









STUDY PROTOCOL

Climate action for health and wellbeing in cities: a protocol for the systematic development of a database of peer-reviewed studies using machine learning methods [version 1; peer review: awaiting peer review]

Kristine Belesova ¹, Max Callaghan ², Jan C Minx ², Felix Creutzig², Catalina Turcu ³, Emma Hutchinson¹, James Milner¹, Melanie Crane ⁴, Andy Haines ¹, Michael Davies⁵, Paul Wilkinson¹

¹Department of Public Health, Environments and Society and Centre on Climate Change and Planetary Health, London School of Hygiene and Tropical Medicine, London, WC1H 9SH, UK

²Mercator Research Institute on Global Commons and Climate Change, Berlin, 10829, Germany

³Bartlett School of Planning, University College London, London, WC1H 0QB, UK

⁴Charles Perkins Centre, Sydney School of Public Health, University of Sydney, Sydney, Australia

⁵Bartlett School Environment, Energy & Resources, University College London, London, WC1H 0QB, UK

V1 First published: 05 Mar 2021, 6:50
<https://doi.org/10.12688/wellcomeopenres.16570.1>

Latest published: 05 Mar 2021, 6:50
<https://doi.org/10.12688/wellcomeopenres.16570.1>

Open Peer Review

Reviewer Status AWAITING PEER REVIEW

Any reports and responses or comments on the article can be found at the end of the article.

Abstract

Cities produce more than 70% of global greenhouse gas emissions. Action by cities is therefore crucial for climate change mitigation as well as for safeguarding the health and wellbeing of their populations under climate change. Many city governments have made ambitious commitments to climate change mitigation and adaptation and implemented a range of actions to address them. However, a systematic record and synthesis of the findings of evaluations of the effect of such actions on human health and wellbeing is currently lacking. This, in turn, impedes the development of robust knowledge on what constitutes high-impact climate actions of benefit to human health and wellbeing, which can inform future action plans, their implementation and scale-up. The development of a systematic record of studies reporting climate and health actions in cities is made challenging by the broad landscape of relevant literature scattered across many disciplines and sectors, which is challenging to effectively consolidate using traditional literature review methods. This protocol reports an innovative approach for the systematic development of a database of studies of climate change mitigation and adaptation actions implemented in cities, and their benefits (or disbenefits) for human health and wellbeing, derived from peer-reviewed academic literature. Our approach draws on extensive tailored search strategies and machine learning methods for article classification and tagging to

generate a database for subsequent systematic reviews addressing questions of importance to urban decision-makers on climate actions in cities for human health and wellbeing.

Keywords

climate change, cities, urban health, mitigation, adaptation, planetary health, case studies, implementation, wellbeing, public health, climate action, evaluation, intervention, solutions, actions

Corresponding author: Kristine Belesova (kristine.belesova@lshtm.ac.uk)

Author roles: **Belesova K:** Conceptualization, Data Curation, Investigation, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Callaghan M:** Data Curation, Investigation, Methodology, Software, Writing – Review & Editing; **Minx JC:** Conceptualization, Investigation, Methodology, Software, Writing – Review & Editing; **Creutzig F:** Conceptualization, Investigation, Methodology, Software, Writing – Review & Editing; **Turcu C:** Investigation, Methodology, Writing – Review & Editing; **Hutchinson E:** Investigation, Methodology, Writing – Review & Editing; **Milner J:** Methodology, Writing – Review & Editing; **Crane M:** Methodology, Writing – Review & Editing; **Haines A:** Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing; **Davies M:** Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing; **Wilkinson P:** Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through an Our Planet, Our Health programme grant for the consortium Complex Urban Systems for Sustainability and Health (CUSSH) [209387, <https://doi.org/10.35802/209387>]. JCM and MC acknowledge funding from the German Federal Ministry of Education and Research [01LA1826A, 03SFK5J0].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Belesova K *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Belesova K, Callaghan M, Minx JC *et al.* **Climate action for health and wellbeing in cities: a protocol for the systematic development of a database of peer-reviewed studies using machine learning methods [version 1; peer review: awaiting peer review]** Wellcome Open Research 2021, 6:50 <https://doi.org/10.12688/wellcomeopenres.16570.1>

First published: 05 Mar 2021, 6:50 <https://doi.org/10.12688/wellcomeopenres.16570.1>

Introduction

Cities are responsible for 71% to 76% of global energy-related carbon emissions, including both consumption and production-related emission (Seto *et al.*, 2014). To limit global average temperature increase to well below 2°C, CO₂ and short-lived climate pollutant emissions need to be reduced to net zero (often abbreviated as the net zero target) within the next 50 years – though some suggest that cities should achieve this much earlier (C40 & Arup, 2017). The achievement of this deadline would require climate action at all scales: individual, city, national and international levels triggering rapid transformation of the ways in which urban societies operate.

In 2015, there were over 10,000 climate actions identified as being undertaken in the 96 cities comprising the C40 cities climate leadership group, with further potential 26,000 actions identified that could be implemented to expand their range of existing climate actions (C40 Cities & Arup, 2015). However, so far urban climate actions have not resulted in rapid and sustained emission reductions that are required to meet the climate goals of the Paris Agreement (Reckien *et al.*, 2014). The overall progress on climate change action so far has been seriously inadequate in comparison to the magnitude of the challenge, as global greenhouse gas emissions have been steadily rising. Although a decline of 4 to 7% (2% to 13%) in global CO₂ emissions is projected in 2020 due to the measures taken in response to the COVID-19 pandemic, those are not suitable for the required sustained long-term emission reduction (Belesova *et al.*, 2020b; Forster *et al.*, 2020; Quéiré *et al.*, 2020).

Climate change mitigation and adaptation actions have implications for human health and wellbeing in cities (see Box 1 for definitions). Mitigation actions can produce considerable health and wellbeing benefits for urban residents, for example, through reductions in air pollution, increased levels of active travel, reduced noise levels, more urban greenspace (Gao *et al.*, 2018). Adaptation actions can benefit health through, for example, reduced risk of extreme weather effects, exposure to vector-borne infectious diseases and greater resilience to socio-economic shocks (Smith *et al.*, 2014). Nevertheless, certain climate actions and their implementation may also result in unintended adverse consequences e.g., the use of air conditioning to manage thermal comfort during heatwaves leads to increased electricity consumption and further exacerbation of urban heat reducing the overall health benefits of reduced indoor heat exposure; likewise, building retrofit could, if not implemented correctly, lead, for example, to poor ventilation and accumulation of pollutants generated indoors with adverse effects on health (Li *et al.*, 2014; Shrubsole *et al.*, 2014; Wenz *et al.*, 2017).

Box 1. Definitions of key concepts

- **Mitigation:** actions that aim to reduce sources or enhance sinks of greenhouse gases (IPCC, 2012).
- **Adaptation:** “the process of adjustment to actual or expected climate and its effects” (IPCC, 2012).

- **Action:** tangible actions to alter institutions, technology, policies, programs, built environments, mandates or behaviours in the effort to reduce the rate of climate change and/or adapt to it (Lesnikowski *et al.*, 2011; Lesnikowski *et al.*, 2013).*
- **Health:** “health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (World Health Organization (WHO), 1948). Although the WHO definition encompasses physical, mental and social health and wellbeing, here we largely reserve the term “health outcomes” for adverse physical and mental health outcomes and use the term ‘wellbeing’ separately.
- **Wellbeing:** here is used to emphasise the wider psychological and social aspects of human health that determine subjective human wellbeing, as defined by Marsh *et al.* (2020), understood as a cognitive sense of life satisfaction and pleasant or unpleasant affect (moods and emotions) (Dodge *et al.*, 2012).

*As a part of action we also include “groundwork activities”, i.e., activities that prepare conditions for mitigation/adaptation, enable mitigation and adaptation actions, inform and prepare stakeholders for the actions, e.g., vulnerability assessments, adaptation and mitigation research, development of conceptual tools, stakeholder networking, and provision of policy recommendations (Lesnikowski *et al.*, 2011; Lesnikowski *et al.*, 2013). These often contribute to the capacity to mitigate or adapt to climate change.

To our knowledge, there are no comprehensive assessments of collective ‘real world’ climate actions and their impacts on human health and well-being in cities, including comprehensive databases of publications available for syntheses or syntheses themselves. Lamb *et al.* used machine learning methods to develop a database of 4,000 urban climate mitigation case studies reported in peer-reviewed literature (Lamb *et al.*, 2019), but their approach does not include climate change adaptation terms, nor the implications of climate actions for human health and wellbeing. The lack of comprehensive syntheses addressing stakeholder needs hinders evidence-based decision-making on, and effective implementation of, the most reliable high-impact actions.

The following protocol presents an approach for the systematic development of a database of peer-reviewed studies to enable synthesising evidence on climate actions implemented in cities and their effects on human health and wellbeing. Identifying such studies requires investigation across a wide pool of literature dispersed across many disciplines, sectors, and topics. The capacity of traditional systematic literature review processes to comprehensively capture this extensive literature is limited. Therefore, here, we extend the machine-learning methods developed by Lamb *et al.* (2019), and apply them to a set of extensive specialised search strategies developed by our multi-disciplinary team to cover studies of both urban climate change mitigation and adaptation actions relevant to human health and wellbeing. Machine learning entails training a computer to perform some parts of work automatically (Berrang-Ford *et al.*, 2021). In the case of systematic literature identification, machine learning can be used to reduce human resources needed to

search, screen, and classify tens of thousands of scientific articles, and hence, cover a much broader scope of literature (Berrang-Ford *et al.*, 2021). This protocol provides an overview of the process and methods for developing the database of studies.

This database is being developed as a part of the Complex Urban Systems for Sustainability and Health (CUSSH) project in collaboration with lead authors of the upcoming 6th Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). The CUSSH project is a Wellcome Trust funded five-year consortium of 13 academic and other institutions collaborating with six cities – London, Rennes, Nairobi, Kisumu, Beijing, and Ningbo – the C40 Cities Climate Leadership Network, and other stakeholders. Its aim is to deliver key global research on the systems that connect urban sustainable development and population health (Belesova *et al.*, 2018). The database will provide evidence to improve understanding of a range of questions about urban mitigation/ adaptation actions and human health impacts — relevant to cities around the world, and to urban policy makers and NGOs who require examples of good practice in climate action, city networks such as C40, and the IPCC. The IPCC increasingly pays attention to demand-side solutions that require evaluation in terms of their impact on health and wellbeing (Creutzig *et al.*, 2018). As most effects of climate action will take place at the scale of human urban settlements, a comprehensive database of the underlying literature will hence be invaluable.

Aim and research areas

Our aim is to test and implement an approach for the systematic development of a database of urban climate change mitigation and adaptation actions (see the definition in [Box 1](#)) relevant to human health and wellbeing that have been implemented in cities and reported in peer-reviewed academic literature. Subsequently, we intend to use this database to address questions about current climate mitigation and adaptation actions for health and well-being that are implemented in cities and reported in peer-reviewed literature in at least three broad areas:

1. Distribution, variation, and diversity of actions across the world's cities
2. The scale of impact and effectiveness of the actions
3. Lessons for their successful implementation

We particularly seek input from policy makers and other urban stakeholders to ensure the relevance of our research questions and strengthen the translational relevance of the evidence to policy and practice. Protocols for each of the subsequent systematic reviews will be reported separately. This protocol covers only the approach to the development of the source database itself. This protocol has been registered in the OSF Registries (Belesova *et al.*, 2020a).

Scope of the database

The scope of the database is framed by the PICO statement and inclusion criteria elaborated below.

PICO statement

Population: populations in any of the cities included in the *GeoNames* database (GeoNames Team, 2020)

Intervention: actions targeting either climate change mitigation and/or adaptation

Comparison: any form of comparison, including, but not limited to change-over-time and between-area controlled studies, as well as descriptive reports

Outcome:

- (1) any form of documented health and wellbeing outcome, whether positive or negative¹,
- (2) change in policy ambitions or in the drivers and processes contributing to climate change mitigation and/or adaptation

Inclusion criteria

Papers will be included according to the following criteria:

- **Literature type:** peer-reviewed scientific papers reporting original research;
- **Year of publication:** from 1990 until present to capture literature produced after the publication in year 1990 of the first assessment report by the IPCC;
- **Language:** English;
- **Study type:** any observational, evaluation, natural experiment or modelling study that reports any action(s) that
 - target **climate change** mitigation and/or adaptation;
 - report actual or potential impact on health and/or wellbeing;
 - are intended to have impact at the level of the whole city or substantial part of the city (neighbourhood-level and larger);
 - have been **implemented**, i.e., their implementation is ongoing, completed, failed.

Exclusion criteria

- **Literature type:** non peer-reviewed literature, commentaries, editorials, systematic and non-systematic literature reviews, opinion pieces, news reports, book chapters, meeting reports/conference proceedings, grey literature;
- **Year of publication:** before 1990;
- **Language:** other than English;

¹ We will also include studies with changes in those exposures that are known to have strong links to health and wellbeing outcomes, even if these studies do not measure the outcomes

- **Study type:** studies that do not report implementation of a relevant action(-s) (as per inclusion criteria above), e.g., association studies of an environmental exposure with an outcome but without an intervention (e.g., studies of the heat—mortality association), methodological studies, carbon accounting studies, conceptual and protocol papers, in silico modelling actions that have not been implemented, i.e., action plans, masterplans, scenarios, scenario modelling, urban visions, guidelines, frameworks (N.B. modelling studies that evaluate impacts of an existing intervention are included, e.g., modelling of the effects of intervention-related changes in air pollution),
- **Scale:** small scale, niche action or pilot studies not intended to have impact at the level of the whole city or substantial part of the city (i.e., actions at scales below neighbourhood-level);
- **Implementation:** actions that are only hypothesized or planned and not yet implemented.

We will include only anticipatory and purposeful adaptation and mitigation and exclude unplanned/ spontaneous efforts (co-)resulting in adaptation and mitigation or entirely hypothetical actions with no attempt at implementation. Hence, the database will focus on tangible actions and preparatory activities specifically intended to reduce GHG emissions or the adverse effects of climate change which report potential or actual impacts on human health and wellbeing.

Literature search, screening, and data management

Our three-stage process of study identification (initial selection, machine learning, and eligibility assessment) will be done in parallel to data management, with both being supported by machine learning and human processing (Figure 1).

Stage I: initial selection

We will make use of a search that has been performed across [Web of Science Core Collections](#) (consisting of SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, and ESCI), [Scopus](#), and [Medline](#) (accessed via Web of Science). These databases allow access to literature from diverse disciplinary perspectives, matching the transdisciplinary nature of our research. With the breadth of our search results approximating a total of 650,000 records (including duplicates and before applying the city filter), we do not search an even wider set of databases due to a lack of an automated workflow that would support other bibliographic databases and grey literature search within our current technical infrastructure and capabilities.

We have developed two separate sets of search terms, one to capture climate change adaptation actions and one climate change mitigation actions (Figure 2). Papers describing actions that are taken to contribute to climate change mitigation or adaptation may not always mention the terms ‘mitigation’ or ‘adaptation’ or their synonyms (e.g., ‘emission reduction’ and ‘climate risk reduction’) explicitly. Therefore, both search strategies

contain terms for actions likely to entail mitigation and/or adaptation effects even if ‘adaptation’ or ‘mitigation’ is not mentioned. Such terms include, for example, ‘energy efficiency policies’ and ‘coastal protection’.

The structure of the search terms is shown in Figure 2. Both searches entail the intersection (‘AND’) of three blocks of terms, representing climate, mitigation/adaptation actions, and health/wellbeing. Specifically, for mitigation the blocks of terms are: (climate terms OR energy terms) AND (explicit mitigation and mitigation policy terms OR sector-specific mitigation terms) AND (health terms OR wellbeing terms). For adaptation the terms are: (climate terms) AND (explicit adaptation terms including resilience OR action-specific adaptation terms) AND (health terms OR wellbeing terms). The sector-specific mitigation terms are grouped by the following ‘sectors’: transport, buildings, urban form and greenspace, waste and food. The action-specific adaptation terms follow sets of adaptation strategies that can be implemented for each category of climate change impacts on human health outlined in the health chapter of the 5th assessment report of the IPCC with further sectoral terms of relevance to health (Smith *et al.*, 2014). Thus, the mitigation search strategy targets all climate change mitigation actions mentioning health or human wellbeing terms in their abstract/title or keywords. The adaptation search strategy targets those actions that have the potential to protect people from health risks of climate change (e.g., coastal defences) and does not cover actions implemented to maintain stability of other sectors (e.g., profits in the tourism sector) without mentioning human health or wellbeing. In the health and wellbeing block of terms, we did not attempt to include terms for well-known exposures through which mitigation and adaptation actions affect human health, e.g., air pollution. Inclusion of only few well-known exposure terms would have biased our search results towards well-known actions and impact mechanisms; inclusion of terms for all possible exposures and impact pathways would have been impractical due to the extensive range of potential pathways. We also did not include terms for long-term health effects, such as cancers, the rates of which can change with the implementation of certain mitigation strategies, as they do not reflect health co-benefits of climate change mitigation – a concept that refers to the benefits occurring the near-term, and thus, providing an additional incentive for the implementation of mitigation actions (West *et al.*, 2013).

The adaptation terms were constructed by KB with additions and edits from all co-authors. The mitigation terms were constructed by JM, FC and KB with additions and edits from all co-authors. Terms were selected based on keystone articles and authors’ expertise, which covers all major urban sectors (transport, buildings, urban form including green and blue space, food systems, waste, energy, water and sanitation, health care, industry) and relevant disciplines (epidemiology, public health, urban planning, building science, climate change and energy, sustainability science, and implementation science). Climate terms for the mitigation search were adapted from Grieneisen & Zhang (2011) and for the adaptation search from Whitmee *et al.* (2015). Adaptation terms for the food sector

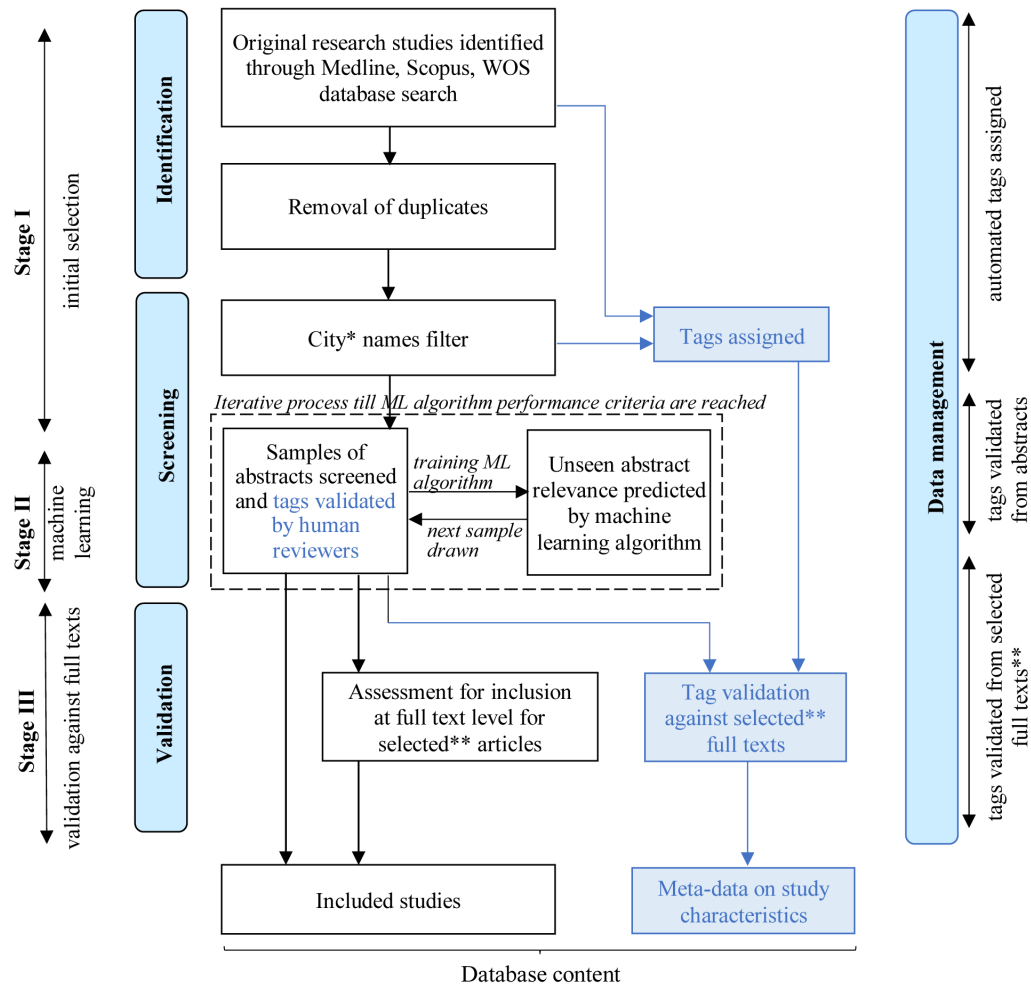


Figure 1. A diagram of the review and database development process. ML: machine learning; WOS: Web of Science. *also includes names of major city networks. **selected articles include the articles whose inclusion was not decided, and a random sample of the articles included following the abstract level assessment.

were adapted from Aleksandrowicz *et al.* (2016). Health terms were adapted from Berrang-Ford *et al.* (2021) with additions based on reviewing the International Statistical Classification of Diseases and Related Health Problems 10th Revision (World Health Organization (WHO), 2019) for terms of potential relevance. Wellbeing terms were constructed by KB drawing on the determinants of subjective wellbeing (Marsh *et al.*, 2020) and published search strategies on mental health and wellbeing (Clark & Paunovic, 2018; Cochrane Common Mental Disorders, 2014; Keynejad *et al.*, 2020; Larvin *et al.*, 2019; Wilczynski *et al.*, 2006).

When constructing the search strategy, we tested the search term relevance and performance of different combinations of search terms, identifying and restricting any terms that contributed excessive numbers of irrelevant results. The search strategy was designed to be comprehensive and allow identifying understudied mitigation and adaptation actions.

Search terms were reviewed by all team members and peer-reviewed by a librarian experienced in systematic review searching, using the PRESS Guidance (McGowan *et al.*, 2016). Recommendations were discussed with the project team and implemented where appropriate. The final set of search terms developed for Medline is available as extended data (Belesova *et al.*, 2021). The search strategies were adapted to the syntax of each bibliographic database. The bibliographic databases were searched by article abstracts, titles, key words, and headings in Scopus and Medline and by article abstracts, titles in Web of Science. Keyword searches in Web of Science were excluded as their inbuilt “keyword plus” function generates additional default searches that are not transparent and generate a large number of irrelevant records (Berrang-Ford *et al.*, 2021).

We will run the searches restricting them to the original research study articles published in English. We will remove



Figure 2. Search strategy structure.

duplicates and discard records where the abstract is missing, as the abstracts are essential to decide on study inclusion into the database. To identify actions implemented in urban areas, we will search the selected abstracts for names of cities and city networks. We use the Geonames database of geographic locations, which aggregates national survey data, travel destinations and open-sourced contributions, specifying a global list of cities with populations greater than 15,000. To this list we will add names of such city networks as C40 city leadership group, Multi-City Multi-Country (MCC) collaborative research network, ILCEI – Local Governments for Sustainability, Covenant of Mayors, 100 Resilient Cities, and others.

Stage II: machine learning assisted screening

We will use supervised machine learning (ML) to accelerate the screening process through an iterative cycle of training, testing against manual classification and re-training of the ML algorithm until a stopping rule is met (Callaghan & Müller-Hansen, 2019). We will use NACSOS research platform for screening process and Python scikit-learn library for machine learning (Callaghan *et al.*, 2020; Pedregosa *et al.*, 2012).

The process is illustrated in Figure 3. First, a random sample of the studies selected by the initial literature search is screened by two human reviewers and classified with respect to relevance, with any inconsistencies between reviewers resolved by a third reviewer as necessary. This classification is treated as the ‘gold standard’ classification of relevance (‘the truth’). The classified random sample is then used to train a machine learning algorithm, which, once trained, is applied to predict the relevance of the remaining studies unseen by human reviewers,

assigning to each study a score from 0 to 1 for the likelihood of relevance.

The studies with the highest ML scores (those with a score closest to 1) are then reviewed by two independent human reviewers whose classification is used to define the proportion of the high ML-score sample that was correctly classified as truly relevant.

The human (‘gold standard’) classification of this sample is then used to improve the training of the ML algorithm in an iterative process of further ML training, re-classification of studies and testing the results against by human reviewers.

After each cycle, a test is applied of whether to continue the iterative cycles or stop. The stopping rule is based on the proportion of correctly classified studies in the (highest score) ML sample after each iteration. When the proportion of truly relevant studies in the new ML sample is very low, it can be concluded that the proportion of relevant studies that are not being identified by the ML algorithm is also small. Formally, we plan to continue the iterations until statistical testing suggests that we can reject a null hypothesis that <90% of truly relevant studies with a confidence level of 90%. However, the stopping rule will be reviewed and amended as appropriate in the light of the results of processing the initial batches of studies to ensure an efficient and manageable process.

Data management

Meta-data classifying studies by their basic characteristics will be added to the database through a semi-automated tagging process. Tags will be assigned to all the studies identified

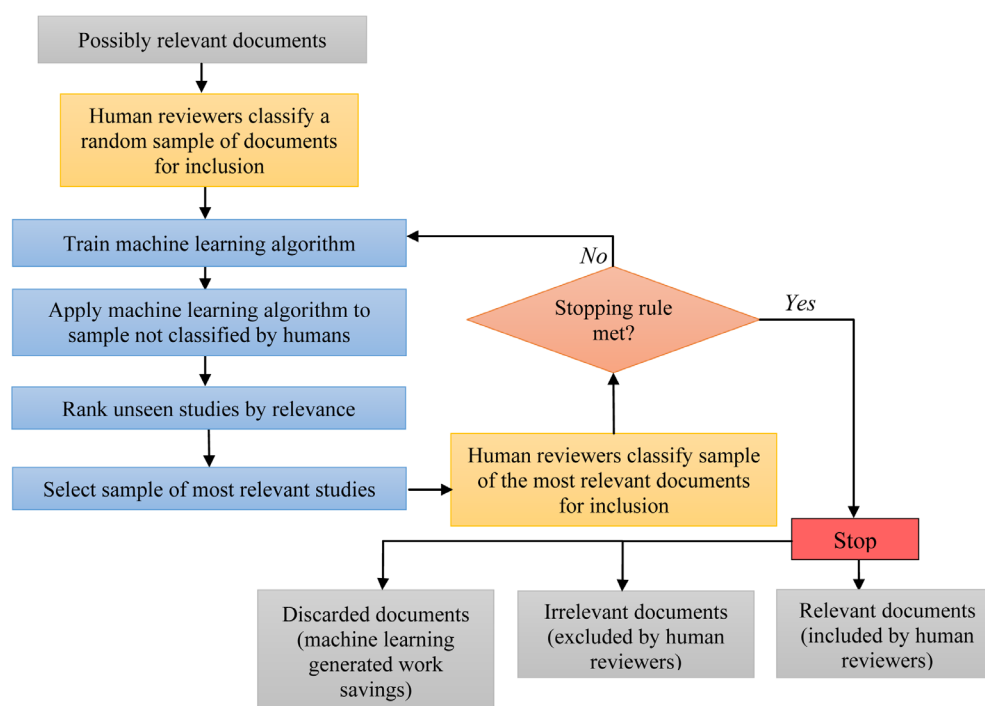


Figure 3. Flowchart illustrating the process of the machine learning assisted abstract screening (Callaghan & Müller-Hansen, 2019).

through bibliographic databases at the point of downloading their records. Tags will correspond to the specific thematic sub-block of the search strategy through which the study is identified (see column 2 in extended data (Belesova *et al.*, 2021)) and to the city name detected by the city names filter:

- (1) **field of study:** mitigation, adaptation, both;
- (2) **health impacts:** general health, nutritional, cardio- and cerebro-vascular, respiratory, renal, thermal effects, other non-communicable diseases, allergies, birth-related, developmental, accidents/injury/violence, vector-borne, helminthic, water-/food-related, other infectious diseases, health behaviours, mental disorders, other, general wellbeing, emotions, cognitive, social psychology.
- (3) **sector(-s):** buildings, transport, urban planning and greenspace, food, waste and circular economy, water, energy, health sector, unclear from the abstract;
- (4) **city:** city name(-s)

For example, if a study is identified by the mitigation search by the combination of its climate terms, terms for mitigation in the transport sector, and terms for accidents, injury and violence and its abstract or title contains the name of the city “Paris”, then it will be assigned the following tag values: field of study – mitigation, health impacts – accidents/injury/violence, sector – transport, city – Paris. If the same study is identified by

multiple combinations of search term sub-blocks, then it will be assigned all the relevant tag values (before removing its duplicate records from the database). All tag values assigned through this process will be validated by human reviewers at the abstract screenings stage for their consistency with the content of the study abstracts, titles, and keywords. For a selection of studies, the tags will also be validated against the full text versions of the papers, as described in the following section.

Stage IV: validation against full texts

Validation of decisions on study inclusion. Those articles whose inclusion could not be agreed and a random sample of the studies whose inclusion was agreed at the abstract review stage will be assessed against the inclusion/exclusion criteria by reviewing their full text versions. These will be reviewed by two independent reviewers, recording reasons for any discrepancies of the decision on inclusion based on full text review *vs* the decision based on abstract review. Any disagreements on inclusion will be discussed and agreed by consensus, involving a third reviewer, if necessary.

Tag validation. Accuracy of the meta-data contained in study tags will be verified during the full text review for the same sample of studies whose inclusion is verified at the full text level as well as any studies whose tags could not be previously credibly validated against the information contained in the

abstract. The verification will be performed by two reviewers independently and any inconsistencies will be resolved consulting a third reviewer, if necessary.

Ethics, outputs, and dissemination

The CUSSH project, of which this research is part, has overall ethical approval from the University College London Ethics Committee. Results of the database, including the distribution, variation, and diversity of actions across the world's cities, will be presented in a peer-reviewed academic paper. Furthermore, a series of protocols for systematic reviews that will be produced drawing on the database content and addressing specific research questions, including questions on the success of climate actions and their implementation strategies. Results of the subsequent systematic reviews will be published in peer-reviewed academic journals and additionally communicated to stakeholders through policy briefs and other preferred dissemination formats.

Study status

The study protocol and search strategy have been completed and searches have been performed; as of publication, screening and subsequent processes have not been started yet.

Discussion

This protocol described the systematic approach for the development of the first comprehensive database of peer-reviewed studies reporting climate change mitigation and adaptation actions of relevance to human health and wellbeing that have been implemented in cities.

A unique aspect of the database is the innovative use of machine learning methods. All approaches to systematic literature searching have their limitations (Berrang-Ford *et al.*, 2021) but our deployment of machine learning to enhance the process is designed to make it possible to screen a larger number of studies than is usually feasible through manual review alone and to improve the consistency of the selection process. This is particularly important in our case because urban climate change adaptation and mitigation measures can take many forms and thus require an extensive set of search terms and judgement on the relevance of a disparate array of interventions and study types. While there is a substantial gain in efficiency in terms of the proportion of articles that need to be reviewed manually, a large investment of effort is still needed for the supervised training of the ML algorithm and the repeated checking of results. We are still exploring the most efficient ways to use the ML methods, including the stopping rules for the supervised training and how much further review is undertaken of the studies classified by the fully trained ML algorithm. We will also assess the overall quality of the search using random samples of studies and use the results to guide the refinement of future searches. Without wishing to pre-judge the results, our hope and expectation is that the support of machine learning will allow our review to be more comprehensive than any similar initiative to date.

Our search strategies capture climate change adaptation and mitigation actions, including relevant adaptation and mitigation

actions that are not explicitly identified in those terms, as well as studies reporting physical, mental and social health and wellbeing. Mental health and wellbeing have been under-represented in the literature on climate change and health but represent a substantial health burden where they have been assessed (Berry *et al.*, 2018; Hayes *et al.*, 2018).

We acknowledge that the database is subject to a range of possible sources of bias that are characteristic of systematic reviews:

- **Reviewer bias** can originate from different interpretation of inclusion and exclusion criteria by reviewers, which can also introduce bias into the article classification by the machine learning algorithm. To minimise the risk of such bias we will train all reviewers in consistent application of inclusion and exclusion criteria prior to the start of the machine learning algorithm training. Furthermore, each abstract classification will be performed by two reviewers independently and reconciled with a third reviewer in cases of any disagreement.
- **Publication bias** is likely to be a limitation of the database, as studies on climate change mitigation and adaptation actions in cities tend to be more frequently conducted in high-income countries than low- and middle-income countries, introducing geographical publication bias. The geographical publication bias is likely to be further exacerbated by our language and publication type inclusion criteria. In low- and middle-income countries there is likely to be a higher fraction of grey literature reports on urban climate action and health in local languages than peer-reviewed scientific studies in English. Furthermore, actions that are perceived as successful are more likely to be reported in academic literature than actions that did not have a marked perceived effect. This is likely to introduce another dimension of likely publication bias in our database.
- **Inconsistent definitions** across studies may add further bias, as, for example, study authors may use such terms as 'climate action' differently. To minimise bias that can result from inconsistent definitions, we set our own broad definitions for the key terms whose definitions are particularly likely to vary across studies. Any data entry and study classification in the development of our database will be based on these unifying definitions.

The publication bias and strength of the body of evidence will be formally assessed as a part of the subsequent systematic reviews using the content of this database.

The database will have a broad and comprehensive scope with the aim of creating a representative evidence base for the subsequent systematic reviews and flexibility to address new research questions. Only including studies published in English may seem an important limitation. However, we tested our search strategy on studies published in languages other

than English (Spanish, Chinese, French, German, Russian, and Arabic) but with an abstract available in English. We found that such studies added only 2% to our overall search results. It is possible that a higher number of studies in other languages would be retrieved if our search strategy was fully translated into these languages. However, in light of resource constraints we took the decision to exclude languages other than English.

We will continue to refine the methods described here as we begin to exploit and learn from the database. Future planned developments of machine learning capabilities include development of searches through further bibliographic databases, including grey literature databases, search engines such as Google Scholar and organisational websites, and consideration of approaches to covering more languages. Together with stakeholders such as C40, we intend to develop a strategy for curation and regular updating of the database and mechanisms for accessing its content. We welcome suggestions for further elaboration and development of the strategy, technical aspects and applications of our database, including suggestions for joint research work.

Conclusion

This protocol describes an approach for the systematic development of a database of climate change mitigation and adaptation actions relevant to human health and wellbeing. To address the challenge of the broad landscape of published research on such actions across multiple disciplines, sectors, and complex systems, we exploit machine learning methods applied to specialised search strategies to develop a comprehensive and updatable collection of relevant peer-reviewed studies classified by their key characteristics. The database is intended to serve as a source for subsequent systematic

reviews addressing specific research questions of relevance to stakeholders in urban climate action and health. The outputs of the systematic reviews will be of value to the scientific community, international networks on city climate action and leadership, such as C40 cities network, urban policy makers, and other stakeholders.

Data availability

Underlying data

There is no data associated with this article.

Extended data

LSHTM Data Compass: Search strategies for: “Climate action for health and wellbeing in cities: a protocol for the systematic development of a database of peer-reviewed studies using machine learning methods”. <https://doi.org/10.17037/DATA.00002094> (Belesova *et al.*, 2021)

Data are available under the terms of the [Creative Commons Attribution 3.0 Unported](#) (CC-BY 3.0).

Acknowledgements

We would like to acknowledge Ms Jane Falconer librarian at the London School of Hygiene and Tropical Medicine for completing the PRESS review of the systematic search strategy. We would also like to thank Dr Jamie Anderson for advice on wellbeing concepts, Dr Simon Lloyd for initial advice on the search structure, Dr Helen Pineo for reviewing building sector mitigation search terms, Dr Phil Symonds for reviewing climate and energy search terms, and Dr Ai Milojevic for taking part in initial discussions on the conception of this work.

References

- Aleksandrowicz L, Green R, Joy EJM, *et al.*: **The impacts of dietary change on greenhouse gas emissions, land use, water use, and health: A systematic review.** *PLoS One.* 2016; **11**(11): e0165797.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Belesova K, Callaghan M, Minx J, *et al.*: **Climate action for health and wellbeing in cities: a protocol for the development of a systematic database of peer-reviewed studies using machine learning methods.** *OSF Registries.* 2020a.
[Publisher Full Text](#)
- Belesova K, Hale J, Alpin J, *et al.*: **Complex Urban Systems for Sustainability and Health (CUSSH): A Structured Approach to Support the Development and Implementation of City Policies for Population and Planetary Health.** In *Proceedings of WHO International Healthy Cities Conference.* Belfast: WHO. 2018; 1–6.
[Reference Source](#)
- Belesova K, Callaghan M, Minx J, *et al.*: **Search strategies for: “Climate action for health and wellbeing in cities: a protocol for the systematic development of a database of peer-reviewed studies using machine learning methods.”** London School of Hygiene and Tropical Medicine, London, United Kingdom. 2021.
<http://www.doi.org/10.17037/DATA.00002094>
- Belesova K, Heymann DL, Haines A: **Integrating climate action for health into covid-19 recovery plans.** *BMJ.* 2020b; **370**: m3169.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Berrang-Ford L, Sietsma A J, Callaghan M, *et al.*: **Systematic Mapping of Global Research on Climate and Health Using Machine Learning.** *SSRN Electronic Journal.* 2021; 1-12.
[Publisher Full Text](#)
- Berry HL, Waite TD, Dear KBG, *et al.*: **The case for systems thinking about climate change and mental health.** *Nat Clim Chang.* 2018; **8**(4): 282–290.
[Publisher Full Text](#)
- C40, Arup: **Deadline 2020.** 2017.
[Reference Source](#)
- C40 Cities, Arup: **Climate Action in Megacities 3.0.** 2015.
[Reference Source](#)
- Callaghan M, Lee YT, Müller-Hansen F, *et al.*: **NACSOS: NLP Assisted Classification, Synthesis and Online Screening.** *Zenodo.* 2020.
[Publisher Full Text](#)
- Callaghan MW, Müller-Hansen F: **Robust Statistical Stopping Criteria for Automated Screening in Systematic Reviews.** *PREPRINT (Version 1) Available at Research Square.* 2019; 1–16.
[Publisher Full Text](#)
- Clark C, Paunovic K: **Who environmental noise guidelines for the European**

region: A systematic review on environmental noise and quality of life, wellbeing and mental health. *Int J Environ Res Public Health.* 2018; **15**(11): 2400.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Cochrane Common Mental Disorders: **Search strategies for the identification of studies Searching.** 2014.

[Reference Source](#)

Creutzig F, Roy J, Lamb WF, *et al.*: **Towards demand-side solutions for mitigating climate change.** *Nat Clim Chang.* 2018; **8**(4): 260–263.

[Publisher Full Text](#)

Dodge R, Daly AP, Huyton J, *et al.*: **The challenge of defining wellbeing.** *International Journal of Wellbeing.* 2012; **2**(3): 222–235.

[Publisher Full Text](#)

Forster PM, Forster HI, Evans MJ, *et al.*: **Erratum: Publisher Correction: Current and future global climate impacts resulting from COVID-19.** *Nat Clim Chang.* 2020; **1**.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Gao J, Kovats S, Vardoulakis S, *et al.*: **Public health co-benefits of greenhouse gas emissions reduction: A systematic review.** *Sci Total Environ.* 2018; **627**(155): 388–402.

[PubMed Abstract](#) | [Publisher Full Text](#)

GeoNames Team: **GeoNames database.** 2020.

Grieneisen ML, Zhang M: **The current status of climate change research.** *Nat Clim Chang.* 2011; **1**(2): 72–73.

[Publisher Full Text](#)

Hayes K, Blashki G, Wiseman J, *et al.*: **Climate change and mental health: Risks, impacts and priority actions.** *Int J Ment Health Syst.* 2018; **12**(1): 28.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

IPCC: **Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.** In C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, ... P. M. Midgley (Eds.): *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change (IPCC)* (pp. 555–567). Cambridge, UK, and New York, NY, USA: Cambridge University Press. 2012.

[Reference Source](#)

Keynejad RC, Hanlon C, Howard LM: **Psychological interventions for common mental disorders in women experiencing intimate partner violence in low-income and middle-income countries: a systematic review and meta-analysis.** *Lancet Psychiatry.* 2020; **7**(2): 173–190.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Lamb WF, Creutzig F, Callaghan MW, *et al.*: **Learning about urban climate solutions from case studies.** *Nat Clim Chang.* 2019; **9**(4): 279–287.

[Publisher Full Text](#)

Larvin H, Peckham E, Prady SL: **Case-finding for common mental disorders in primary care using routinely collected data: a systematic review.** *Soc Psychiatry Psychiatr Epidemiol.* 2019; **54**(10): 1161–1175.

[PubMed Abstract](#) | [Publisher Full Text](#)

Lesnikowski AC, Ford JD, Berrang-Ford L, *et al.*: **National-level factors affecting planned, public adaptation to health impacts of climate change.** *Glob Environ Change.* 2013; **23**(5): 1153–1163.

[Publisher Full Text](#)

Lesnikowski AC, Ford JD, Berrang-Ford L, *et al.*: **Adapting to health impacts of climate change: A study of UNFCCC Annex I parties.** *Environ Res Lett.* 2011; **6**(4): 044009.

[Publisher Full Text](#)

Li C, Zhou J, Cao Y, *et al.*: **Interaction between urban microclimate and electric air-conditioning energy consumption during high temperature season.** *Appl Energy.* 2014; **117**: 149–156.

[Publisher Full Text](#)

Marsh HW, Huppert FA, Donald JN, *et al.*: **The well-being profile (WB-Pro): Creating a theoretically based multidimensional measure of well-being to advance theory, research, policy, and practice.** *Psychol Assess.* 2020; **32**(3): 294–313.

[PubMed Abstract](#) | [Publisher Full Text](#)

McGowan J, Sampson M, Salzwedel DM, *et al.*: **PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement.** *J Clin Epidemiol.* 2016; **75**: 40–46.

[PubMed Abstract](#) | [Publisher Full Text](#)

Pedregosa F, Varoquaux G, Gramfort A, *et al.*: **Scikit-learn: Machine Learning in Python.** *JMLR.* 2012; **12**(85): 2825–2830.

[Reference Source](#)

Quéré CL, Jackson RB, Jones MW, *et al.*: **Temporary reduction in daily global CO2 emissions during the COVID-19 forced confinement.** *Nat Clim Chang.* 2020; **10**: 647–653.

[Publisher Full Text](#)

Reckien D, Flacke J, Dawson RJ, *et al.*: **Climate change response in Europe: what's the reality? Analysis of adaptation and mitigation plans from 200 urban areas in 11 countries.** *Clim Change.* 2014; **122**(1): 331–340.

[Publisher Full Text](#)

Seto KC, Dhakal S, Bigio A, *et al.*: **Human Settlements, Infrastructure, and Spatial Planning.** In T. Z. and J. C. M. Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow (Ed.). *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 2014; 923–1000.

[Publisher Full Text](#)

Shrubsole C, Macmillan A, Davies M, *et al.*: **100 Unintended consequences of policies to improve the energy efficiency of the UK housing stock.** *Indoor Built Environ.* 2014; **23**(3): 340–352.

[Publisher Full Text](#)

Smith KR, Woodward A, Campbell-Lendrum D, *et al.*: **Human Health: Impacts, Adaptation, and Co-Benefit.** In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, ... L. L. White (Eds.). *Climate Change 2014: Impacts, Adaptation, and Vulnerability.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 2014; 709–754.

[Publisher Full Text](#)

Wenz L, Levermann A, Auffhammer M: **North-south polarization of European electricity consumption under future warming.** *Proc Natl Acad Sci U S A.* 2017; **114**(38): E7910–E7918.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

West JJ, Smith SJ, Silva RA, *et al.*: **Co-benefits of mitigating global greenhouse gas emissions for future air quality and human health.** *Nat Clim Chang.* 2013; **3**(10): 885–889.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Whitmee S, Haines A, Beyrer C, *et al.*: **Safeguarding human health in the Anthropocene epoch: Report of the Rockefeller Foundation-Lancet Commission on planetary health.** *Lancet.* 2015; **386**(10007): 1973–2028.

[PubMed Abstract](#) | [Publisher Full Text](#)

Wilczynski NL, Haynes RB, Hedges T: **Optimal search strategies for identifying mental health content in MEDLINE: An analytic survey.** *Ann Gen Psychiatry.* 2006; **5**: 4.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

World Health Organization: **Constitution of the World Health Organization.** Geneva: WHO. 1948.

[Reference Source](#)

World Health Organization (WHO): **International Statistical Classification of Diseases and Related Health Problems 10th Revision.** 2019.

[Reference Source](#)