2020;371:m3811.
14 Chersich MF, Scorgie F, Filippi V, et al. Increasing global temperatures threaten gains in maternal and newborn health in Africa: A review

of impacts and an adaptation framework. *Int J Gynaecol Obstet* 2023;160:421–9.

15 Samuels L, Nakstad B, Roos N, *et al*. Physiological mechanisms of the impact of heat during pregnancy and the clinical implications: review of the evidence from an expert group meeting. *Int J Biometeorol* 2022;66:1505–13.

16 Booth GL, Luo J, Park AL, *et al*. Influence of environmental temperature on risk of gestational diabetes. *CMAJ* 2017;189:E682–9.

17 Su W-L, Lu C-L, Martini S, *et al*. A population-based study on the prevalence of gestational diabetes mellitus in association with temperature in Taiwan. *Sci Total Environ* 2020;714:S0048-9697(20)30257-6.

18 Part C, le Roux J, Chersich M, *et al*. Ambient temperature during pregnancy and risk of maternal hypertensive disorders: A time-to-event study in Johannesburg, South Africa. *Environ Res* 2022;212:S0013-9351(22)00923-9.

19 Lakhoo DP, Blake HA, Chersich MF, *et al*. The effect of high and low ambient temperature on infant health: A systematic review. *Int J Environ Res Public Health* 2022;19:9109.

20 Smith CJ. Pediatric thermoregulation: considerations in the face of global climate change. *Nutrients* 2019;11:2010.

21 Xu Z, Sheffield PE, Hu W, *et al*. Climate change and children's health-a call for research on what works to protect children. *Int J Environ Res Public Health* 2012;9:3298–316.

22 Xu Z, Etzel RA, Su H, *et al*. Impact of ambient temperature on children's health: a systematic review. *Environ Res* 2012;117:120–31.

23 Xu Z, Sheffield PE, Su H, *et al*. The impact of heat waves on children's health: a systematic review. *Int J Biometeorol* 2014;58:239–47.

24 Basu R, Pearson D, Sie L, *et al*. A case-crossover study of temperature and infant mortality in California. *Paediatr Perinat Epidemiol* 2015;29:407–15.

25 Basagaña X, Sartini C, Barrera-Gómez J, *et al*. Heat waves and Causespecific mortality at all ages. *Epidemiology* 2011;22:765–72.

26 Son J-Y, Lee J-T, Lane KJ, *et al*. Impacts of high temperature on adverse birth outcomes in Seoul, Korea: disparities by Individual- and community-level characteristics. *Environ Res* 2019;168:460–6.

27 Khalaj B, Lloyd G, Sheppeard V, *et al*. The health impacts of heat waves in five regions of New South Wales, Australia: a case-only analysis. *Int Arch Occup Environ Health* 2010;83:833–42.

28 Lam LT. The association between Climatic factors and childhood illnesses presented to hospital emergency among young children. *Int J Environ Health Res* 2007;17:1–8.

29 Yan S, Wei L, Duan Y, *et al*. Short-term effects of meteorological factors and air Pollutants on hand, foot and mouth disease among children in Shenzhen, China, 2009-2017. *Int J Environ Res Public Health* 2019;16:3639.

30 Yin F, Zhang T, Liu L, *et al*. The association between ambient temperature and childhood hand, foot, and mouth disease in Chengdu, China: A distributed lag non-linear analysis. *Sci Rep* 2016;6:27305.

31 Chapman S, Birch CE, Marsham JH, *et al*. Past and projected climate change impacts on heat-related child mortality in Africa. *Environ Res Lett* 2022;17.

32 Chersich M, Kovats S, Part C, *et al*. Systematic review of the effect of ambient heat on maternal health outcomes. *ISEE Conference Abstracts* 2021;2021.

33 Molina O, Saldarriaga V. The perils of climate change: in utero exposure to temperature variability and birth outcomes in the Andean region. *Econ Hum Biol* 2017;24:111–24.

34 Davis RE, Novicoff WM. The impact of heat waves on emergency Department admissions in Charlottesville. *Int J Environ Res Public Health* 2018;15:1436.

35 Kakkad K, Barzaga ML, Wallenstein S, *et al*. Neonates in Ahmedabad, India, during the 2010 heat wave: a climate change adaptation study. *J Environ Public Health* 2014;2014:946875.

36 Sheffield PE, Herrera MT, Kinnee EJ, *et al*. Not so little differences: variation in hot weather risk to young children in New York City. *Public Health* 2018;161:119–26.

37 Kim J, Lee A, Rossin-Slater M. What to expect when it gets hotter. *Am J Health Econom* 2021;7:281–305.

38 Zhao Q, Li S, Coelho MSZS, *et al*. Assessment of Intraseasonal variation in hospitalization associated with heat exposure in Brazil. *JAMA Netw Open* 2019;2:e187901.

39 Qu Y, Zhang W, Ryan I, *et al*. Ambient extreme heat exposure in summer and transitional months and emergency Department visits and hospital admissions due to pregnancy complications. *Sci Total Environ* 2021;777:S0048-9697(21)01201-8.

40 Chersich M, Blaauw D, Dumbaugh M, *et al*. Mapping of research on maternal health interventions in Low- and middle-income countries: a review of 2292 publications between 2000 and 2012. *Global Health* 2016;12:52.

41 Stewart LA, Tierney JF. To IPD or not to IPD? advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;25:76–97.

42 Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:bmj.c221.

43 DSI-Africa. *DSI-Africa 2023*. Available: https://dsi-africa.org

44 HEAT Center. About HEAT center. 2023. Available: https://heatcenter.wrhi.ac.za/About-HEAT-Center

45 Stewart LA, Clarke M, Rovers M, *et al*. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. *JAMA* 2015;313:1657–65.

46 The University of Texas: School of Public Health. Search filters for various databases. 2021. Available: https://libguides.sph.uth.tmc.edu/search_filters/pubmed_filters

47 Chersich M, Becerril-Montekio V, Becerra-Posada F, *et al*. Perspectives on the methods of a large systematic mapping of maternal health interventions. *Global Health* 2016;12:51.

48 Chersich MF, Blaauw D, Dumbaugh M, *et al*. Local and foreign authorship of maternal health Interventional research in Low- and middle-income countries: systematic mapping of publications 2000-2012. *Global Health* 2016;12:35.

49 Chersich MF, Martin G. Priority gaps and promising areas in maternal health research in Low- and middle-income countries: summary findings of a mapping of 2292 publications between 2000 and 2012. *Global Health* 2017;13:6.

50 Footman K, Chersich M, Blaauw D, *et al*. A systematic mapping of Funders of maternal health intervention research 2000-2012. *Global Health* 2014;10:72.

51 Thomas J, Graziosi S, Brunton J, *et al*. EPPI-reviewer: advanced software for systematic reviews, maps and evidence synthesis. EPPI-centre, UCL social research Institute, University college London. 2022.

52 WWARN. Accessing data, . 2002Available: https://www.wwarn.org/working-together/sharing-accessing-data/accessing-data

53 World Air Quality Index. World air quality index project. 2022. Available: https://aqicn.org/contact

54 OpenAQ. Open Air Quality 2022, Available: https://openaq.org

55 Dewidar O, Riddle A, Ghogomu E, *et al*. PRIME-IPD SERIES part 1. The PRIME-IPD tool promoted verification and standardization of study Datasets retrieved for IPD meta-analysis. *J Clin Epidemiol* 2021;136:227–34.

56 Bonhoeffer J, Kochhar S, Hirschfeld S, *et al*. Global alignment of immunization safety assessment in pregnancy - the GAIA project. *Vaccine* 2016;34:5993–7.

57 Duan T, Anand A, Ding DY, *et al*. Ngboost: natural gradient boosting for probabilistic prediction. Hal D, Aarti S, eds. Proceedings of the 37th International Conference on Machine Learning; Proceedings of Machine Learning Research: PMLR; 2020:2690–700

58 Sekulić A, Kilibarda M, Heuvelink GBM, *et al*. Random forest spatial interpolation. *Remote Sensing* 2020;12:1687.

59 Barreto NJC, Mesquita M d. S, Mendes D, *et al*. Maximum covariance analysis to identify Intraseasonal Oscillations over tropical Brazil. *Clim Dyn* 2017;49:1583–96.

60 Core Team R. R: A Languge and environment for statistical computing. *Vienna, Austria R Foundation for Statistical Computing* 2020.

61 vanG. *Python Library Reference*. Amsterdam: Centrum voor Wiskunde en Informatica (CWI), 2006.

62 Deeks JJ, Higgins JP, Altman DG. Chapter 10: analysing data and undertaking meta-analyses. In: *Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022)*. 2022. Available: www.training.cochrane.org/handbook

63 Guilherme Marmerola. xgbse: Improving XGBoost for Survivial Analysis, . 2021Available: https://towardsdatascience.com/xgbse-improving-xgboost-for-survival-analysis-393d47f1384a

64 de Bont J, Stafoggia M, Nakstad B, *et al*. Associations between ambient temperature and risk of Preterm birth in Sweden: A comparison of Analytical approaches. *Environ Res* 2022;213:S0013-9351(22)00913-6.

65 Loef B, Wong A, Janssen NAH, *et al*. Using random Forest to identify longitudinal predictors of health in a 30-year cohort study. *Sci Rep* 2022;12:10372.

66 Wang J, Keusters WR, Wen L, *et al*. Ipdmada: an R shiny tool for analyzing and Visualizing individual patient data meta-analyses of diagnostic test accuracy. *Res Synth Methods* 2021;12:45–54.