**Overview:**
This living map (the Map) consists of bibliographic records of research articles on COVID-19 published since 2019.a

The Map is now maintained using a fully automated workflow, hosted in [EPPI-Reviewer Web (ER-Web)](https://eppi.ioe.ac.uk/CMS/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4&).

**Using this map:**
There are various ways to view the map content.

***Quick Start:***
All the bibliographic records in the Map have been tagged (or 'keyworded') with one 'Topic' code, listed in the first expandable list on the left. To list all records of 'Treatment evaluation' studies, open the 'Topic' heading, click on 'Treatment evaluation' and then on the button above 'List records'. Select 'Home' to return to the previous screen. To see how many records are in each category under the 'Topic' code, select 'Topic', and click on the 'Frequencies' button above it. This will open a table in the middle of the screen that tells you how many records have been categorised with each heading. You can then list the records by clicking on the corresponding number. Similarly, all records have one version code, as well as being tagged with 'All versions'. To see how many records are in each version of the Map, select 'Version' or 'Previous versions' and click on the 'Frequencies' button. Clicking on each number will display that set of records.

***Downloading data and exploring individual records:***
The search features described above (and below) will result in a list of bibliographic records being displayed. This list can be downloaded as a plain text file, in Excel format, or as a RIS file for importing into reference manager software (such as Zotero or EndNote). Clicking an individual title in the list will result in the detailed information about that record being displayed. This defaults to standard bibliographic information plus the abstract, but can be expanded to include all bibliographic fields in the database. This screen also contains a 'show coding' button, which opens up the display to show the 'Topic' and 'Version' codes assigned to that specific record.

***Displaying a map or cross-tabulation:***
Crosstabulation operates using the grouped headings of codes. To show a table of all records with topic across the top and version down the side, click on the heading 'Topic' and, at the bottom right of the screen, click on the 'Set X axis' button; then click on the 'Version' heading (on the left of the screen; or Previous versions... below Version) and click on the 'Set Y axis' button; finally, click on the 'Get Crosstab' button (bottom right of the screen). The resulting page will display a matrix showing the intersections of the categories under these two headings. By clicking on the 'Bubble map' button, this matrix can be changed from a table to a bubble map, with bubbles indicating the relative number of records in each cell (similar to the display in EPPI-mapper). The numbers / bubbles in the cells are clickable, and clicking on them will display a list of the records in that cell (below the table or bubble map - this may take some time to appear).

***Finding a specific record or set of records:***
There is a free-text search at the top of the 'Home' screen. This defaults to searching the title and abstract fields, but specific fields can be selected using the drop-down menu next to it.

**Identifying the evidence**:

***Searches 1-34***Prior to implementing the *OpenAlex*-enabled (formerly *MAG*-enabled) workflow described below (see Versions 35 to 44), evidence in this Map was identified by conducting weekly searches of MEDLINE and Embase, beginning on Wednesday 4th March 2020, and updating the search strategy as necessary.

MEDLINE search strategy at October 2020:

Database: Ovid MEDLINE(R) ALL <1946 to October 01, 2020>
Search Strategy:
--------------------------------------------------------------------------------
1 ("20200925" or "20200926" or "20200927" or "20200928").dt. (14732)
2 preprint.pt (1048)
3 1 not 2 (14732)
4 limit 3 to covid-19 (1119)

The Embase search strategy at October 2020:

Database: Embase <2016 to 2020 Week 40>
Search Strategy:
--------------------------------------------------------------------------------
1 "202040".em. (120134)
2 limit 1 to covid-19 (5705)

For further details of these MEDLINE/Embase search strategies, please see current OVID Covid-19 Expert Searches developed by Wolters Kluwer, available from: https://tools.ovid.com/ovidtools/expertsearches.html#corona - for example (October 2020):

Coronavirus (Covid-19) 2019-nCoV on MEDLINE1. exp Coronavirus/
2. exp Coronavirus Infections/
3. (coronavirus\* or corona virus\* or OC43 or NL63 or 229E or HKU1 or HCoV\* or ncov\* or covid\* or sars-cov\* or sarscov\* or Sars-coronavirus\* or Severe Acute Respiratory Syndrome Coronavirus\*).mp.
4. (or/1-3) and ((20191\* or 202\*).dp. or 20190101:20301231.(ep).) [this set is the sensitive/broad part of the search]
5. 4 not (SARS or SARS-CoV or MERS or MERS-CoV or Middle East respiratory syndrome or camel\* or dromedar\* or equine or coronary or coronal or covidence\* or covidien or influenza virus or HIV or bovine or calves or TGEV or feline or porcine or BCoV or PED or PEDV or PDCoV or FIPV or FCoV or SADS-CoV or canine or CCov or zoonotic or avian influenza or H1N1 or H5N1 or H5N6 or IBV or murine corona\*).mp. [line 5 removes noise in the search results]
6. ((pneumonia or covid\* or coronavirus\* or corona virus\* or ncov\* or 2019-ncov or sars\*).mp. or exp pneumonia/) and Wuhan.mp.
7. (2019-ncov or ncov19 or ncov-19 or 2019-novel CoV or sars-cov2 or sars-cov-2 or sarscov2 or sarscov-2 or Sars-coronavirus2 or Sars-coronavirus-2 or SARS-like coronavirus\* or coronavirus-19 or covid19 or covid-19 or covid 2019 or ((novel or new or nouveau) adj2 (CoV on nCoV or covid or coronavirus\* or corona virus or Pandemi\*2)) or ((covid or covid19 or covid-19) and pandemic\*2) or (coronavirus\* and pneumonia)).mp.
8. COVID-19.rx,px,ox. or severe acute respiratory syndrome coronavirus 2.os.
9. ("32240632" or "32236488" or "32268021" or "32267941" or "32169616" or "32267649" or "32267499" or "32267344" or "32248853" or "32246156" or "32243118" or "32240583" or "32237674" or "32234725" or "32173381" or "32227595" or "32185863" or "32221979" or "32213260" or "32205350" or "32202721" or "32197097" or "32196032" or "32188729" or "32176889" or "32088947" or "32277065" or "32273472" or "32273444" or "32145185" or "31917786" or "32267384" or "32265186" or "32253187" or "32265567" or "32231286" or "32105468" or "32179788" or "32152361" or "32152148" or "32140676" or "32053580" or "32029604" or "32127714" or "32047315" or "32020111" or "32267950" or "32249952" or "32172715").ui. [Articles not captured by this search when created in April 2020, pending further indexing by NLM]
10. or/6-9 [Lines 6 to 9 are specific to COVID-19]
11. 5 or 10
12. 11 and 20191201:20301231.(dt).
13. remove duplicates from 12

From 20th July 2020 (Search 20) until 26th October 2020 (Search 34), all unique (following de-duplication) 'new' MEDLINE/Embase records were also scored using a binary machine learning (ML) classifier described below (see Versions 35 to 44). MEDLINE/Embase records scoring above an identified threshold score were retained for screening; while those scoring below this threshold score were set aside**.**

***Versions 35 to 44***From 9th November 2020 (Version 35) onwards, we stopped searching MEDLINE and Embase (see Searches 1-34, above) and began to identify the evidence using automated continuous prospective surveillance of the *Microsoft Academic Graph* (*MAG*) dataset (knowledge graph) [1]. A *MAG*-enabled workflow was operationalised using the new *MAG Browser* suite of tools in [*EPPI-Reviewer Web*](https://eppi.ioe.ac.uk/CMS/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4&) (*ER-Web*) [2]. The full *MAG* dataset, at this time, comprised >245 million bibliographic records of research articles on all topics across science, connected in a large network graph of conceptual, citation, author and other relationships [3].

Each time an updated copy of the *MAG* dataset was released by *Microsoft*TM (initially, every two weeks), all 'new' *MAG* records (i.e. records of articles not indexed in any preceding versions of the *MAG* dataset - up to one million new records per update) and their associated metadata were automatically imported into *MAG Browser*(*ER-Web*) systems. New *MAG* records were then automatically scored by our novel *Auto-Update* machine learning (ML) ‘recommender’ model.[b] The *Auto-Update* model exploits both network graph features [3] and text features of new *MAG*records (with reference to the same features of known include *MAG* records identified and coded in preceding versions of this Map) to score and prioritise (by ranking them from highest to lowest, by score) the new records for potential manual screening-coding (see Coding the evidence, below). Preprints, and articles from specific sources[c] that are invariably excluded from this Map, were automatically filtered out and discarded. We then retained the top scoring (on *Auto-Update*) new, filtered *MAG* records, between 3,000 and 10,000 per *MAG*update, contingent on the time elapsed since the preceding update and the total number of new records included in each update. From August 2020 (Version 68) onwards, we also deployed our binary ML classifier (see below) to score these new *MAG* records and automatically exclude those scoring below the identified threshold score (see below), before importing the remaining records (scoring at the threshold or above) into ER-Web.

Next, we used ER-Web de-duplication tools to identify and remove duplicate new *MAG* records. Then we re-scored the remaining top-scoring new *MAG* records (i.e. highest ranked by our *Auto-Update* model) using a binary ML classifier that we designed to distinguish between title-abstract records included in (positive class), and those excluded from (negative class), this Map[d]. New *MAG* records scoring above an identified threshold score on the binary ML classifier were retained; while those scoring below this threshold score were set aside. Finally, we retained the remaining new *MAG* records for potential manual screening and coding (see Coding the evidence).

***Versions 45 to 79***From Version 45 onwards, we supplemented the set of top-scoring new records (on the *Auto-Update* model) with a second set of *MAG* records, from each update of the *MAG* dataset, identified using a COVID-19 'custom search' that we developed and executed using *MAG Browser* tools in *ER-Web*.

OR(And(W='severe',W='acute',W='respiratory',W='syndrome',W='coronavirus'),And(W='coronavirus',W='19'),And(W='coronavirus',W='2019'),And(W='covid',W='19'),And(W='covid',W='2019'),W='covid19',And(W='2019',W='ncov'),And(W='middle',W='east',W='respiratory',W='syndrome',W='coronavirus'),And(W='corona',W='virus',W='disease',W='2019'),And(W='new',W='coronavirus'),And(W='novel',W='coronavirus'),And(W='sars',W='cov2'),And(W='sars',W='cov',W='2'),And(W='sars',W='coronavirus',W='2'),Composite(F.FId=3008058167),Composite(F.FId=3007834351),Composite(F.FId=3006700255))

We restricted this custom search to *MAG* records, in each update of the *MAG* dataset, with publication dates on or after 6th July 2020, and (automatically) only imported those records that had not previously been either (a) imported into *ER-Web* from *MAG* or (b) matched from its corresponding MEDLINE/Embase record. These records were then processed into our Priority Screening workflow in precisely the same way as the set of top-scoring new *MAG* records on our *Auto-Update* model (see 'Versions 35 to 44', above), except that we have periodically updated our ML Classifier, until its final version (up to Version 97 - March 2022).

***Versions 80 to 89***With each *MAG* update, four sets of records were imported. The first set was identified using an *Auto-Update* search based on COVID-19 included records. Records were discarded using the ML Classifier, and a key-term search (those that do not contain one of these key terms: 'COVID' OR ‘COVID19’ OR ‘nCov’ OR 'coronavirus' OR 'corona virus’ OR 'pandemic' OR ‘SARS’ OR ‘Moderna’ OR 'Ad26.COV2.S’ OR ‘BNT162b2'). The second set was the custom search, described above, and limited to the previous six months, with records discarded in the same way as for the first set. The third set was an *Auto-Update* search based on records included in the Long COVID section of the Map, with records discarded using the Long COVID Classifier and the key-term search. The fourth set is a network graph search, which automatically retrieves all records that appear in the forwards (cited by) or backwards (cites) citation networks, and/or in the forwards (recommended by) and/or backwards (recommends) ‘related publications’ networks of records included in the Long COVID ‘segment’ of the Map, with records discarded by the Long COVID Classifier, the key-term search and a search for pre-prints and other excluded items.

***Version 90 to date***From January-February 2022, we switched over to using automated searches of the [*OpenAlex* dataset](https://openalex.org/) [6] - which replaced and superseded the *MAG* dataset at the end of 2021 (i.e. the *OpenAlex* dataset incorporates and maintains the *MAG* dataset (knowledge graph)). *OpenAlex* searches are conducted using our re-engineered [*OpenAlex Browser* tools](https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3819) (which have replaced *MAG Browser* tools) in ER-Web. The types of *OpenAlex* searches used to maintain the Map (*Auto-Update* searches x2, custom search and network graph search) remain the same as those used for the *MAG* dataset (see Versions 80 to 89).

**Coding the evidence**:

***Versions 30 to 97***
From 28th September 2020 (Version 30) to March 2022 (Version 97), screening and coding of new retained MEDLINE/Embase records (up to Search 34) or *MAG* records (from Version 35 onwards) was conducted using priority screening mode in ER-Web. In priority screening mode, retained records (i.e. top-scoring on the *Auto-Update* model and above the threshold score on the binary ML classifier - see Identifying the evidence) were screened in prioritised rank order (highest to lowest) based on scores assigned by the binary ML classifier; and the rank order of those records awaiting screening was periodically reprioritised based on all preceding coding decisions (i.e. active learning [4, 5]).

From 28th September 2020 (Version 30) to March 2022 (Version 97), each team member had a fixed target to screen and code to reach a total of 1,500 records, each week, using priority screening mode. All retained records that were not screened-coded by the team in a given week were carried forward, along with new records from the next updated version of the *MAG* dataset (or, for Searches 30 to 34, from the next MEDLINE-Embase searches), to the pool of records to be reprioritised for screening-coding during a subsequent week. The option of referring selected records for a second opinion remained in the *MAG*-enabled workflow, and second opinion records were still resolved by team discussion and consensus.

From 26th November 2021 (Version 83 - *MAG* dataset updated on 8th November) onwards, we applied our Bidirectional Encoder Representations from Transformers (BERT) ‘COVID-19 Categories’ model – which we trained, calibrated and evaluated on records manually coded for the Map – to assign imported records to topic codes, if they scored above the probability thresholds that were calibrated to achieve over 0.95 accuracy across all codes. All remaining records were added to the pool of records for manual screening and coding, in priority-screening mode, as above.

***Version 98 to date***
From March 2022 (Version 97) onwards, we have discontinued manual coding (see below). We continue to import records when the *OpenAlex* dataset is updated, process them in the same way described above, and upload those records coded as ‘included’ (i.e. assigned to one of eleven ‘included’ Topic codes) by the BERT model.

***Manual Coding***

Criteria used to manually code evidence for inclusion in the Map under each category heading (topic code) were as follows.

Primary empirical data, systematic review,\* modelling,\*\* full economic evaluation,\*\*\* or novel analysis on COVID-19.

* 1. Treatment evaluation
		+ Any intervention aimed at treatment, prevention and/or rehabilitation of COVID-19 (i.e. with either a population of COVID-19 patients or COVID-19 incidence as an outcome), including vaccines
		+ Prospective outcome studies with comparison between researcher-allocated groups (i.e. randomised trials, quasi-randomised trials, and non-randomised trials with researcher allocation)
		+ Include systematic reviews\* that aimed to include studies meeting these criteria, whether or not any were located
		+ Exclude observational/retrospective studies including treatment as an exposure and uncontrolled studies (code as Treatment development)
		+ Exclude case reports with some information about treatment (code as Case reports)
		+ Exclude basic science with claimed relevance to interventions, but without evaluation of effectiveness in human patients (code as Treatment development or Vaccine development)
	2. Transmission / risk / prevalence
		+ Epidemiological modelling of spread (incl. studies which aim to model health outcomes or health system outcomes; include population mortality rates (i.e. deaths relative to total population); exclude case fatality rates (i.e. deaths relative to COVID-19 cases)); include genetic epidemiology if the main focus is on spread of disease (if the focus is on characterising strains, code as Genetics / biology)
		+ Risk modelling
		+ Studies of viral persistence in bodily secretions/tissues or in the environment, e.g. on surfaces (including methods to inactivate the virus in these contexts)
		+ Population prevalence studies (including seroprevalence)
		+ Studies of risk factors for developing COVID-19 at individual level (not risks of developing more severe disease/complications among people infected with COVID-19; code as Health impacts) or at population level
		+ Studies of the effectiveness of non-drug prevention strategies e.g. masks, contact tracing (only if data on prevalence reported). Include modelling studies of impacts of vaccines on prevalence at population level.
		+ Exclude data on preventive behaviour outcomes only (code as Social / economic / indirect impacts). Exclude studies of disinfection, aerosolisation etc. if they do not include data on COVID-19 (code as Not on topic)
	3. Diagnosis
		+ Sensitivity and specificity of tests for COVID-19 (including antibody tests)
		+ Training clinicians in diagnosis
		+ Include studies of clinical signs if the main focus is on their diagnostic value (if the focus is on health outcomes or prognosis, code under Health impacts)
		+ Exclude diagnosis of other conditions/comorbidities in COVID-19 patients
	4. Health impacts of COVID-19
		+ Any observational study with a population of COVID-19-infected patients measuring physical health outcomes (incl. case fatality rates, QALYs or DALYs) and/or somatic indicators (code studies reporting prevalence data here if they also present health impacts)
		+ Include studies of prognostic factors, indicators of disease progress or severity
		+ Studies of comorbidities (e.g. coinfections), if not explicitly analysed as risk factors for infection
	5. Vaccine development
		+ Basic science aimed at development of vaccines
		+ Include animal studies testing human vaccines only
		+ Studies looking at vaccines but not meeting methodological criteria for Treatment evaluation, e.g. observational/retrospective studies including vaccine receipt as an exposure (even if measuring prevalence as an outcome), studies without researcher-allocated control group, or pre-post studies of antibody response
		+ Studies of vaccine safety/side-effects
		+ (Studies of vaccine hesitancy/intentions to be vaccinated/attitudes towards vaccination - code under Social / economic / indirect impacts)
	6. Treatment development
		+ Basic science aimed at development of treatment, e.g. drug discovery (including *in silico*molecular docking studies)
		+ Include animal studies testing human treatments only
		+ Studies looking at treatments but not meeting methodological criteria for Treatment evaluation, e.g. observational/retrospective studies including treatment as an exposure, studies without researcher-allocated control group, or modelling based on evaluation data (but exclude studies of outcomes which simply state that treatment was administered without relating outcomes to treatments - code under Health impacts)
		+ Studies of treatment safety/side-effects
		+ Training clinicians to deliver interventions
		+ Include studies of drug treatments used to prevent infection
		+ (Studies of relevance to both vaccine and treatment development - code under Treatment development)

* 1. Genetics / biology of virus
		+ Any data on the genetic or biological characteristics of the virus, or of mechanisms or responses to infection (including antibody responses or humoral immunity, if not clearly aimed at diagnosis/vaccine development)
		+ Include modelling on the basis of secondary data analysis
		+ Exclude studies of biological mechanisms theoretically linked to COVID-19 infection, but without data which actually concern COVID-19
		+ If explicitly aimed at treatment (resp. vaccine) development, code as Treatment (resp. Vaccine) development; if quantifying seroprevalence, code as Transmission / risk / prevalence
	2. Case reports - patients
		+ Medical case reports of small numbers of patients considered as individuals
		+ Include any case with confirmed COVID-19 or symptoms or history suggestive of COVID-19 infection (otherwise code as Not on topic, case studies of health professionals and mental health consequences of lockdown are also not on topic)
		+ Include mental health cases tested using the "Fear of COVID" scale or equivalent
		+ Include case reports of adverse effects of vaccines
	3. Case study - organisation
		+ Descriptive studies setting out organisational responses/strategies to COVID-19
		+ Surveys of professionals/institutions on organisational responses (not broader knowledge or attitudes to COVID-19 - code these as Social / economic / indirect impacts); any studies focused on service delivery e.g. performance of clinical procedures pre- and post-COVID-19 (or delivery of treatments, if no outcome data)
		+ Include any organisation (healthcare or other) and any form of response to COVID-19, whether directly concerning COVID-19 patients or not
		+ Exclude guidance or recommendation papers which do not describe the recommended measures being implemented in a specific case
	4. Social / economic / indirect impacts
		+ Include studies mainly focusing on behaviour, attitudes etc.
		+ Studies of information (e.g. analysis of websites or social media)
		+ Surveys of professionals if not mainly focused on organisational responses
		+ Studies of behaviour or health outcomes of patients without diagnosed COVID-19 (including total excess mortality, unless separable data on COVID-19 mortality are available)
		+ Studies of other impacts of COVID-19 or COVID-19 control measures (e.g. environmental impacts of lockdown)
		+ Include indirect health impacts on healthcare workers (e.g. from PPE use)
		+ Studies of vaccine hesitancy/intention to be vaccinated, vaccine uptake/coverage
		+ Studies of access to services (vaccination, treatment, testing, etc.) for COVID-19
		+ Indirect impacts of diagnostic procedures (e.g. radiation)
	5. Mental health impacts
		+ Include both COVID-19 patients and/or indirect mental health impacts on the broader population (or healthcare workers, etc.)
		+ Include mental health status (anxiety, depression, etc.) and sleep-related outcomes
		+ Where studies have an equal focus on mental health impacts and health and/or indirect impacts, code as mental health impacts
		+ Exclude if there are no COVID-related measures *and* only one time point (no before-and-during or -after comparison)

\* Define systematic review as any paper reporting secondary data which reports: some search terms; clearly defined inclusion criteria; and some information on the selection process (at least the number of references located by searches and the number of studies included). Include any systematic review which aimed to include studies on COVID-19, whether or not any were located. Include updates to systematic reviews and living reviews if the report presents new data and the original review meets the criteria above.

\*\* Include modelling studies which are at least partly based on empirical data related to COVID-19 (e.g. data used as inputs to the model, or data against which the model is being calibrated or tested); code purely theoretical modelling as not primary data.

\*\*\* Include full economic evaluations (i.e. cost-effectiveness analyses, cost-minimisation analyses, cost-utility analyses, or cost-benefit analyses - see <https://yhec.co.uk/resources/glossary/>). Include model-based and single-study based economic evaluations. Code topic based on the main focus/ aim of the study (e.g. code cost-effectiveness analyses of clinical treatments or management strategies for COVID-19 as Treatment evaluation).

In general, code using the main aim (or the main focus) of the paper if it covers more than one topic. Code systematic reviews by the inclusion criteria or the focus of the included papers.

Two exclude codes were originally displayed in the Map, but since 21st April 2020 these are no longer shown. Since 12th February, these have been combined into one exclude code.

1. Other viruses (SARS, MERS, etc.)
	* Anything on human coronaviruses other than COVID-19; include both primary data and non-data papers
2. No primary empirical data, systematic review or modelling
	* Thinkpieces, non-systematic reviews, guidance, consensus statements, hypotheses, etc.
	* Protocols for studies or reviews which do not report findings data
	* Systematic reviews which do not report findings data (mapping reviews; reviews which only contain guideline documents / opinion pieces)
	* Methods papers (including validation of data collection methods if usable primary data not reported)
	* Corrections, errata, retractions
	* Responses or replies which do not report substantive new data or analysis
	* Items in data repositories (e.g. Protein Data Bank, Mendeley Data, OSF), and patents

The remaining excludes were all not on topic, pre-prints (that will be published in a journal, if they pass peer review, including all Rxiv, Research Square, Authorea, and SSRN articles), or duplicates identified while screening and coding.

To access a RIS file for any of these codes, please email: theo.lorenc@york.ac.uk

**Results:**

For results, see: <https://eppi.ioe.ac.uk/eppi-vis/login/open?webdbid=7>

**Endnotes:**

a Primary empirical data, systematic review, modelling, full economic evaluation, or novel analysis on COVID-19 - see Coding the evidence.

b Our ContReview model was built and tested (in collaboration with MicrosoftTM) using *MAG* records of COVID-19 research articles that we had matched to records included in the Map up to Search 19 (and which are therefore also indexed in MEDLINE and/or Embase).

c New Scientist, The Conversation, NEJM Journal Watch, Veterinary Record, Chemical & Engineering News and Physics Today.

d The binary ML classifier was built and tested in ER-Web using *MAG* records of COVID-19 research articles that we had matched to records included in, and excluded from, our Map up to Search 19 (and which are therefore also indexed in MEDLINE and/or Embase). It was calibrated to achieve at least 0.95 recall among MEDLINE-Embase records included in this Map, with a corollary workload reduction of ~30% (compared with screening all MEDLINE-Embase records).

**Suggested citation for this Map**:

Lorenc T, Khouja C, Raine G, Shemilt I, Sutcliffe K, D'Souza P, Burchett H, Hinds K, Macdowall W, Melton H, Richardson M, South E, Stansfield C, Thomas S, Kwan I, Wright K, Sowden A, Thomas J (2020) COVID-19: living map of the evidence. London: EPPI-Centre, Social Science Research Unit, UCL Social Research Institute, University College London.

**Funding:**

This Map was commissioned by the National Institute for Health Research (NIHR) Policy Research Programme (PRP) for the Department of Health and Social Care (DHSC) and Public Health England (PHE). It was funded through the NIHR PRP contract with the EPPI Centre at UCL (Reviews facility to support national policy development and implementation, PR-R6-0113-11003). Any views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the DHSC.

**Conflicts of interest:**

None.

**Contributions:**

Any opinions expressed in this publication are not necessarily those of the EPPI-Centre or the funders. Responsibility for any views expressed remains solely with the authors.

**References:**

1. Sinha A, Shen Z, Song S, Ma H, Eide D, Hsu B-J, Wang K. An Overview of Microsoft Academic Service (MA) and Applications. In Proceedings of the 24th International Conference on World Wide Web (WWW '15 Companion): 243-246. ACM, New York, NY, USA. 2015. https://explore.openalex.org/works/W1932742904
2. Thomas J, Graziosi S, Brunton J, Ghouze Z, O'Driscoll P, Bond M (2020). EPPI-Reviewer: advanced software for systematic reviews, maps and other evidence synthesis [Software]. https://eppi.ioe.ac.uk/CMS/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4
3. Shemilt I, Thomas J. MAG-Net-ise it! How the use of Microsoft Academic Graph with machine learning classifiers can revolutionise study identification for systematic reviews. Oral paper accepted for presentation at the 26th Cochrane Colloquium, Santiago, Chile, 22-25 October 2019. https://colloquium2019.cochrane.org/abstracts/mag-net-ise-it-how-use-microsoft-academic-graph-machine-learning-classifiers-can
4. Miwa M, Thomas J, O'Mara-Eves A, Ananiadou S. Reducing systematic review workload through certainty-based screening. Journal of Biomedical Informatics 2014; 51: 242-253. https://explore.openalex.org/works/W2099883114
5. O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. Systematic Reviews 2015; 4:5. https://explore.openalex.org/works/W2147469877
6. Chawla DS. Massive open index of scholarly papers launches [News]. Nature 2022, 24th January. https://explore.openalex.org/works/W4207002271

**Copyright:**

Authors of the systematic reviews on the EPPI-Centre website http://eppi.ioe.ac.uk hold the copyright for the text of their reviews. The EPPI-Centre owns the copyright for all material on the website it has developed, including the contents of the databases, manuals, and keywording and data-extraction systems. The centre and authors give permission for users of the site to display and print the contents of the site for their own non-commercial use, providing that the materials are not modified, copyright and other proprietary notices contained in the materials are retained, and the source of the material is cited clearly following the citation details provided. Otherwise users are not permitted to duplicate, reproduce, re-publish, distribute, or store material from this website without express written permission.

**The NIHR Policy Research Programme** **Reviews Facility is a collaboration between the following:**

[UCL Social Science Research Institute](https://www.ucl.ac.uk/ioe/departments-and-centres/departments/ucl-social-research-institute)

[EPPI-Centre (Evidence for Policy and Practice Information and Co-ordinating Centre)](https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Feppi.ioe.ac.uk%2Fcms%2FDefault.aspx%3Ftabid%3D53&data=02%7C01%7C%7C958803178cf94d41dc4808d79f368f56%7C1faf88fea9984c5b93c9210a11d9a5c2%7C0%7C0%7C637152928987900541&sdata=uxLWGLE%2FGJZpduMddYXAbvn5YtvleFBiM5lK7MA04I8%3D&reserved=0)

[London School of Hygiene & Tropical Medicine](https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.lshtm.ac.uk%2F&data=02%7C01%7C%7C958803178cf94d41dc4808d79f368f56%7C1faf88fea9984c5b93c9210a11d9a5c2%7C0%7C0%7C637152928987910537&sdata=41tucT56JomQwzaNkzn6n%2BjxkshxJP7z6o4BWhFa8UU%3D&reserved=0)

[University of York, Centre for Reviews and Dissemination](https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.york.ac.uk%2Fcrd%2F&data=02%7C01%7C%7C958803178cf94d41dc4808d79f368f56%7C1faf88fea9984c5b93c9210a11d9a5c2%7C0%7C0%7C637152928987910537&sdata=3owdqm4DZsWnz%2Bq%2Ftzc0UpdwkThpXY1YHwxLk58Egn0%3D&reserved=0)