Reviewing the animal literature: how to describe and choose between different types of literature reviews

Cathalijn Leenaars[1,2] ⓘ, **Katya Tsaioun**[3], **Frans Stafleu**[2], **Kieron Rooney**[4], **Franck Meijboom**[2], **Merel Ritskes-Hoitinga**[5,6] **and André Bleich**[1] ⓘ

Abstract

Before starting any (animal) research project, review of the existing literature is good practice. From both the scientific and the ethical perspective, high-quality literature reviews are essential. Literature reviews have many potential advantages besides synthesising the evidence for a research question. First, they can show if a proposed study has already been performed, preventing redundant research. Second, when planning new experiments, reviews can inform the experimental design, thereby increasing the reliability, relevance and efficiency of the study. Third, reviews may even answer research questions using already available data. Multiple definitions of the term literature review co-exist. In this paper, we describe the different steps in the review process, and the risks and benefits of using various methodologies in each step. We then suggest common terminology for different review types: narrative reviews, mapping reviews, scoping reviews, rapid reviews, systematic reviews and umbrella reviews. We recommend which review to select, depending on the research question and available resources. We believe that improved understanding of review methods and terminology will prevent ambiguity and increase appropriate interpretation of the conclusions of reviews.

Keywords

Narrative review, mapping review, scoping review, systematic review, rapid review, umbrella review

Introduction

Literature reviews should provide a comprehensive overview of the publicly available data on a topic, summarising the available evidence and identifying gaps in the currently available knowledge. Besides, reviews can be used to inform planned primary studies, where they can limit several types of systematic error and bias (e.g. those caused by experimenters' expectations). Transparent and comprehensive literature reviews could also benefit the reliability, reproducibility and translational value of scientific studies by assembling and weighing all available data. Thereby they can play a role in solving the reproducibility crisis.

Before starting a project, the relevant literature should be reviewed, to take full advantage of the current state of the art, and to prevent wasting resources on conducting redundant research. This is also encouraged by the organisations funding scientific research;

[1]Institute for Laboratory Animal Science, Hannover Medical School, Germany
[2]Department of Animals in Science and Society, Utrecht University, the Netherlands
[3]Evidence-based Toxicology Collaboration, Johns Hopkins Bloomberg School of Public Health (EBTC), USA
[4]Charles Perkins Centre, Faculty of Medicine and Health, University of Sydney, Australia
[5]SYRCLE, Department for Health Evidence (section HTA), Radboud Institute for Health Sciences, The Netherlands
[6]AUGUST, Department for Clinical Medicine, Aarhus University, Denmark

Corresponding author:
Cathalijn Leenaars, Department of Animals in Science and Society, Utrecht University, Yalelaan 2, 3584 CM, Utrecht, The Netherlands.
Email: Leenaars.Cathalijn@mh-hannover.de

according to the Ensuring Value in Research (EViR) Funders' Collaboration and Development Forum: 'Research should only be funded if set in the context of one or more existing systematic reviews of what is already known or an otherwise robust demonstration of a research gap'.[1] Several funding initiatives encourage systematic reviews of animal experiments specifically.[2,3]

However, the word 'review' has multiple meanings. Relevant dictionary references define a review as 'a critical evaluation'[4] or 'a formal assessment of something with the intention of instituting change if necessary'.[5] Several interpretations for the term 'literature review' co-exist.[6,7] Therefore, students assigned to perform an unspecified literature review are frequently uncertain about which type of review to perform and how to perform it. Besides, readers can be confused about which type of review they are reading.

In this paper, we suggest common terminology for different review types, based on the literature and our expert team's experience with several review types. The review types are summarised in Table 1.

We recommend scientists which review type to select, based on the research question and the available resources. We believe that more consistent use of terminology in the process of selecting, preparing, conducting and reporting of the published reviews will lead to better understanding of the review methodology and improved interpretation of the conclusions.

The review process consists of several steps, and different review types vary in the methodology used in each of these steps. Each procedural choice affects the potential bias in the review results. To understand the differences between different review types and the associated risks of drawing incorrect conclusions, we first outline the review process.

The review process

We provide minimal information on methodology to help selecting the right review type for a specific review question, without writing a detailed manual. Detailed resources already exist[6,8–10] and are referred to throughout. As with any new scientific method, it is challenging to familiarise oneself properly with review methods without proper training and guidance. Specific training and coaching from experienced peers are strongly recommended.

The review process can be very time consuming; Cochrane reviews of clinical trials take 67.3 weeks on average from registered start to publication, with a range of 6 to 186 weeks.[11] As time may be limited, we provide estimates on the time needed for the different steps in the review process, which depend on the amount of literature included. At the start of a review, testing the literature searches and analysing subsets of the search results as a pilot study can help in creating a realistic plan.

Defining the research question

The research question for a literature review can be wide; for example, 'What has been published on compound X', or 'Which animal models exist for disorder Y', or 'What methods are available to measure Z'. Alternatively, it can be very specific and follow what is now widely known as the PICO format.[12,13] PICO stands for population, intervention, comparison and outcome.

The population can be wide; for example, 'all animal species' or very specific; for example, 'ΔF508 mutant mice with a BL6 background'. For questions concerning animal studies with relevance to human outcomes, restricting to a single preclinical species cannot be recommended. While comparing apples and oranges is not advised when we are only interested in apples, we should consider both when we are interested in fruit.

The intervention, or exposure in observational studies (PECO),[14] can be a drug of interest, a specific experimental technique, or an aspect of animal husbandry. As animal studies are more varied in their design than randomised clinical trials, analysing the literature on an

Table 1. Suggested terminology for review types (further clarified below).

Review type	Definition
Narrative review	Non-systematic review contributing an idea or opinion to scientific discourse
Mapping review	Review aiming to provide a high level overview of the complete literature, partially using systematic methodology (i.e. systematised)
Scoping review	Review aiming to provide more or less detailed evidence based on an incomplete convenience sample, partially using systematic methodology (i.e. systematised)
Rapid review	Scoping review following the PICO format
Systematic review	Review comprising a full search resulting in a complete literature overview, inclusion of papers following strict criteria, tabulation of extracted data, risk of bias assessment of included studies, and meaningful (qualitative or quantitative) synthesis of the data
Umbrella review	Review of reviews

experimental technique can be extremely valuable (see, for example, van der Mierden et al.).[15]

The comparator/control can be within or between subjects, and may include sham, untreated, placebo, etc., or any control. Depending on the research question, uncontrolled studies can also be included. As adequate controls are essential to achieve reliable results, including any type of control (instead of only the most appropriate control) can decrease the reliability of the review. However, including uncontrolled studies will increase the number of included studies, and thereby the power of the review results.

The outcome can be a specific behavioural, physiological, histological or other measure. Alternatively, all outcomes can be considered. Mainly for selecting outcome measures, there is a risk of the review authors' bias affecting the inclusion criteria to align the included set of studies with their expectations. Therefore, it is important that these criteria are defined and registered before starting the screening for the search results.

Several alternatives to PICO have been described: PICOT[16] (adding time), PICOS[17] (statistical analysis), PICOT-D[18] (data), SPICE (setting, perspective, intervention, comparison, evaluation)[19] and SPIDER (sample, phenomenon of interest, design, evaluation and research type).[20] PICO searches retrieve more relevant papers than SPIDER searches.[20]

The research question should normally be defined before the authors start searching and selecting papers. While a review can be set up for any review question, the literature may not be available to answer it. Preliminary searches can be used to determine the viability of a review. If a full systematic search does not result in enough information to answer a review question, we suggest publishing the search strategy and the lack of results as a short communication.

Writing a protocol

We highly recommend the template from SYRCLE[21] for writing the review protocol. Systematic review protocols should be posted online for transparency. It is not yet common to post protocols for other review types, but posting prevents cherry picking of the nicest results; 'hypothesising after results are known' (HARKing); and multiple groups performing the same review. Several places are available for protocol posting; for example, PROSPERO (<seurld>www.crd.york.ac.uk/PROSPERO/</seurld>) for those with human health-related outcomes and SYRF (<seurld>http://syrf.org.uk/</seurld>). Both are easily searched for ongoing reviews on a specific topic, and are used well by reviewers.

Besides web-posting, protocols can be published. Publication of a protocol in a scientific journal has the added advantage of external peer review before starting the review efforts, which is expected to increase the quality of the work. Examples are Leenaars et al.,[22] Matta et al.,[23] Pires et al.[24] and van Luijk et al.[25]

Developing a search and searching the literature

The aim of a search strategy is to be complete; to retrieve all relevant papers, which allows us to draw reliable conclusions from our review. For systematic and mapping reviews, multiple databases (e.g. Embase, Psychinfo, Web of Science) should be searched.[26,27] Besides, we search all relevant searchable fields within the databases: title, abstract, (author-defined) keywords and, when available, indexed terms from the database's thesaurus. Different authors use different spellings and synonyms, and complete search strategies should retrieve all of them, by using all relevant terms combined with the Boolean operator 'OR'.

Literature searches generally combine two or three of the PICO elements, combined with the Boolean operator 'AND' (Figure 1). For mapping reviews, but also for relatively rare compounds and models, focusing the search on a single element and screening all the results can be more efficient. Certain elements may not be mentioned in the abstract and including them as a separate element in the search would result in missing relevant papers. For example, if you would include uncontrolled studies, you would not add a search element for the comparison. Another example, if you would want to analyse activity measurements as the outcome, you would miss relevant studies by searching
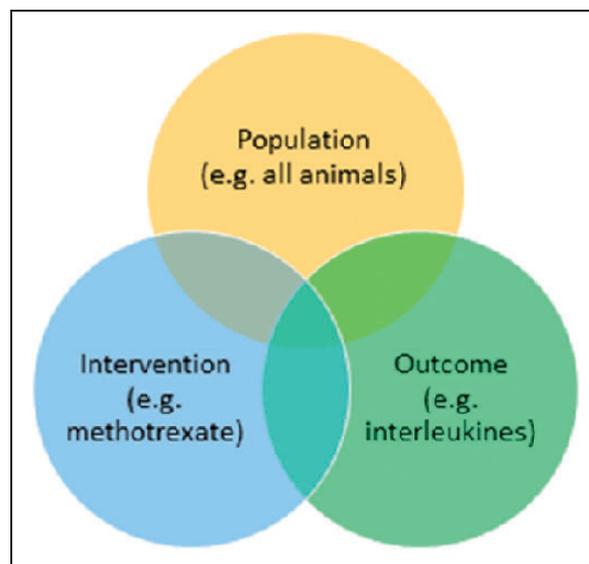

Figure 1. Example search structure; we are interested in the triple-overlapping part of the search strings, in the middle part of the figure.

for them, as activity measurements are often neither reported in the title, abstract and keywords, nor indexed (e.g. Leenaars and colleagues).[28–30]

To develop a search, we normally test each individual term to see how many results it will give us, and whether they are relevant. Development of a search string for an individual element can therefore easily take a month. Reuse of search elements from other reviews and purpose-developed filters (e.g. animal filters)[31,32] will help to save time. Of note, published search strategies need to be evaluated critically before re-use; search strategies in papers that claim to be systematic reviews are often far from comprehensive and will miss a substantial part of the relevant literature.

Besides database searches, alternative strategies can ensure the retrieval of all relevant studies. Mainly small and negative studies may not have been published in journals that are indexed in literature databases. Screening the reference lists of included studies is the most common alternative strategy, but also useful is contacting authors of relevant papers,[33] citation searches and screening conference proceedings, tables of contents of specific journals and dedicated online platforms.

Screening and selecting papers

The first screening of titles and abstracts should be inclusive; when in doubt, a paper should be included for the following full-text screening as the relevant information may not be presented in the abstract. Selecting the relevant papers is one of the more subjective phases of the review; reviewers' preferences could affect which studies are included, and thereby the review's conclusion. Several tools are available to aid the screening process.[34]

Previous work has estimated the average time to screen one title and abstract at around one minute.[35,36] With well-defined criteria, in our experience a trained screener can screen up to 200 titles and abstracts per hour, but not for more than 2 hours per day. Full-text screening takes more time; but with well-defined criteria up to 50–100 papers per hour, for 2 hours per day, is still manageable. Subsequent time needs to be planned for retrieving the full-text papers, waiting for ordered papers to arrive and, if two independent screeners are involved, for discrepancy resolution.

Data extraction

Extraction of the study characteristics and outcome data can be relatively straightforward if only a few parameters are extracted. Alternatively, data extraction can initiate discussions on, for example, which data to

include; how to manage repeated measures (refer to Borenstein et al.[37] for advice); and unanticipated outcome measures that are relevant for the review question. Ideally, the protocol pre-specifies everything to prevent reviewers' theories affecting the selection of data, but this is only possible if the reviewers are very familiar with the literature upfront.

If only a few parameters are extracted, data can be extracted from over 10 papers per hour. For full data extraction, it can easily take over an hour per paper. With multiple data extractors, the extracted data will need harmonisation.

Study quality assessment and risk of bias analysis

The reliability of the conclusion of our review depends on the reliability of the included data. Ideally, we therefore assess reporting quality and/or risk of bias for all included studies.

We distinguish reporting quality from study quality and risk of bias. Reporting quality evaluates if study details have been reported, that is, with the ARRIVE guidelines.[38] Risk of bias is the risk of the study design obscuring the true effect of the intervention under investigation.[39,40] Without proper reporting, we cannot evaluate the risk of bias. For example, when we do not know if an experiment was properly blinded (reporting quality), we cannot estimate to what extent experimenters' expectations have affected the outcome measurements (risk of bias). Many systematic reviews conclude that the risk of bias is unclear for the majority of included studies.

Risk of bias and reporting quality assessment may take around an hour for each individual reference, although most reviewers will manage to work faster with increasing experience. Time can be gained by combining these assessments with the data extraction.

Data analysis

Nearly all reviews will tabulate relevant study characteristics and the main outcomes. Quantitative analysis (formal meta-analysis or another mathematical summary) is optional, and not always informative. We generally advise against performing meta-analyses without performing a risk of bias assessment, as the correct interpretation of the meta-analytical results heavily depend on it.

Sufficient time should be planned for tabulation and qualitative data analysis. Running simple meta-analyses is not necessarily time consuming, they may be performed within 2 hours, but cleaning and import of the data and writing the script for non-standard

analyses can increase the time needed up to several weeks for large datasets or complex analyses.

Reporting

In performing reviews, as in any type of research, we are at risk of biasing our results. Consciously assessing the risks associated with our selected methods and transparently describing them will prevent misinterpretation of the review results. Bias can arise, for example in paper selection; differences in study outcomes between selectively included and excluded publications can occur, for example for English-only papers (language bias), locally available papers only (availability bias), free papers only (cost bias), papers that the review authors are personally familiar with (familiarity bias) and papers with outcomes in line with the expectations (outcome bias).[41]

The review's methods section should clearly outline all methodological steps described above, as decisions made in each step affect the risk of altering the review's conclusions. Guidelines and checklists for reporting of reviews are available, and following them will aid correct interpretation of the review's results. The PRISMA (preferred reporting items for systematic reviews and meta-analyses) checklist comprises 27 items that should be part of every published systematic or systematised review.[42,43] Twelve of these items detail how to describe the review methods.

Review types

In this section the different review types are described further.

Narrative review

The narrative review is the traditional literature review. Historically it was the main and only review type. In the early days of science, it was manageable for one or a few scientists engaged in a narrative review to read all the available literature from their field and stay up to date with new publications. Experts were thus familiar with all relevant literature and could write a thorough review with a solid conclusion, based solely on their expertise.

At the current pace of publication, staying up to date is only possible for a very narrow topic, and even then, it is easy to miss relevant research. Thus, narrative reviews are at an increasing risk of being biased to a single opinion of particular researcher(s), and their conclusions need to be treated accordingly.

However, well written and researched expert narrative reviews occupy an important place in scientific research. These substantiated expert opinion papers can contribute to scientific discourse, or define a vision for the future. The creativity and novelty in narrative reviews can shed new light on scientific fields. While we cannot use them to inform trial design or experimental model choice, they can help to generate testable hypotheses, for example, describing speculative mechanistic insights. Posing fundamental questions and sharing our theories is essential to advance science. Experts performing a narrative review can incorporate elements of the systematic approach (e.g. using a broad literature search strategy and reporting it) to make the narrative review more valuable and sustainable for future generations of scientists.

The narrative reviews of most value are written by experts in the field. They base their opinions on decades of experience, including debates with other experts. Students new to a topic are particularly in danger of selection bias by only being aware of a subset of the relevant research in the field. We therefore do not recommend new students to perform narrative reviews.

We consider the current paper an example of a narrative review which is based on the literature that we are familiar with. It can contribute to developing the animal review field and advance the scientific discussion on review terminology.

Mapping review

A mapping review, also known as a systematic map, is a high-level review with a broad research question, not restricted to the PICO format. It can focus on a specific population, intervention, comparison or outcome, but will not combine all four of these elements. The mapping review comprises a full comprehensive systematic search of a broad field, and presents the global results of the relevant studies in a user-friendly format.[44–48] It identifies the available literature on a topic with basic characteristics, as well as clear evidence gaps. While several authors have described the mapping review as a type of scoping review,[49–51] we find it important to distinguish these review types; the mapping review can be conclusive in describing the available evidence and identifying gaps, while the scoping review is explorative in nature.

To limit bias, ideally two reviewers independently screen all references for inclusion. Because of the large size of a mapping review, data extraction is limited to key study characteristics and outcomes. Reporting quality and risk of bias assessments are usually out of scope. Study characteristics reflecting study and reporting quality can be extracted, but if the mapping review is exploratory, bias cannot be evaluated against a clearly defined hypothesis.

A mapping review answered the review question of how many publications had described intracerebral microdialysis measurements for several amino acids

in animals.[52] It provides lists of the publications on these amino acids, allowing for easy retrieval. Another mapping review was performed to answer the review question 'What are the currently available animal models for cystic fibrosis'.[53] That review provides a full overview of basic animal model characteristics with a list of the outcomes addressed in the primary publications.

Scoping review

Scoping reviews are preliminary explorative assessments of (the potential size and scope of) the available literature on a topic.[6,7,51] A scoping review can be performed for many different types of research questions. The research question may follow the PICO format, but it can also be differently phrased. If the question does not follow the full PICO format, it usually still focusses on specific populations, interventions, comparisons or outcomes. Scoping reviews are generally smaller than mapping reviews, and include more thorough analyses of the included literature.

The scoping review can take several shortcuts compared to the full systematic review. For example, it can search only one database, use a simple or alternative instead of a full search strategy, restrict the included papers to those in English or to certain publication dates, use only one reviewer for the selection of papers, use a convenience sample, etc. Note that the validity of the outcome of a scoping review is restricted to the sample; the results cannot be extrapolated outside this sample. The scoping review comprises a relatively complete description of a well-defined subset of the literature, providing a good idea of the findings reported on a certain topic. As it is explorative, it generally skips the reporting quality/risk of bias assessment.[51]

A scoping review following the PICO format describes the effects of several clinically used drugs for inflammatory bowel disease (intervention) on colon length, histology, myeloperoxidase, daily activity index, macroscopic damage and weight (outcome) in experimental colitis animal models (population) compared to a negative control (comparison).[54] The main shortcut taken in this review is searching only one database. Another one synthesises the animal and human (population) evidence on low-energy sweeteners (intervention) and body weight (outcome).[55] The main shortcut taken in this second scoping review is the lack of a risk of bias assessment for the included animal studies, besides the paper selection probably not being performed by two independent reviewers.

A scoping review not following the PICO format is a retrospective harm–benefit analysis of preclinical animal testing;[56] which analyses the severity of the experiments described in the included papers. This scoping review used convenience sampling; it was based on 228 animal studies included in a preceding systematic review from other authors.[57]

Rapid review

The rapid review is a specific type of scoping review, aimed at answering a specific question following the PICO format. The rapid review is under development by (among others) the Cochrane collaboration.[58] Like scoping reviews, rapid reviews can take shortcuts compared to the full systematic review,[59] for example searching only one database, using a simple instead of a full search strategy, restricting the included papers to those in English or to certain publication dates, using only one reviewer for the selection of papers, skipping the reporting quality/risk of bias assessment, extracting only limited data, etc.

We do not expect rapid reviews to become common in animal research, as the need for answering a research question will rarely be urgent enough to warrant shortcuts. However, they could be useful; for example, when an unexpected adverse event is first observed after human exposure. A rapid review of preclinical data could then help in risk assessment. A recent rapid review[60] analyses the effect of organophosphate pesticide exposure (intervention) on breast cancer risk (outcome) in humans, animals and cells (population). It was inspired by the publication of a primary observational study.

Full systematic review

A systematic review is designed to locate, appraise and synthesise the evidence to answer a specific research question in an evidence-based manner.[9] According to the Cochrane handbook, the key characteristics of a systematic review are a clearly stated (set of) objective (s), predefined eligibility criteria for inclusion, explicit and reproducible methodology, a comprehensive systematic search, assessment of the validity of the findings, a systematic presentation of the characteristics and findings of the included studies, and a systematic synthesis of the evidence, ideally using meta-analysis. Systematic reviews are the review type least prone to bias and will therefore answer research questions most reliably. They are therefore increasingly encouraged in the animal sciences. Much information on performing systematic reviews is available from several sources (<seurld>www.SYRCLE.nl</seurld>, tools section, <seurld>www.SYRF.org.uk</seurld>, for toxicological subjects <seurld>www.EBTox.org</seurld> and Hoffmann et al.[61]).

While it is theoretically possible to perform a full systematic review for any research question, they generally follow the PICO format. A systematic review fully following the methods described above addresses the review question 'what are the metabolic and behavioural effects (outcome) in offspring (population) exposed prenatally to non-nutritive sweeteners (intervention)'.[62] However, a systematic review can also be used to answer a different type of review question, for example 'are the experimental designs of preclinical animal studies comparable to those of the clinical trials?'.[63]

Other review types

Scientists can review different subsets of the literature, for example pooling the data from several experiments from an individual research group or a research consortium. Pooling these datasets in a meta-analysis can increase the power compared to individual experiments and is therefore informative. However, the validity of the outcome is restricted to the sample; external validity is limited. Regardless of this, these types of analyses can provide us with valuable information, for example, on the use of body weight reduction as a humane endpoint.[64] For clarity, the term 'review' should be replaced by 'data synthesis', and the research group or consortium mentioned in the title.

A second example is the umbrella review, a review of reviews on a specific topic. While evidence can efficiently be collated to provide a comprehensive overview, results of umbrella reviews are limited by the quality of the primary reviews they include. An umbrella review of 103 systematic reviews and meta-analyses of animal studies concluded that reporting of meta-analyses was inadequate.[65] Other examples of umbrella reviews focus on, for example, the risk of bias analyses,[66] dissemination bias[67] and the translational value of animal studies.[33]

A published analysis of review types identified several other review types.[7] In our opinion, other review types can all be classified as one of the review types described here. Unspecified literature reviews, expert reviews, critical reviews and state-of-the-art reviews are usually narrative reviews. Mixed methods reviews and qualitative reviews can be mapping, scoping or full systematic reviews comprising a qualitative text analysis. Systematised reviews are usually either mapping or scoping reviews.

A summary of methods for different review types

The methods used in different review types are summarized in Table 2.

Table 2. A summary of review types.

Review type	Research question	Number of references	Complete literature	Protocol	Search	Screening	Data extraction	RoB	Analysis	Resources
Narrative	Any	Any	No	No	Unstructured	None	Any	Recommended	Any	Low–medium
Mapping	Wide	Large	Yes	Yes	Full	Single/duplo	Low level of detail	Recommended	Global	Medium
Scoping	Rather specific	Medium	No	Yes	Partial	Single/duplo	Medium–highly detailed	Recommended	Any	Medium
Systematic	Specific	Small	Yes	Yes	Full	Duplo	Highly detailed	Required	Detailed	High
Rapid	Specific	Medium	No	Yes	Partial	Single/duplo	Medium–highly detailed	Required	Detailed	Medium

RoB: risk of bias.

What review type to select?

The type of review to select depends on several factors. If reviewers have a specific question and sufficient resources, we recommend performing a full systematic review, as this review type is least prone to bias. If the results of the review are urgently needed to assess safety risks, the rapid review is preferred. If the research question is not specific enough to follow the PICO format, the review type depends on the research field under consideration. Starting with a wide mapping review can be extremely useful, as all relevant literature (on a specific intervention, technique or set of animal models) will be gathered. The mapping review can also be used as a basis for further reviews, both by the original authors and by other teams. The scoping review can be used to explore more or less specific research questions.

A simplified flow scheme to help choose the appropriate review type is provided in Figure 2. While it may be challenging to find sufficient resources for a full systematic review, we hope that being aware of the speed versus quality trade-off will limit the effect of 'perverse incentives' on selecting a review type.

Conclusion and discussion

In this paper, we outline important steps in reviewing the literature, and describe different review types and their uses. Systematic and systematised reviews can be used to put a planned experiment into perspective, to refine the experimental design, to show that the experiment has not yet been done, to identify alternative research methodologies, and occasionally even as an alternative to de novo experiments (answering a new research question based on available data). The more we use systematic review methodology, the lower the risk of bias in the conclusions, and the lower the chance that we disregard important evidence contradicting our preconceived ideas. However, as the time and resources needed to complete a review are considerable, it is important to consider viability upfront.

A literature review can be part of the ethical review process. The literature review can put the planned study into perspective, but it may also be used to refine the experimental design.[68,69] Unfortunately, systematic reviews show considerable repetition of experiments for some topics[70] indicating that reviews are not yet optimally used. In certain cases, systematic reviews with meta-analyses can be used as an efficient alternative to a new experiment by answering a novel research question based on existing data (e.g. van der Mierden).[15] The value of systematic reviews of animal studies is increasingly recognised by research funders; for example, The Netherlands Organization for Health Research and Development[3] and the German Federal Ministry of Education and Research[2] have created specific calls to fund them.

Reviews can help to advance science in several ways. First, reviews of primary studies can prevent the use of funds on conducting redundant experiments, as cumulative meta-analyses can show when a treatment's overall estimated effect size no longer changes with further studies.[71] Second, they can stimulate the adoption of evidence-based experimental design for future experiments,[68,72] thus improving the quality, consistency and reproducibility of primary research. Third, systematic
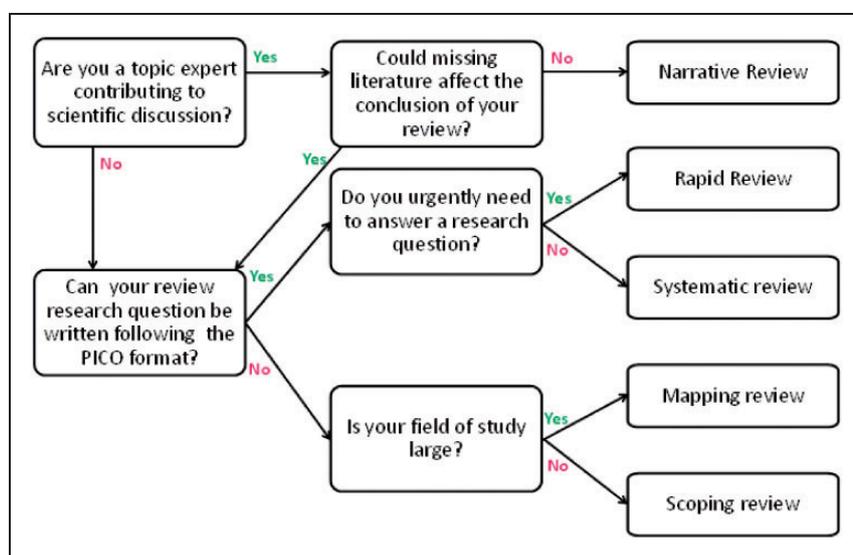

Figure 2. Simplified flow scheme to determine the optimal review type. Note: this flow scheme disregards the amount of time and money available.

reviews with meta-analyses can be used to answer research questions without using new animals.[15]

Timely reviews of animal studies may improve success rates and patients' safety in clinical trials. Several clinical trials with undesirable outcomes would possibly not have started if the animal literature had been analysed more thoroughly upfront. Examples comprise PROPATRIA (probiotics);[73] trials for calcium antagonists after stroke;[74,75] trials for the MVA85A vaccine;[76] and STRIDER (sildenafil).[77] However, these reviews were performed after the clinical trials; we are not sure if more timely reviews would have influenced the decision to start.

While the quality of the included studies remains a limiting factor in all types of evidence synthesis, thorough review of the data can prevent wasting resources on suboptimal and unnecessary studies. The term 'literature review' still may comprise any of the review types described in this paper. We hope that the consequent use of the names of the described review types in the titles of published reviews will improve the understanding and interpretation of these literature reviews. Besides, we anticipate that starting scientists will benefit from our tables and figure in deciding which review type to select, depending on their experience, research question and available resources. Moreover, we expect that this text will inspire all literature reviewers to describe clearly the review methodology used in their publications.

ORCID iDs

Cathalijn Leenaars (iD) https://orcid.org/0000-0002-8212-7632
André Bleich (iD) https://orcid.org/0000-0002-3438-0254

References

1. EViR. *Guiding Principles*. https://sites.google.com/view/evir-funders-forum/guiding-principles (accessed 24 April 2019).
2. BMBF and research) BfBuFMfea. *Richtlinie zur Förderung von konfirmatorischen präklinischen Studien – Qualität in der Gesundheitsforschung*. 2018. https://www.gesundheitsforschung-bmbf.de/de/8344.php (accessed 24 January 2020).
3. ZonMW (the Netherlands national funding organisation for health research). Programme More Knowledge with Fewer Animals. 2020. 2020. https://www.zonmw.nl/en/programma-opslag-en/more-knowledge-with-fewer-animals/ (2020; accessed 24 January 2020).
4. Merriam-Webster. *Merriam-Webster Online Dictionary*, 2019. https://www.merriam-webster.com/dictionary/review (2019; accessed 26 July 2019).
5. Oxford. *Lexico online dictionary*, 2019. https://www.lexico.com/definition/review (accessed 26 July 2019).
6. Booth A, Papaioannou D and Sutton A. *Systematic approaches to a successful literature review*. London: Sage, 2012.
7. Grant MJ and Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009; 26: 91–108. 2009/06/06. DOI: 10.1111/j.1471-1842.2009.00848.x.
8. Higgins JPT and Green S. *Cochrane handbook for systematic reviews of interventions*. London: Wiley-Blackwell & the Cochrane Collaboration, 2008, pp. 649.
9. Boland A, Cherry MG and Dickson R. *Doing a systematic review – a student's guide*. London: Sage, 2014, pp. 210.
10. Jesson JK, Matheson L and Lacey FM. *Doing your literature review – traditional and systematic techniques*. London: Sage, 2011.
11. Borah R, Brown AW, Capers PL, et al. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* 2017; 7: e012545. 2017/03/01. DOI: 10.1136/bmjopen-2016-012545.
12. Eriksen MB and Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *J Med Libr Assoc* 2018; 106: 420–431. 2018/10/03. DOI: 10.5195/jmla.2018.345.
13. Richardson WS, Wilson MC, Nishikawa J, et al. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995; 123: A12–A13. 1995/11/01.
14. Morgan RL, Whaley P, Thayer KA, et al. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. *Environ Int* 2018; 121: 1027–1031. 2018/09/01. DOI: 10.1016/j.envint.2018.07.015.
15. van der Mierden S, Savelyev SA, IntHout J, et al. Intracerebral microdialysis of adenosine and adenosine monophosphate – a systematic review and meta-regression analysis of baseline concentrations. *J Neurochem* 2018; 147: 58–70. 2018/07/20. DOI: 10.1111/jnc.14552.

16. Echevarria IM and Walker S. To make your case, start with a PICOT question. *Nursing* 2014; 44: 18–19. 2014/01/17. DOI: 10.1097/01.NURSE.0000442594.00242.f9.

17. Saaiq M and Ashraf B. Modifying "Pico" question into "Picos" model for more robust and reproducible presentation of the methodology employed in a scientific study. *World J Plast Surg* 2017; 6: 390–392. 2017/12/09.

18. Elias BL, Polancich S, Jones C, et al. Evolving the PICOT method for the digital age: the PICOT-D. *J Nurs Educ* 2015; 54: 594–599. 2015/10/03. DOI: 10.3928/01484834-20150916-09.

19. Phelan S, Lin F, Mitchell M, et al. Implementing early mobilisation in the intensive care unit: an integrative review. *Int J Nurs Stud* 2018; 77: 91–105. 2017/10/27. DOI: 10.1016/j.ijnurstu.2017.09.019.

20. Methley AM, Campbell S, Chew-Graham C, et al. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv Res* 2014; 14: 579. 2014/11/22. DOI: 10.1186/s12913-014-0579-0.

21. De Vries RBM, Hooijmans CR, Langendam MW, et al. A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. *Evid Based Preclin Med* 2015; 1, 1, 1–9, e00007, DOI: 10.1002/ebm2.7

22. Leenaars CHC, van der Mierden S, Durst M, et al. Measurement of corticosterone in mice: a protocol for a mapping review. *Lab Anim* 2020; 54: 26–32.

23. Matta K, Ploteau S, Coumoul X, et al. The associations between exposure to environmental endocrine disrupting chemicals and endometriosis in experimental studies: a systematic review protocol. *Environment Int* 2019; 124: 400–407. DOI: 10.1016/j.envint.2018.12.063.

24. Pires GM, Bezerra AG, de Vries RBM, et al. Effects of experimental sleep deprivation on aggressive, sexual and maternal behaviour in animals: a systematic review protocol. *BMJ Open Sci* 2018; 2: e000041. doi:10.1136/bmjos-2017-000041.

25. van Luijk J, Popa M, Swinkels J, et al. Establishing a health-based recommended occupational exposure limit for nitrous oxide using experimental animal data – a systematic review protocol. *Environ Res* 2019; 178: 108711. 2019/09/15. DOI: 10.1016/j.envres.2019.108711.

26. Leenaars M, Hooijmans CR, van Veggel N, et al. A step-by-step guide to systematically identify all relevant animal studies. *Lab Anim* 2020; 54: 330–340.

27. Roi AJ and Grune B. *The EURL ECVAM search guide – good search practice on animal alternatives.* Luxembourg: Publications Office of the European Union, 2013, pp. 124.

28. Leenaars CH, Girardi CE, Joosten RN, et al. Instrumental learning: an animal model for sleep dependent memory enhancement. *J Neurosci Methods* 2013; 217: 44–53. 2013/04/23. DOI: 10.1016/j.jneumeth.2013.04.003.

29. Leenaars CH, Joosten RN, Kramer M, et al. Spatial reversal learning is robust to total sleep deprivation. *Behav Brain Res* 2012; 230: 40–47. 2012/02/11. DOI: 10.1016/j.bbr.2012.01.047.

30. Leenaars CH, Joosten RN, Zwart A, et al. Switch-task performance in rats is disturbed by 12 h of sleep deprivation but not by 12 h of sleep fragmentation. *Sleep* 2012; 35: 211–221. 2012/02/02. DOI: 10.5665/sleep.1624.

31. de Vries RB, Hooijmans CR, Tillema A, et al. A search filter for increasing the retrieval of animal studies in Embase. *Lab Anim* 2011; 45: 268–270. 2011/09/06. DOI: 10.1258/la.2011.011056.

32. de Vries RB, Hooijmans CR, Tillema A, et al. Updated version of the Embase search filter for animal studies. *Lab Anim* 2014; 48: 88. 2013/07/10. DOI: 10.1177/0023677213494374.

33. Menon JML, Nolten C, Achterberg EJM, et al. Brain microdialysate monoamines in relation to circadian rhythms, sleep, and sleep deprivation – a systematic review, network meta-analysis, and new primary data. *J Circadian Rhythms* 2019; 17: 1. 2019/01/24. DOI: 10.5334/jcr.174.

34. Van der Mierden S, Tsaioun K, Bleich A, et al. Software tools for literature screening in systematic reviews in biomedical research. *ALTEX* 2019; 36: 508–517. 2019/05/22. DOI: 10.14573/altex.1902131.

35. Bramer WM, Milic J and Mast F. Reviewing retrieved references for inclusion in systematic reviews using EndNote. *J Med Libr Assoc* 2017; 105: 84–87. 2017/01/18. DOI: 10.5195/jmla.2017.111.

36. Shemilt I, Khan N, Park S, et al. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. *Syst Rev* 2016; 5: 140. 2016/08/19. DOI: 10.1186/s13643-016-0315-4.

37. Borenstein M, Hedges LV, Higgins JPT, et al. *Introduction to meta-analysis.* Chichester: John Wiley & Sons Ltd., 2009, pp. 421.

38. Kilkenny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010; 8: e1000412. 2010/07/09. DOI: 10.1371/journal.pbio.1000412.

39. Hooijmans CR, Rovers MM, de Vries RB, et al. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol* 2014; 14: 43. 2014/03/29. DOI: 10.1186/1471-2288-14-43.

40. Krauth D, Woodruff TJ and Bero L. Instruments for assessing risk of bias and other methodological criteria of published animal studies: a systematic review. *Environ Health Perspect* 2013; 121: 985–992. 2013/06/19. DOI: 10.1289/ehp.1206389.

41. Rothstein HR, Sutton AJ and Borenstein M. *Publication bias in meta-analyses.* Chichester: Wiley, 2005, pp. 356.

42. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; 339: b2535. 2009/07/23. DOI: 10.1136/bmj.b2535.

43. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473. 2018/09/05. DOI: 10.7326/M18-0850.

44. James KL, Randall NP and Haddaway NR. A methodology for systematic mapping in environmental sciences. *Environmental Evidence* 2016; 5: 1–13.

45. Miake-Lye IM, Hempel S, Shanman R, et al. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Syst Rev* 2016; 5: 28. 2016/02/13. DOI: 10.1186/s13643-016-0204-x.

46. O'Leary BC, Woodcock P, Kaiser MJ, et al. Evidence maps and evidence gaps: evidence review mapping as a method for collating and appraising evidence reviews to inform research and policy. *Environmental Evidence* 2017; 6: 1–9. DOI: 10.1186/s13750-017-0096-9.

47. Snilstveit A, Vojtkova M, Bhavsar A, et al. Evidence & gap maps: a tool for promoting evidence informed policy and strategic research agendas. *J Clin Epidemiol* 2016; 79: 120–129. DOI: 10.1016/j.jclinepi.2016.05.015.

48. Wolffe TAM, Whaley P, Halsall C, et al. Systematic evidence maps as a novel tool to support evidence-based decision-making in chemicals policy and risk management. *Environ Int* 2019; 130: 104871. 2019/06/30. DOI: 10.1016/j.envint.2019.05.065.

49. Arksey H and O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005; 8: 19–32. DOI: 10.1080/1364557032000119616.

50. Levac D, Colquhoun H and O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010; 5: 69. 2010/09/22. DOI: 10.1186/1748-5908-5-69.

51. Rumrill PD, Fitzgerald SM and Merchant WR. Using scoping literature reviews as a means of understanding and interpreting existing literature. *Work* 2010; 35: 399–404. 2010/04/07. DOI: 10.3233/WOR-2010-0998.

52. Leenaars CHC, Freymann J, Jakobs K, et al. A systematic search and mapping review of studies on intracerebral microdialysis of amino acids, and systematized review of studies on circadian rhythms. *J Circadian Rhythms* 2018; 16: 12. 2018/11/30. DOI: 10.5334/jcr.172.

53. Leenaars CH, De Vries RB, Heming A, et al. Animal models for cystic fibrosis: a systematic search and mapping review of the literature – Part 1: genetic models. *Lab Anim* 2019; 23677219868502. 2019/08/15. DOI: 10.1177/0023677219868502.

54. Zeeff SB, Kunne C, Bouma G, et al. Actual usage and quality of experimental colitis models in preclinical efficacy testing: a scoping review. *Inflamm Bowel Dis* 2016; 22: 1296–1305. 2016/04/23. DOI: 10.1097/MIB.0000000000000758.

55. Rogers PJ, Hogenkamp PS, de Graaf C, et al. Does low-energy sweetener consumption affect energy intake and body weight? A systematic review, including meta-analyses, of the evidence from human and animal studies. *Int J Obes (Lond)* 2016; 40: 381–394. 2015/09/15. DOI: 10.1038/ijo.2015.177.

56. Pound P and Nicol CJ. Retrospective harm benefit analysis of pre-clinical animal research for six treatment interventions. *PLoS One* 2018; 13: e0193758. 2018/03/29. DOI: 10.1371/journal.pone.0193758.

57. Perel P, Roberts I, Sena E, et al. Comparison of treatment effects between animal experiments and clinical trials: systematic review. *BMJ* 2007; 334: 197. 2006/12/19. DOI: 10.1136/bmj.39048.407928.BE.

58. Cochrane. *Rapid Reviews Methods Group.* 2019. https://methods.cochrane.org/rapidreviews/about-us (2019; accessed 24 April 2019).

59. Tricco AC, Antony J, Zarin W, et al. A scoping review of rapid review methods. *BMC Med* 2015; 13: 224. 2015/09/18. DOI: 10.1186/s12916-015-0465-6.

60. Yang KJ, Lee J and Park HL. Organophosphate pesticide exposure and breast cancer risk: a rapid review of human, animal, and cell-based studies. *Int J Environ Res Public Health* 2020; 17: 5030; doi:10.3390/ijerph171450301-25.

61. Hoffmann S, de Vries RBM, Stephens ML, et al. A primer on systematic reviews in toxicology. *Arch Toxicol* 2017; 91: 2551–2575. 2017/05/16. DOI: 10.1007/s00204-017-1980-3.

62. Morahan HL, Leenaars CHC, Boakes RA, et al. Metabolic and behavioural effects of prenatal exposure to non-nutritive sweeteners: a systematic review and meta-analysis of rodent models. *Physiol Behav* 2020; 213: 112696. 2019/10/28. DOI: 10.1016/j.physbeh.2019.112696.

63. Leenaars C, Stafleu F, de Jong D, et al. A systematic review comparing experimental design of animal and human methotrexate efficacy studies for rheumatoid arthritis: lessons for the translational value of animal studies. *Animals (Basel)* 2020; 10: 1–20. 2020/06/21. DOI: 10.3390/ani10061047.

64. Talbot SR, Biernot S, Bleich A, et al. Defining body-weight reduction as a humane endpoint: a critical appraisal. *Lab Anim* 2020; 54: 99–110. 2019/11/02. DOI: 10.1177/0023677219883319.

65. Peters JL, Sutton AJ, Jones DR, et al. A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting. *J Environ Sci Health B* 2006; 41: 1245–1258. 2006/08/23. DOI: 10.1080/10934520600655507.

66. van Luijk J, Bakker B, Rovers MM, et al. Systematic reviews of animal studies; missing link in translational research? *PLoS One* 2014; 9: e89981. 2014/03/29. DOI: 10.1371/journal.pone.0089981.

67. Mueller KF, Briel M, Strech D, et al. Dissemination bias in systematic reviews of animal research: a systematic review. *PLoS One* 2014; 9: e116016. 2014/12/30. DOI: 10.1371/journal.pone.0116016.

68. de Vries RB, Wever KE, Avey MT, et al. The usefulness of systematic reviews of animal experiments for the design of preclinical and clinical studies. *ILAR J* 2014; 55: 427–437. 2014/12/30. DOI: 10.1093/ilar/ilu043.

69. Sena ES, Currie GL, McCann SK, et al. Systematic reviews and meta-analysis of preclinical studies: why perform them and how to appraise them critically. *J Cereb Blood Flow Metab* 2014; 34: 737–742. 2014/02/20. DOI: 10.1038/jcbfm.2014.28.

70. Yauw ST, Wever KE, Hoesseini A, et al. Systematic review of experimental studies on intestinal anastomosis. *Br J Surg* 2015; 102: 726–734. 2015/04/08. DOI: 10.1002/bjs.9776.

71. Sena ES, Briscoe CL, Howells DW, et al. Factors affecting the apparent efficacy and safety of tissue plasminogen activator in thrombotic occlusion models of stroke: systematic review and meta-analysis. *J Cereb Blood Flow Metab* 2010; 30: 1905–1913. 2010/07/22. DOI: 10.1038/jcbfm.2010.116.

72. de Vries RB, Buma P, Leenaars M, et al. Reducing the number of laboratory animals used in tissue engineering research by restricting the variety of animal models. Articular cartilage tissue engineering as a case study. *Tissue Eng Part B Rev* 2012; 18: 427–435. 2012/05/11. DOI: 10.1089/ten.TEB.2012.0059.

73. Hooijmans CR, de Vries RB, Rovers MM, et al. The effects of probiotic supplementation on experimental acute pancreatitis: a systematic review and meta-analysis. *PLoS One* 2012; 7: e48811. 2012/11/16. DOI: 10.1371/journal.pone.0048811.

74. Horn J, de Haan RJ, Vermeulen M, et al. Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 2001; 32: 2433–2438. 2001/10/06.

75. Horn J and Limburg M. Calcium antagonists for ischemic stroke: a systematic review. *Stroke* 2001; 32: 570–576. 2001/02/07.

76. Kashangura R, Sena ES, Young T, et al. Effects of MVA85A vaccine on tuberculosis challenge in animals: systematic review. *Int J Epidemiol* 2015; 44: 1970–1981. 2015/09/10. DOI: 10.1093/ije/dyv142.

77. Symonds ME and Budge H. Comprehensive literature search for animal studies may have saved STRIDER trial. *BMJ* 2018; 362: k4007. 2018/09/27. DOI: 10.1136/bmj.k4007.

Résumé

Avant de commencer un projet de recherche (animale), l'examen de la littérature existante est une bonne pratique. Du point de vue scientifique et éthique, des examens de la littérature de haute qualité sont essentiels. Les examens de littérature présentent de nombreux avantages potentiels en plus de synthétiser les données probantes concernant une question de recherche. Ils peuvent tout d'abord indiquer si l'étude proposée a déjà été effectuée, empêchant ainsi la recherche redondante. Ensuite, lors de la planification de nouvelles expériences, les examens peuvent éclairer la conception expérimentale, augmentant ainsi la fiabilité, la pertinence et l'efficacité de l'étude. Enfin, les examens peuvent également répondre aux questions de recherche en utilisant les données déjà disponibles. Plusieurs définitions du terme examen de la littérature coexistent. Dans le présent document, nous décrivons les différentes étapes du processus d'examen, ainsi que les risques et les avantages liés à l'utilisation de diverses méthodologies à chaque étape. Nous suggérons ensuite une terminologie commune pour différents types d'examen: les examens narratifs, les examens de cartographie, les examens de portée, les examens rapides, les examens systématiques et les examens de coordination. Nous recommandons l'examen à sélectionner, en fonction de la question de recherche et des ressources disponibles. Nous estimons qu'une meilleure compréhension des méthodes d'examen et de la terminologie permettra d'éviter l'ambiguïté et d'améliorer l'interprétation appropriée des conclusions des examens.

Abstract

Vor Beginn eines (Tier-)Forschungsprojekts ist die Durchsicht der vorhandenen Literatur (Review) eine gute Praxis. Sowohl in wissenschaftlicher als auch in ethischer Hinsicht sind qualitativ hochwertige Literaturüberblicke unerlässlich. Literaturüberblicke haben neben der Zusammenstellung der Fakten für eine Forschungsfrage viele potenzielle Vorteile. Erstens können sie zeigen, ob eine vorgeschlagene Studie bereits durchgeführt wurde, wodurch überflüssige Forschung vermieden wird. Zweitens können Übersichten bei der Planung neuer Experimente in das Versuchsdesign einfließen und dadurch die Zuverlässigkeit, Relevanz und Effizienz von Studien erhöhen. Drittens können Reviews sogar Forschungsfragen unter Verwendung bereits verfügbarer Daten beantworten. Es existieren mehrere parallele Definitionen des Begriffs Literaturüberblick. In diesem Veröffentlichung beschreiben wir die verschiedenen Schritte des Begutachtungsprozesses sowie die Risiken und Vorteile der Anwendung verschiedener Methoden in jedem Schritt. Anschließend schlagen wir eine gemeinsame Terminologie für verschiedene Arten von Reviews vor: narrative Überblicke, Mapping-Überblicke, Scoping-Reviews, Schnellüberblicke, systematische Überblicke und Umbrella-Reviews. Je nach Forschungsfrage und verfügbaren Ressourcen empfehlen wir, welche Review zu wählen ist. Wir glauben, dass ein besseres Verständnis der Reviewmethoden und -terminologie Ambivalenzen verhindert und die angemessene Interpretation der Schlussfolgerungen von Reviews verbessert.

Resumen

Antes de empezar cualquier proyecto de investigación (con animales), es aconsejable revisar la bibliografía existente. Tanto desde la perspectiva científica como ética, unas revisiones bibliográficas de alta calidad son esenciales. Las revisiones bibliográficas presentan muchas posibles ventajas más allá de sintetizar las pruebas para un tema de investigación. En primer lugar, pueden evidenciar si un estudio propuesto ya ha sido realizado, evitando así cualquier duplicación en la investigación. En segundo lugar, al planificar nuevos experimentos, las revisiones pueden dar información sobre el diseño experimental, incrementando por tanto la fiabilidad, la relevancia y la eficiencia del estudio. En tercer lugar, las revisiones pueden incluso responder preguntas de investigación utilizando datos ya disponibles. Existen múltiples definiciones del término «revisión bibliográfica». En este estudio, describimos los distintos pasos del proceso de revisión, así como los riesgos y beneficios de utilizar varias metodologías en cada paso. Asimismo, sugerimos una terminología en común para distintos tipos de revisiones: revisiones narrativas, revisiones de mapeo, revisiones de alcance, revisiones rápidas, revisiones sistemáticas y revisiones paraguas. Recomendamos qué revisiones seleccionar, según el tema de investigación y los recursos disponibles. Creemos que una mejora del entendimiento de los métodos de revisión y de la terminología puede evitar la ambigüedad y mejorar la interpretación de las conclusiones de las revisiones.