TUTORIAL

Visualization toolkits for enriching meta-analyses through evidence maps, bibliometrics, and alternative impact metrics

Yefeng Yang [1], Malgorzata Lagisz [1,2] and Shinichi Nakagawa [1,2]

[1]Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia
[2]Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

Corresponding authors: Yefeng Yang and Shinichi Nakagawa; Emails: yefeng.yang1@unsw.edu.au; s.nakagawa@unsw.edu.au

Abstract

Data visualization is crucial for effectively communicating knowledge in meta-analysis. However, existing visualization methods in meta-analysis have predominantly focused on quantitative aspects, such as forest plots and funnel plots, thereby neglecting qualitative information that is equally important for end-users in science, policy, and practice. We introduce a framework consisting of a series of visualization toolkits designed to enrich meta-analyses by borrowing approaches from other research synthesis methods, including systematic evidence mapping (scoping reviews), bibliometrics (bibliometric analysis), and alternative impact metric analysis. These "enrichment" toolkits aim to facilitate the synthesis of both quantitative and qualitative evidence, along with the assessment of the academic and nonacademic influences of the meta-analytic evidence base. While the meta-analysis yields quantitative insights, the enrichment analyses, and visualizations provide user-friendly summaries of qualitative information on the evidence base. For example, a systematic evidence map can visualize study characteristics, unraveling knowledge gaps and methodological differences. Bibliometric analysis offers a visual assessment of the nonindependent evidence, such as hyper-dominant authors and countries, and funding sources, potentially informing the risk of bias. Alternative impact metric analysis employs alternative metrics to gauge societal influence and research translation (e.g., policy and patent citations) of studies in the meta-analysis. We provide a dedicated webpage showcasing sample visualizations and providing step-by-step implementation in open-source software R (https://yefeng0920.github.io/MA_Map_Bib/). Additionally, we offer a guide on leveraging three commercially free large language models (LLMs) to help adapt the sample script, enabling users with less R coding experience to visualize their own meta-analytic evidence base.

Highlights
What is already known

- Data visualization enhances meta-analysis, aiding in identifying patterns and increasing the accessibility of evidence to end-users in science, policy, and practice.
- Quantitative information in meta-analyses is commonly presented through user-friendly formats, such as forest plots and funnel plots, with effect size estimates and precisions.

- Meta-analyses include qualitative information on study characteristics and bibliographic attributes, which are rarely visualized.

What is new

- Integrating the elements of scoping reviews and bibliometrics into meta-analyses enriches research synthesis by visualizing qualitative evidence.
- Scoping reviews (systematic mapping) can enhance meta-analyses by providing a user-friendly summary of study characteristics and identifying knowledge gaps and gluts.
- Bibliometric analysis can visualize nonindependent evidence, revealing dominant authors and countries, informing potential risk of bias (RoB).
- Further, alternative impact metrics, such as Altmetric attention scores, patents, and policy citations, can help capture the societal impacts of meta-analyses beyond academia.

Potential impact for RSM readers

- The proposed visualization toolkits transform qualitative information into easily digestible formats for end-users, requiring minimal additional data extraction effort from meta-analysts.
- The proposed visualization toolkits' capacity can be extended to other methods in research synthesis ecosystems, such as the systematic-review family, review-of-review family, and rapid-review family, to convey quantitative information in a user-friendly manner for every evidence base.

1. Introduction

Data visualization is an effective tool for knowledge communication. It also works in meta-analysis,[1–3] a statistical method for quantitatively synthesizing evidence.[4–12] The incorporation of data visualization can yield two potential benefits to meta-analysis. First, data visualization can facilitate the identification of patterns, trends, and associations, not immediately apparent within the complex meta-analytic data.[13] Second, it could increase the uptake and interpretation of meta-analytic findings by the end-users in science, policy, and practice.[2,14,15]

Typical meta-analytic data include both quantitative and qualitative information. Quantitative information usually refers to the effect size estimates, sampling variances, and quantitative moderators. Qualitative information comes in two types: study characteristics (e.g., contextual factors of the performed research) and bibliographic attributes (e.g., author and country information). Convention-ally, the meta-analytic visualization methods have primarily focused on the quantitative aspects,[1,13] exemplified by forest plots and funnel plots for representing effect sizes and their precisions.[3] However, qualitative information is rarely visualized or tends to be conveyed in a format not immediately digestible by end-users.[16] Comprehensive and clear visual representation of moderator variables can facilitate customizable evidence synthesis, identifying context-specific drivers and delivering more tailored evidence for science, policy, and practice, improving the generalizability and transferability of evidence.[17] Yet, these qualitative details are no less valuable than their quantitative counterparts.

A user-friendly summary of study characteristics can add value to meta-analyses. For example, visualizing the intersections between the types of interventions and outcomes (or other combinations of two categorical variables) can identify knowledge gaps,[18] thus setting the agenda for future research to fill such gaps. At the same time, identifying evidence "gluts" is of urgent importance given research waste due to preventable problems,[19] such as designing new experiments while ignoring the existing evidence.[20] The need for such visualizations aligns with the tenets of scoping reviews, sometimes termed evidence review or map (see [21,22] for a debate).[23,24] Interestingly, network meta-analyses offer a precursor to integrating evidence mapping into meta-analyses. The geometry of network graphs can visually unveil the state of the evidence, distinguishing well-studied interventions from those warranting further investigation.[25]

Similarly, visualizing bibliographic attributes remains relatively rare in meta-analyses. Yet, such visualizations can reveal the nonindependence of evidence.[26] For example, bibliometric analyses

can map author collaboration networks,[27] define research groups, and detect hyper-dominant author clusters (see [28] for an example).[29] Such hyper-dominance induces nonindependent evidence and could influence between-study level risk of bias (RoB). Bibliometric analysis of traditional impact metrics (e.g., citations) can be used to map the conceptual influences within academia. Additionally, alternative impact metrics[30–33] have been proposed to complement traditional citation impact metrics. These alternative impact metrics, such as Altmetric attention scores, can reflect the social impact of a paper based on online news media (e.g., Facebook), policy documents, patents, Wikipedia, and Stack Overflow, reflecting the influences outside academia.

In this work, we present a series of visualization toolkits that enrich traditional meta-analyses by seamlessly integrating visualizations used in three research synthesis approaches: evidence mapping, bibliometrics, and alternative impact analyses (Figure 1). We explain the rationale of each "enrichment" approach and corresponding visualization toolkits. We then present a dedicated webpage with online step-by-step implementations of all these toolkits in R (https://yefeng0920.github.io/MA_Map_Bib/). For users with less R coding experience, we offer a guide on using three commercially free large language models (ChatGPT [https://openai.com/chatgpt/], Gemini [https://gemini.google.com/], and Copilot [https://copilot.microsoft.com/]) to modify the sample R script easily. Also, we use examples from diverse disciplines to illustrate the principles and procedures. The elegance of these proposed visualization toolkits lies in converting the qualitative information into readily digestible formats for end-users with minimal additional data extraction effort from meta-analysts. The visualization toolkits can also enrich other research synthesis methods in the research synthesis method ecosystem, including the systematic-review family (e.g., systematic reviews without meta-analysis), review-of-review family (e.g., overview, umbrella review), and rapid-review family (e.g., rapid review).[34,35]

2. Visualization toolkits for integrating evidence maps, bibliometrics, and altmetrics into meta-analysis

The three enrichment analyses of qualitative information (details see below) can be implemented in conjunction with a broad spectrum of meta-analytic approaches used across different fields, including conventional random-effects, multilevel, multivariate, dose–response, longitudinal, network, location-scale, individual participant data (IPD), phylogenetic comparative models, and meta-meta-analysis.[4,11] Below, we present the details and case studies of the three enrichment approaches.

3. Evidence map in meta-analytic evidence base

3.1. Rationale

Evidence maps (scoping reviews) represent a form of research synthesis approach that systematically collates and catalogs existing evidence on a specific topic, yet is broader than a systematic review.[36] While similar approaches exist with varying terminologies, such as mapping reviews, systematic evidence maps, and evidence and gap maps,[21–23,25,37,38] here we treat them as synonyms because they share a common core feature: mapping the state of the evidence, which makes them valuable for meta-analysis enrichment. We use the term "evidence mapping," defining it as a process that employs various visualization methods to provide a user-friendly summary of study characteristics (qualitative evidence) within a meta-analysis (the end-product).[16,22] Incorporating evidence mapping into a meta-analysis offers two benefits. First, it inherits the merits of a conventional evidence map, such as identifying knowledge gaps, directing research priorities, and providing support for funding and policy decisions.[16,25] Second, evidence mapping can inform methodological decisions (e.g., *post hoc* analyses) and interpretations of meta-analytic findings. Various visualization approaches exist, such as heatmap and bubble plots[39] and other displays used in the 3ie Evidence Gap Map
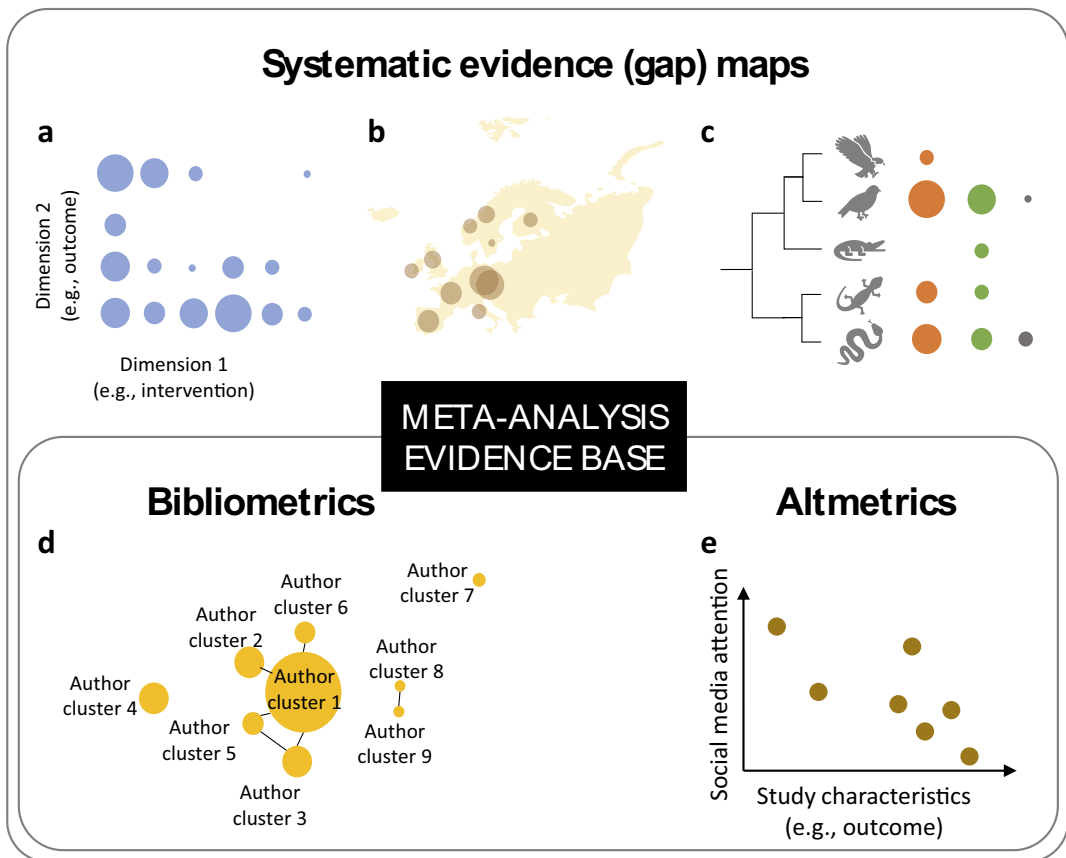
Figure 1. *A conceptual diagram of the visualization framework illustrating the use of the proposed visualization approaches to enrich meta-analyses. These "enrichment" visualization approaches are adapted from research synthesis methods in the research synthesis method ecosystem, including systematic evidence map, bibliometrics (bibliometric analysis), and alternative impact metric analysis. Presented are selected hypothetical examples for the types of plots discussed in the main text. Panels (a) (grid-like graph), (b) (geographical map), and (c) (phylogenetic tree) are examples of plots used in systematic evidence maps. Panel (d) (co-authorship network) is an example of a plot used in bibliometrics. Panel (e) is a visual representation of the relationship between social media attention (e.g., Altmetric score) and study characteristics in a meta-analysis evidence base. For real examples, refer to Sections 3–5. A step-by-step implementation guide is available at https://yefeng0920.github.io/MA_Map_Bib/.*

(https://www.3ieimpact.org/evidence-hub/evidence-gap-maps)[16] and Evidence-based Policing Matrix (https://cebcp.org/evidence-based-policing/the-matrix/),[40] for mapping the evidence landscape.

We recommend using grid-like graphs, where one dimension (e.g., intervention) is placed on the *x*-axis, and the other (e.g., outcomes) is placed on the *y*-axis, with statistics (e.g., the number of studies, number of effects, sample size) displayed at the intersection of the *x*- and *y*-axes.[18] Additional dimensions, such as study design, effect size, study quality, or RoB measures, can also be incorporated (see more on this point in the next section). We recommend using alluvial or Sankey diagrams to visualize the flow or overlaps in the composition of context-dependence drivers, summarizing their connections and co-linearity, and missing data patterns in an accessible manner. Additionally, a geographical map of study participants could reveal the representation status of historically underrepresented regions or groups, by incorporating location, gender, ethnic, or social

backgrounds into visualizations. Understanding such patterns, in turn, can facilitate actions toward greater equity, diversity, and inclusion (EDI) in research. For phylogenetic meta-analytic analyses, visual representation of species information, such as species phylogenetic trees, can intuitively convey the breadth of taxa and underlying phylogenetic heterogeneity.

3.2. Case studies

We demonstrate the application of evidence mapping using three publicly available meta-analytic datasets (Data 1 to Data 3) from various disciplines.[41–43] Note that these applications are provided for illustrative purposes, and any scientific, clinical, or policy implications should not be drawn from them. The data and code for replicating all visualizations can be found at https://yefeng0920.github.io/MA_Map_Bib/.

Data 1: Hodkinson et al.[41] conducted a network meta-analysis (including 68 studies and 80 effect size estimates) to assess the efficacy of different self-management interventions (multidisciplinary case management, regularly supported self-management, and minimally supported self-management) in enhancing the quality of life among asthma patients.

Data 2: Mertens et al.[42] employed a multilevel meta-analytic model to synthesize evidence (including 212 studies and 447 effect size estimates) on the effectiveness of choice architecture interventions (often referred to as nudges) for behavior change across various techniques, behavioral domains, and other study characteristics (e.g., populations and locations).

Data 3: Sanders et al.[43] used a Bayesian meta-analytic model to synthesize evidence (including 126 studies and 1,304 effect size estimates) regarding the impacts of artificial light at night on physiological, phenological, life history, activity patterns, and population/community-based outcomes. This meta-analysis included more than 180 species. For illustration, we used the subset that focused on physiological outcomes.

Evidence maps (scoping reviews) often use structured formats to map the evidence of a given field, such as participant, context, concept (PCC) and units, treatments, outcomes, or settings (UTOS).[18,21] For Data 1, we used grid-like graphs with the intervention variable on the *x*-axis and the outcome variable on the *y*-axis.[18] This aimed to identify knowledge gaps in self-management interventions. Each cell displayed two statistics: the number of studies and the number of effect sizes. We further quantified the population mean effect size estimate for each cell using a robust point and variance estimation approach.[44] For Data 2, we utilized an alluvial diagram to illustrate the connections (collinearities) between moderator variables related to the hypotheses tested. Data 3 featured a phylogenetic tree to visualize the taxonomic breadth and relatedness among the species involved. Data processing and visualizations were conducted using packages *ggplot2*,[45] *ape*,[46] *ggtree*,[47] and *metafor*.[48]

3.3. Results and discussion

Figure 2 shows the landscape of evidence on the effectiveness of the self-management interventions in Data 1. For example, panels (a) and (b) visually represent where randomized controlled trials (RCTs) are available to examine the interventions' effectiveness, offering insights into which interventions have received more clinical attention. Panel (c) conveys critical information about both which interventions have been examined in the RCTs and their associated effectiveness. Panel (d) reveals a potential demographic bias, as most self-management interventions were trialed on adults. This information can inform funding strategies (e.g., funding missing RCTs, or in-depth investigation) and help clinicians gauge the volume and effectiveness of their interventions of interest.

Figure 3 highlights the diversity of experimental designs in the primary studies included, suggesting potential heterogeneity in the meta-analytic evidence in Data 2. The tested moderator variables display minimal collinearity, indicating that each variable represents a unique contextual influence. Importantly, Figure 3 provides useful visual clues to identify the contexts requested by decision-makers, facilitating the assessment of the effectiveness of interventions in the context of interest (e.g., target
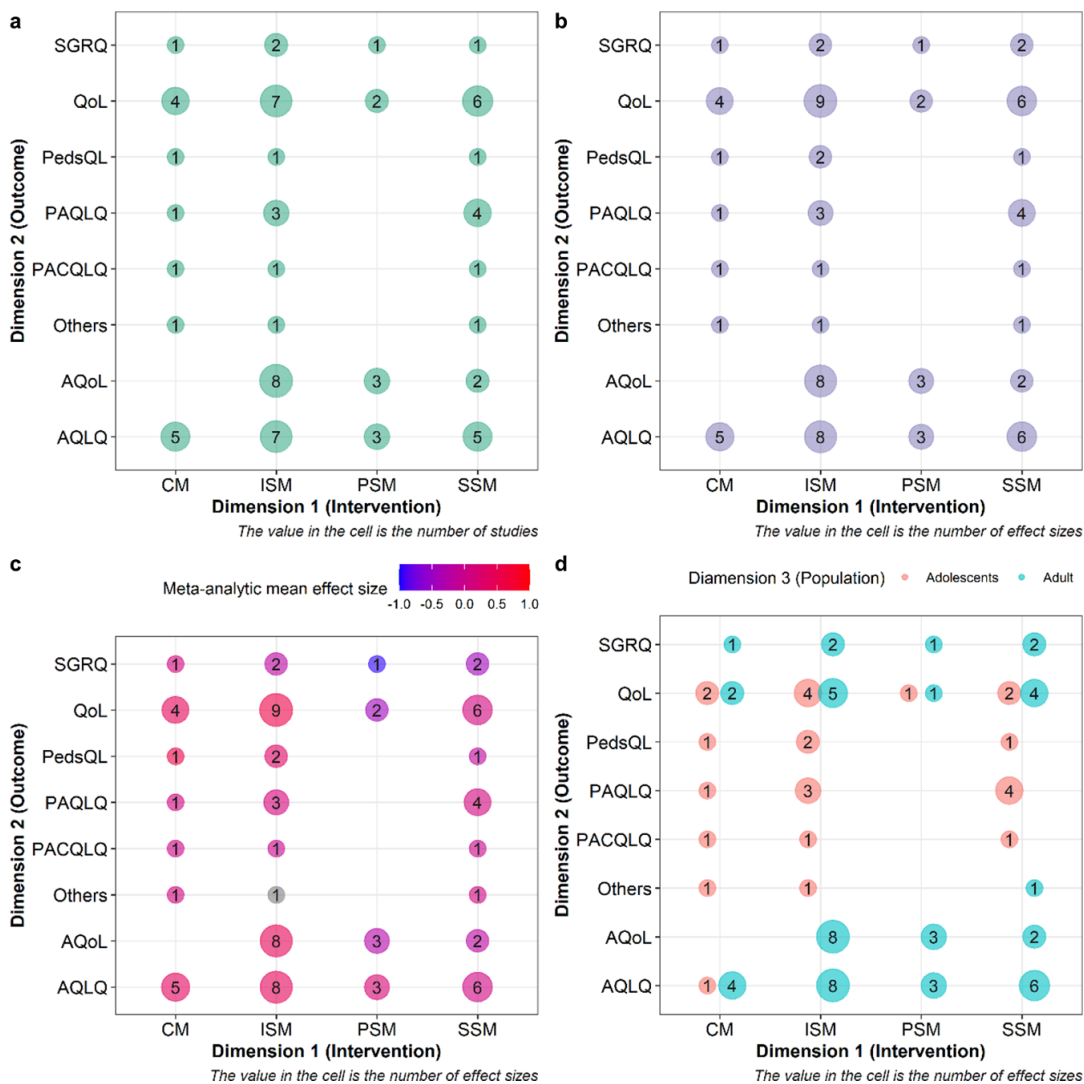
Figure 2. *Examples of evidence maps visualizing study characteristics. (a) A typical grid-like graph with the intervention variable as the first dimension, the outcome variable as the second dimension, and the bubble size representing the number of studies. (b) The bubble sizes are changed to represent the number of effect sizes. (c) The color scale is applied to the bubbles to denote the magnitude of the overall mean effect size. Note that the application of Figure 2 requires that outcomes and interventions (or any other two variables reflecting study characteristics) could be categorized into different levels via moderators, as shown in Figure 2. Hedge's g was used as the effect size measure to quantify mean differences in outcomes between an intervention and a control group. A multivariate fixed effect model was used to derive the overall average effect of each paired combination of intervention and outcome type.[44] (d) The population variable is mapped to the shape serving as the third information dimension. The full names of different types of interventions are as follows: CM = case management, ISM = intensive self-management, PSM = pure self-management, SSM = support self-management. The full names of different types of outcomes are as follows: SGRQ = St Georges Respiratory Questionnaire, QoL = Quality of Life Scale, PedsQL = Pediatric Quality of Life Inventory Generic Core Scales, PAQLQ = Pediatric Asthma Quality of Life Questionnaire, PACQLQ = Pediatric Asthma Caregiver's Quality of Life Questionnaire, AQoL = Assessment of Quality of Life, and AQLQ = Asthma Quality of Life Questionnaire. For the details of the interventions and outcomes, refer to the original paper.[41] R code was adapted from Polanin et al.[18]*
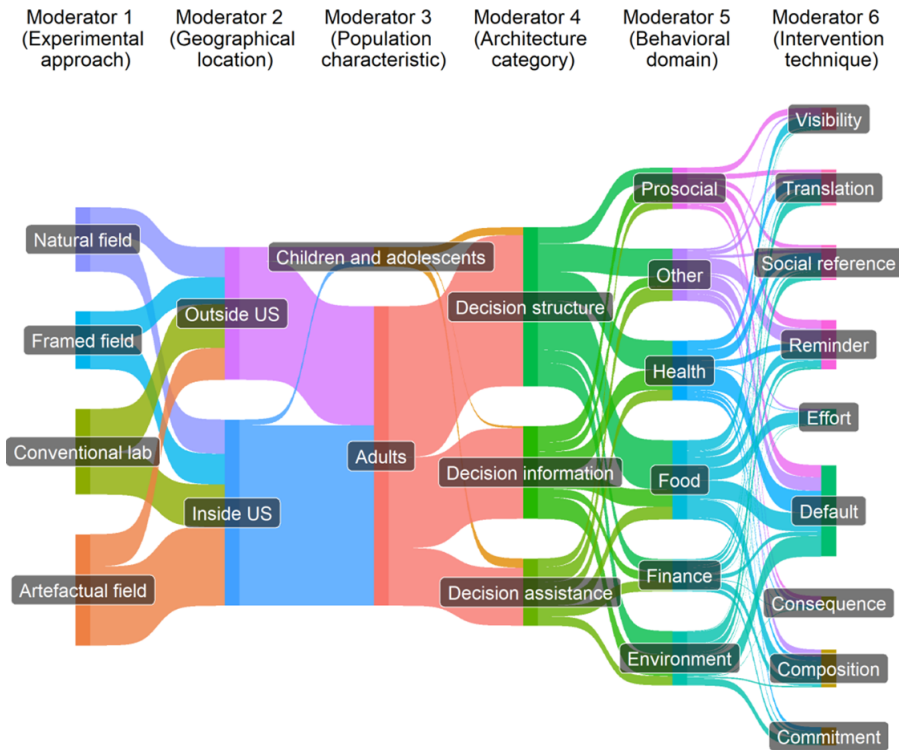
Figure 3. *An example of an alluvial diagram showing the overlaps in the composition of moderator variables considered as potential context-dependence drivers of effect sizes in a meta-analysis.*

population and location). A follow-up customizable evidence synthesis can be conducted to improve the generalizability and transferability of meta-analytic evidence (see [17] for more details).

Figure 4 presents a typical phylogenetic tree revealing the broad coverage of taxa used in artificial light at night experiments, including birds, mammals, insects, reptiles, and arachnids in Data 3. For a more in-depth statistical analysis, constructing a phylogenetic correlation matrix can quantify the effect of the shared evolutionary history among species in the meta-analytic evidence base.

4. Bibliometric map in meta-analytic evidence base

4.1. Rationale

Bibliometric analysis, often called bibliometrics, is a technique that effectively synthesizes bibliometric data to provide an overview of the performance and intellectual structure of a field.[49] It holds a unique position in the research synthesis method ecosystem because it synthesizes the bibliometric characteristics (literature metadata, not research content),[26] such as authors, countries, funders, and citation counts.[50] Literature-level characteristics are to scientific papers as epidemiology is to patients.[51] Diagnosing the "epidemiological susceptibility risk" of studies can uncover hidden biases across the evidence base. Such a process can also extend the scope of RoB assessment from within-study to between-study levels (note that between-study RoB is usually known as publication bias[52] while within-study RoB is just "risk of bias"). Studies are often not independent in terms of their conduct, with the same researchers potentially involved in multiple studies. Thus, highly represented authors (or research groups) may potentially dominate the production of evidence, limiting methodological diversity and generalizability. Indeed, centralized collaboration communities, which involve shared authors and employ similar protocols, have been shown to yield less replicable findings than decentralized ones,
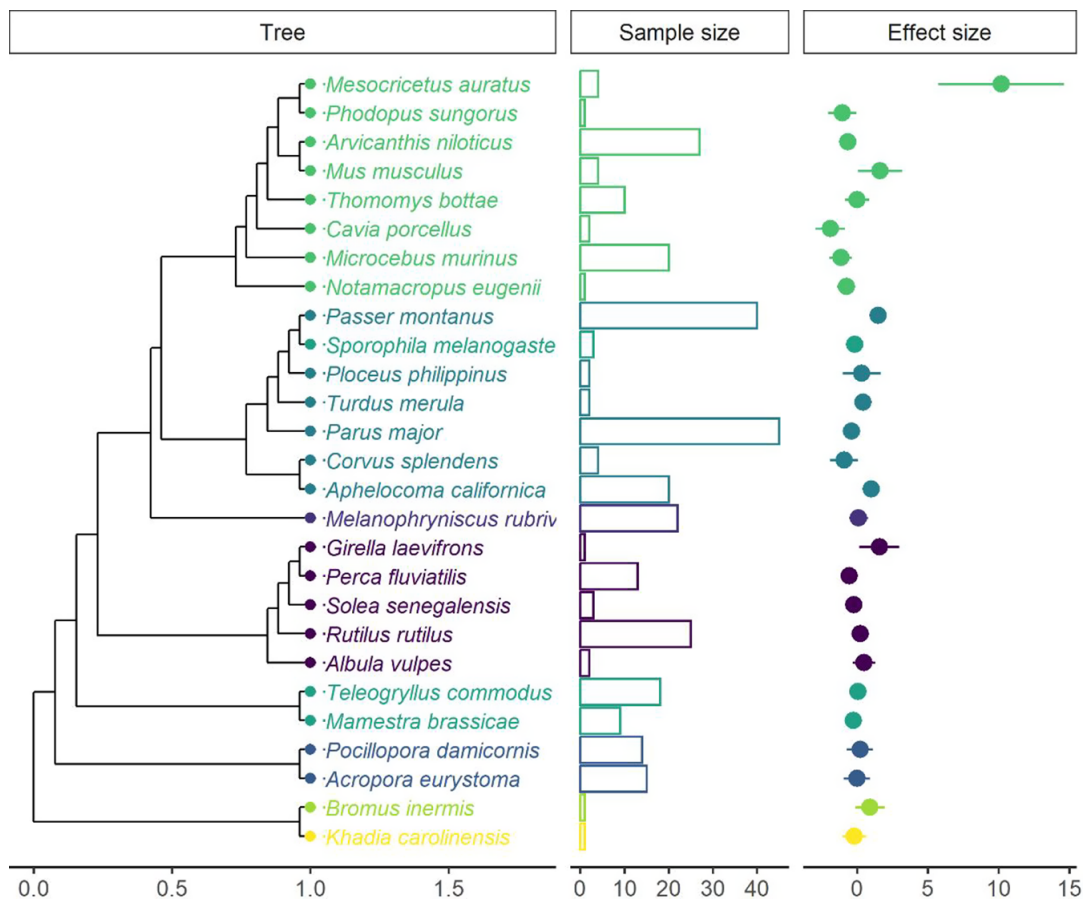
Figure 4. *An example of a phylogenetic tree visualizing the breadth of taxa and underlying phyloge-netic heterogeneity. The effect size estimates for each species were weighted based on the inverse of the sampling variance–covariance matrix. Confidence intervals for each species were constructed using a Ward-type approach, with standard errors as the square root of the sampling variance of the weighted mean. For simplicity, we assumed a sampling correlation of 0.5 for the calculation of confidence intervals (the actual correlation should be informed from ancillary data or guesstimates by domain experts). The dot represents the average mean effect size of each species. The whisker represents the 95% confidence interval. Different colors represent different phylogenetic classes. We approximated the branch lengths of the phylogenetic tree, using the Grafen method, setting the power parameter of 1 to adjust the "height" of branch lengths at the tips of the phylogenetic tree. The correlation matrix of the phylogenetic tree was estimated, assuming that the evolution of the trait follows the Brownian motion.*

which incorporate more diverse independent research groups.[53] Moreover, the publication of false, exaggerated, and falsified effects is believed to be more common in countries with a "publish or perish" culture.[54,55] For example, meta-scientific evidence suggests that studies from the United States tend to overestimate effect sizes and exhibit a larger publication bias.[56–58]

Bibliometrics can bring two major additions to meta-analyses. Bibliometrics can construct a co-authorship network for the studies included in a meta-analysis. Mapping the co-authorship network plot can intuitively reveal authorship dependence.[29] Network metrics, such as the degree of centrality, can provide quantitative insights. Bibliometrics can also identify dominant and unrepresented countries of author affiliation, or reveal interdependences between countries. Following similar principles,

bibliometrics can identify research location bias (e.g., studies from high-impact or predatory journals),[59,60] funding source bias,[61,62] linguistic bias,[63,64] and time-lag bias.[65,66] Notably, a quantitative assessment of those factors can be accomplished by conducting subgroup analysis or meta-regression.[4]

4.2. Case studies

We use Data 3[43] to illustrate the application of bibliometric analysis in meta-analytic evidence base. We collected bibliographic data for the primary studies included in the meta-analysis from Scopus using digital object identifier) searches. We created two bibliographic network typologies using the associated data: co-authorship and country of affiliation. The co-authorship network was constructed through collaboration analysis by multiplying two matrices, *Author × Paper* and *Paper × Author*. For the country network, bibliographic coupling was used, involving multiplication of *Paper × Cited paper* and *Cited paper × Paper*. We used *bibliometrix* package[27] to construct network. Packages *igraph*[67] and *circlize*[68] were used to project the networks.

4.3. Results and discussion

Figure 5 visually depicts the co-author network, with vertices (circles) representing authors and edges (lines) representing co-authorships in Data 3. Highly represented authors and dominant research groups can produce the majority of the evidence, reducing its methodological diversity and generalizability (certain groups might report certain results that are caused by bias in experimental designs and analyses).[29]

In the example meta-analytic data set, we observe nearly 90 author clusters, each of which, on average, included six authors. The largest cluster had 58 authors. To quantitatively detect the effect of hyper-dominant research groups on meta-analytic estimates, a leave-one-out analysis can be employed[4] to compare the meta-analytic effect size estimates of each research group with that of the rest. Figure 6 shows the country of affiliation citation network with 34 countries of author affiliation. While there is no obvious indication of dominant countries (with country contributions shown as proportion of the circle's perimeter), the United States, the United Kingdom, and Germany are the most prominent players in this field, and all countries appear to be linked well via article citations. At the statistical follow-up work, a multilevel meta-analytic model with random effects at the levels of author and country of affiliation clusters can be used to correct potential biases in country influence.[4,29]

5. Alternative metric analysis of meta-analytic evidence bases

5.1. Rationale

Analysis of alternative impact metrics, termed "Altmetrics," is an approach used to estimate research impact beyond academia.[33] It can quantify societal impact by tracking activities on social media (e.g., Facebook and X [formerly Twitter]), Internet pages (e.g., Wikipedia), policy-related documents, and patent applications.[33,69–71] Incorporating Altmetrics into a meta-analysis can reveal the uptake of research findings by stakeholders outside academia. For example, funding agencies have explored utilizing Altmetrics to assess public engagement with the research they support.[72–74] Similarly, policymakers have used it to determine early engagement with research within the policy domain.[75] We recommend visualizing the overall scores (e.g., Altmetric attention score), policy citations, patent citations, and overall social media attention received by studies within a meta-analysis to identify the extent to which these studies are translated into policy, practical applications, or general knowledge dissemination, respectively.
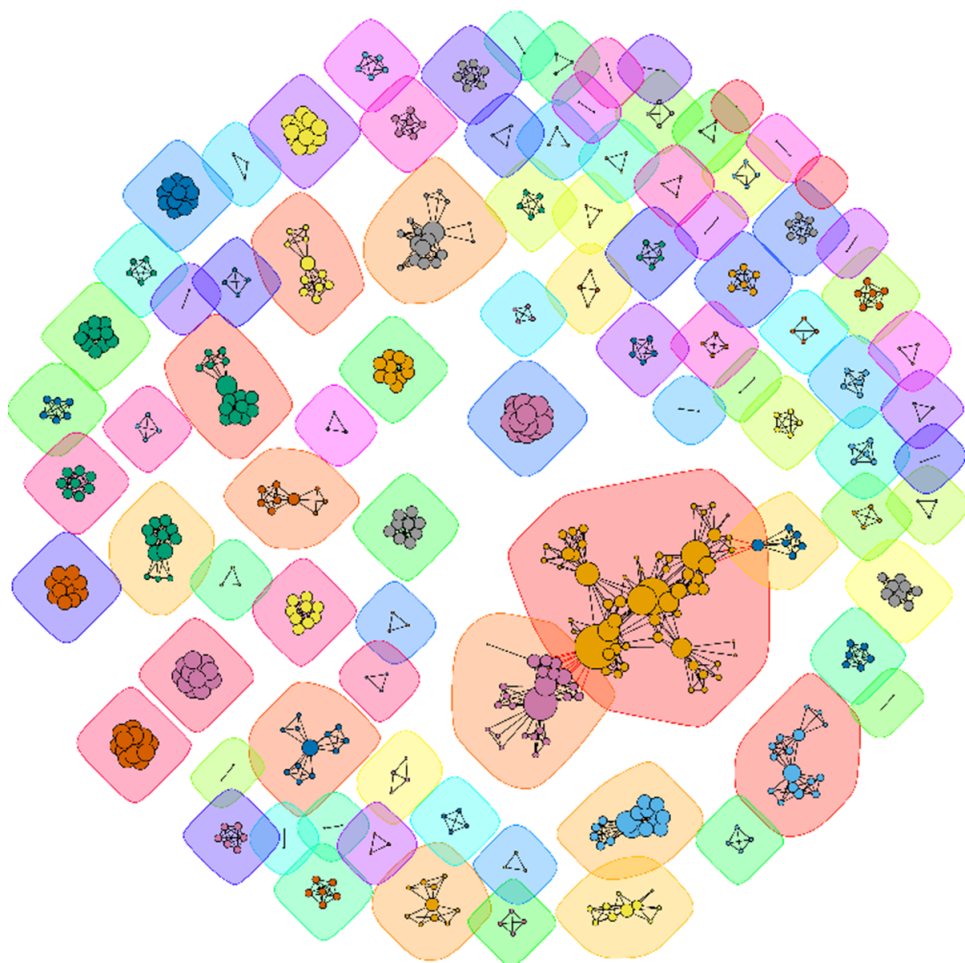
Figure 5. *An example of a co-authorship network visualizing the diversity of the research groups and the degree of centralization of the scientific community. The vertices (nodes) and edges (links) denote authors and co-authorships, respectively. Bubbles of the same color represent the same author cluster or research team. Each color denotes a co-authorship cluster (or research group). The figure was inspired by Moulin et al.,[29] who originally implemented it using Matlab and VOSviewer.*

5.2. Case studies

We use one meta-analytic data to examine the replicability of the preclinical cancer biology.[76] to illustrate the application of Altmetrics analysis, based on data from Altmetrics online service (www.altemetric.com) (Data 4). All data and code necessary for reproducing the subsequent visualizations can be accessed at the https://yefeng0920.github.io/MA_Map_Bib/. The data used for this demonstration are derived from the work of Errington et al.,[76] who conducted a meta-analysis on the differences between the original effect size estimates and replicated effect size estimates across 50 cancer biology experiments. These experiments were sourced from 23 papers published in high-profile journals, such as *Nature*, *Science*, *Cell*, and *PNAS*.

We obtained Altmetric scores for each paper included in the meta-analytic dataset using the Altmetrics online service. This online service also provides data on the counts of policy documents and patent applications directly mentioned by organizational websites. We used API (application programming interface) to automatically retrieve the total Altmetric score, policy, and patent citation
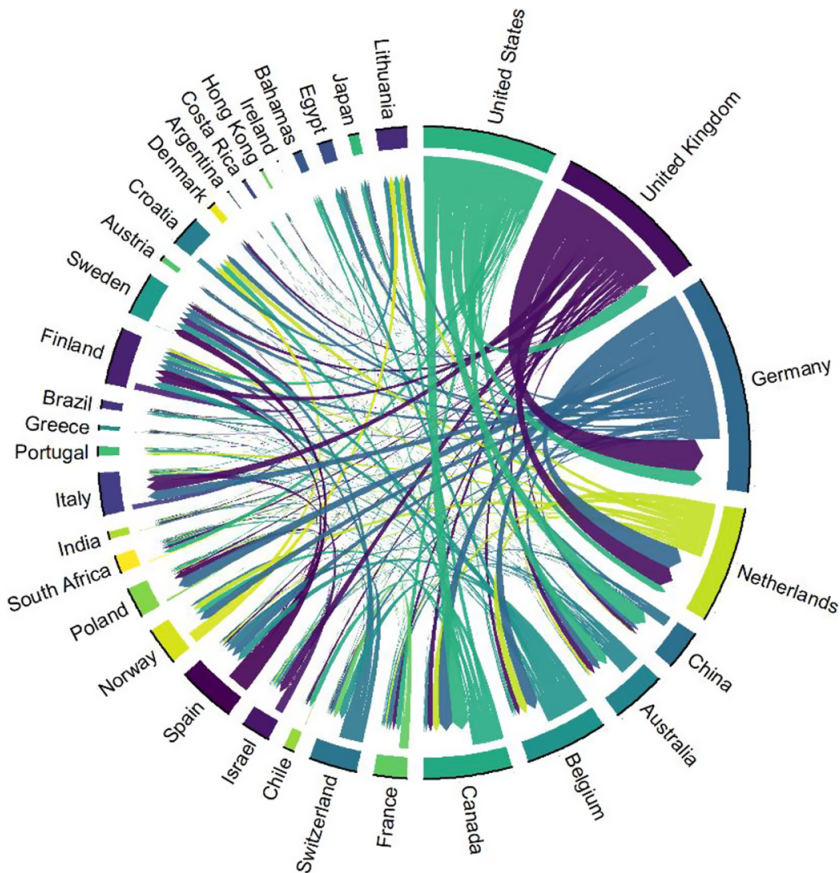
Figure 6. *An example of a chord diagram showing the epistemological interdependences between different countries of author affiliation in the meta-analytic evidence base. These inter-dependences are quantified using a bibliographic coupling approach. Two countries are coupled when the cumulative bibliographies of their respective papers share one or more cited references. The coupling strength, an indicator of dominance, increases as the number of co-cited references between them increases.*

for each original publication. We visually represented the results using the orchard plot[77] and grid-like graph.[18]

5.3. Results and discussion

Figure 7a indicates the substantial social media attention earned by cancer biology studies included in the meta-analytic evidence base. Many of these studies were mentioned in policy documents and patent applications, indicating a potential degree of practical translation. Notably, fully replicated studies exhibited relatively higher Altmetric scores, and larger policy and patent citation counts compared to that of studies that were only partially replicated and not replicated. Among the fully replicated studies, those published in *PNAS*, *Cancer Biology*, and *Cell*, exhibited higher impact metrics than those published in *Nature* and *Science*.

6. Conclusions

Meta-analysis has become a widely accepted synthesis method across various disciplines. Meanwhile, scoping reviews (evidence maps), bibliometrics, and altmetrics are gaining popularity. We leverage the
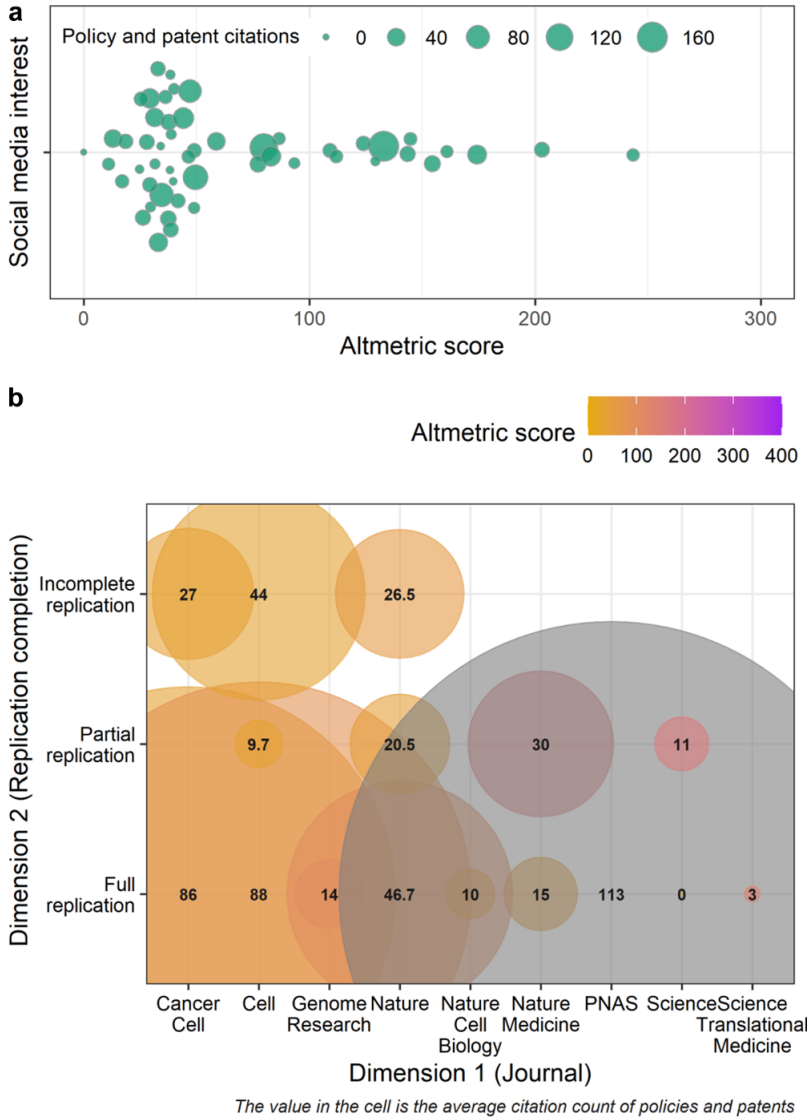
Figure 7. *Examples of visualization showing the results of impact metric analysis. (a) An orchard plot with Altmetric score as the bubble and impact metric related to indicators of practical application (i.e., patent and policy citation counts) as the bubble size. (b) A grid-like graph where: 1) the color and size of the bubbles correspond to the Altmetric score, and 2) the impact metric counts, related to the indicator of practical translation (i.e., patent and policy citation counts), respectively. The grey bubble indicates that the Altmetric score exceeds 400. The categories of "full," "partial," and "no replication" denote that replication of studies was as fully replicated, partially replicated, or not replicated, respectively.*

visualization approaches used in these three research synthesis methods to synthesize both quantitative and qualitative evidence. For meta-analysts, the proposed toolkits require minimal effort to locate, screen, code, and appraise additional evidence, as these procedures are typically integral to a standard meta-analysis. For end-users (e.g., researchers, policymakers, and practitioners), they summarize the qualitative part of the meta-analytic evidence base in immediately digestible graphs enriched with additional dimensions. These graphs can better visualize the landscape of the meta-analytic evidence

base, facilitate the identification of knowledge gaps and methodological decisions, uncover hidden biases, and map the societal impact and degree of research translation. To facilitate the adoption of the proposed visualization toolkits, we have created a dedicated website offering step-by-step implementation guides using open-source software R (https://yefeng0920.github.io/MA_Map_Bib/). Given the dynamic nature of data visualization methods and R packages, we are committed to regularly updating our website to incorporate potentially useful emerging visualization toolkits to provide a "living" arsenal for visualizing qualitative information in the meta-analytic evidence base.

References

[1] Nikolakopoulou A, Chaimani A. More than words: novel visualizations for evidence synthesis. *Res Synth Method*. 2021;12(1): 2–3.

[2] Schild AH, Voracek M. Less is less: a systematic review of graph use in meta-analyses. *Res Synth Method*. 2013;4(3): 209–219.

[3] Anzures-Cabrera J, Higgins JP. Graphical displays for meta-analysis: an overview with suggestions for practice. *Res Synth Methods.* 2010;1(1): 66–80.

[4] Yang Y, Macleod M, Pan J, Lagisz M, Nakagawa S. Advanced methods and implementations for the meta-analyses of animal models: current practices and future recommendations. *Neurosci Biobehav Rev*. 2022;146: 105016.

[5] Seidler AL, Hunter KE, Cheyne S, Ghersi D, Berlin JA, Askie L. A guide to prospective meta-analysis. *BMJ*. 2019;367: l5342.

[6] Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ.* 1997;315(7121): 1533–1537.

[7] Pigott TD, Polanin JR. Methodological guidance paper: high-quality meta-analysis in a systematic review. *Rev Educ Res*. 2020;90(1): 24–46.

[8] Cheung MW-L. A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychol Rev.* 2019;29(4): 387–396.

[9] Stanley TD, Doucouliagos H, Giles M, et al. Meta-analysis of economics research reporting guidelines. *J Econ Surv.* 2013;27(2): 390–394.

[10] Hansen C, Steinmetz H, Block J. How to conduct a meta-analysis in eight steps: a practical guide. *Manag Rev Quart.* 2021;72: 1–19.

[11] Nakagawa S, Yang Y, Macartney EL, Spake R, Lagisz M. Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environ Evid.* 2023;12(1): 1–19.

[12] Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature.* 2018;555(7695): 175–182.

[13] Kossmeier M, Tran US, Voracek M. Charting the landscape of graphical displays for meta-analysis and systematic reviews: a comprehensive review, taxonomy, and feature analysis. *BMC Med Res Methodol.* 2020;20(1): 1–24.

[14] Tricco AC, Cardoso R, Thomas SM, et al. Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: a scoping review. *Implement Sci.* 2015;11: 1–20.

[15] Lavis J, Davies H, Oxman A, Denis J-L, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Pol.* 2005;10: 35–48.

[16] Snilstveit B, Vojtkova M, Bhavsar A, Stevenson J, Gaarder M. Evidence & gap maps: a tool for promoting evidence informed policy and strategic research agendas. *J Clin Epidemiol.* 2016;79: 120–129.

[17] Martin PA, Christie AP, Shackelford GE, et al. Flexible synthesis can deliver more tailored and timely evidence for research and policy. *Proc Nat Acad Sci.* 2023;120(25): e2221911120.

[18] Polanin JR, Zhang Q, Taylor JA, Williams RT, Joshi M, Burr L. Evidence gap maps in education research. *J Res Educ Effect.* 2023;16(3): 532–552.

[19] Purgar M, Klanjscek T, Culina A. Quantifying research waste in ecology. *Nat Ecol Evol.* 2022;6(9): 1390–1397.

[20] Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet.* 2009;374(9683): 86–89.

[21] Khalil H, Tricco AC. Differentiating between mapping reviews and scoping reviews in the evidence synthesis ecosystem. *J Clin Epidemiol.* 2022;149: 175–182.

[22] Chen Y-F, Takwoingi Y, Grove A. Mapping review could be seen as a subtype of scoping review and differentiating between the action of mapping evidence and presentation of evidence as maps may be helpful: response to Khalil et al. *J Clin Epidemiol.* 2023;157: 110–111.

[23] Haddaway NR, Bernes C, Jonsson B-G, Hedlund K. The benefits of systematic mapping to evidence-based environmental management. *Ambio.* 2016;45: 613–620.

[24] Miake-Lye IM, Hempel S, Shanman R, Shekelle PG. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Systemat Rev.* 2016;5(1): 1–21.

[25] Alahdab F, Murad MH. Evidence maps: a tool to guide research agenda setting. *BMJ Evid Based Med.* 2019;24(6): 209–211.

[26] Nakagawa S, Samarasinghe G, Haddaway NR, et al. Research weaving: visualizing the future of research synthesis. *Trend Ecol Evol.* 2019;34(3): 224–238.

[27] Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J Inform.* 2017;11(4): 959–975.

[28] Clements JC, Sundin J, Clark TD, Jutfelt F. Extreme original data yield extreme decline effects. *Plos Biol.* 2023;21(2): e3001996.

[29] Moulin TC, Amaral OB. Using collaboration networks to identify authorship dependence in meta-analysis results. *Res Synth Method.* 2020;11(5): 655–668.

[30] Trueger NS, Thoma B, Hsu CH, Sullivan D, Peters L, Lin M. The altmetric score: a new measure for article-level dissemination and impact. *Annal Emerg Med.* 2015;66(5): 549–553.

[31] Knowlton SE, Paganoni S, Niehaus W, et al. Measuring the impact of research using conventional and alternative metrics. *Am J Phys Med Rehabil.* 2019;98(4): 331–338.

[32] Nakagawa S, Koricheva J, Macleod M, Viechtbauer W. Introducing our series: research synthesis and meta-research in biology. *BMC Biol.* 2020;18(1): 1–3.

[33] Erdt M, Nagarajan A, Sin S-CJ, Theng Y-L. Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics.* 2016;109: 1117–1166.

[34] Sutton A, Clowes M, Preston L, Booth A. Meeting the review family: exploring review types and associated information retrieval requirements. *Health Inform Lib J.* 2019;36(3): 202–222.

[35] Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Inform Lib J.* 2009;26(2): 91–108.

[36] Munn Z, Pollock D, Khalil H, et al. What are scoping reviews? Providing a formal definition of scoping reviews as a type of evidence synthesis. *JBI Evid Synth.* 2022;20(4): 950–952.

[37] Saran A, White H. Evidence and gap maps: a comparison of different approaches. *Campbell Systemat Rev.* 2018;14(1): 1.

[38] Campbell F, Tricco AC, Munn Z, et al. Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same but different—the "Big Picture" review family. *Systemat Rev.* 2023;12(1): 45.

[39] Haddaway NR, Feierman A, Grainger MJ, et al. EviAtlas: a tool for visualising evidence synthesis databases. *Environ Evid.* 2019;8(1): 1–10.

[40] Lum C, Koper CS, Telep CW. The evidence-based policing matrix. *J Exper Criminol.* 2011;7: 3–26.

[41] Hodkinson A, Bower P, Grigoroglou C, et al. Self-management interventions to reduce healthcare use and improve quality of life among patients with asthma: systematic review and network meta-analysis. *BMJ.* 2020;370: m2521.

[42] Mertens S, Herberz M, Hahnel UJ, Brosch T. The effectiveness of nudging: a meta-analysis of choice architecture interventions across behavioral domains. *Proc Nat Acad Sci.* 2022;119(1): e2107346118.

[43] Sanders D, Frago E, Kehoe R, Patterson C, Gaston KJ. A meta-analysis of biological impacts of artificial light at night. *Nat Ecol Evol.* 2021;5(1): 74–81.

[44] Yang Y, Lagisz M, Williams C, Noble DW, Pan J, Nakagawa S. Robust point and variance estimation for meta-analyses with selective reporting and dependent effect sizes. *Method Ecol Evol.* 2024. https://doi.org/10.1111/2041-210X.14377.

[45] Wickham H, Chang W, Wickham MH. Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. *R package.* 2016;2(1): 1–189.

[46] Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics.* 2004;20(2): 289–290.

[47] Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Method Ecol Evol.* 2017;8(1): 28–36.

[48] Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Satist Softw.* 2010;36(3): 1–48.

[49] Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. *J Busin Res.* 2021;133: 285–296.

[50] Godin B. On the origins of bibliometrics. *Scientometrics.* 2006;68(1): 109–133.

[51] Lewison G, Devey ME. Bibliometric methods for the evaluation of arthritis research. *Rheumatology (Oxford).* 1999;38(1): 13–20.

[52] Yang Y, Sánchez-Tójar A, O'Dea RE, et al. Publication bias impacts on effect size, statistical power, and magnitude (Type M) and sign (Type S) errors in ecology and evolutionary biology. *BMC Biol.* 2023;21(1): 1–20.

[53] Danchev V, Rzhetsky A, Evans JA. Centralized scientific communities are less likely to generate replicable results. *Elife.* 2019;8: e43094.

[54] Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS One.* 2009;4(5): e5738.

[55] Anderson MS, Ronning EA, De Vries R, Martinson BC. The perverse effects of competition on scientists' work and relationships. *Sci Eng Ethics.* 2007;13: 437–461.

[56] Fanelli D, Ioannidis JP. US studies may overestimate effect sizes in softer research. *Proc Nat Acad Sci.* 2013;110(37): 15031–15036.

[57] Munafò MR, Attwood AS, Flint J. Bias in genetic association studies: effects of research location and resources. *Psychol Med.* 2008;38(8): 1213–1214.

[58] Doucouliagos H, Laroche P, Stanley TD. Publication bias in union-productivity research? *Relat Industriell.* 2005;60(2): 320–347.

[59] Jannot A-S, Agoritsas T, Gayet-Ageron A, Perneger TV. Citation bias favoring statistically significant studies was present in medical research. *J Clin Epidemiol.* 2013;66(3): 296–301.

[60] Barker TH, Pollock D, Stone JC, et al. How should we handle predatory journals in evidence synthesis? A descriptive survey-based cross-sectional study of evidence synthesis experts. *Res Synth Method.* 2023;14(3): 370–381.

[61] Barden J, Derry S, McQuay HJ, Moore RA. Bias from industry trial funding? A framework, a suggested approach, and a negative result. *Pain.* 2006;121(3): 207–218.

[62] Lynch JR, Cunningham MR, Warme WJ, Schaad DC, Wolf FM, Leopold SS. Commercially funded and United States-based research is more likely to be published; good-quality studies with negative outcomes are not. *J Bone Joint Surg.* 2007;89(5): 1010–1018.

[63] Nuñez MA, Amano T. Monolingual searches can limit and bias results in global literature reviews. *Nat Ecol Evol.* 2021;5(3): 264–264.

[64] Amano T, Berdejo-Espinola V, Akasaka M, et al. The role of non-English-language science in informing national biodiversity assessments. *Nat Sustain.* 2023: 1–10.

[65] Koricheva J, Kulinskaya E. Temporal instability of evidence base: a threat to policy making? *Trend Ecol Evol.* 2019;34(10): 895–902.

[66] Yang Y, Lagisz M, Nakagawa S. Decline effects are rare in ecology: Comment. *Ecology.* 2023;104: e4069.

[67] Csardi G, Nepusz T. The igraph software package for complex network research. *Inter J.* 2006;1695(5): 1–9.

[68] Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular visualization in R. *Bioinformatics.* 2014;30(19): 2811–2812.

[69] Bornmann L, Haunschild R, Adams J. Do altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of altmetrics based on data from the UK research excellence framework (REF). *J Informet.* 2019;13(1): 325–340.

[70] Haustein S. Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics.* 2016;108: 413–423.

[71] Thelwall M. Measuring societal impacts of research with altmetrics? Common problems and mistakes. *J Econ Surv.* 2021;35(5): 1302–1314.

[72] Dinsmore A, Allen L, Dolby K. Alternative perspectives on impact: the potential of ALMs and altmetrics to inform funders about research impact. *PLoS Biol.* 2014;12(11): e1002003.

[73] Thelwall M, Kousha K, Dinsmore A, Dolby K. Alternative metric indicators for funding scheme evaluations. *Aslib J Inform Manag.* 2016;68(1): 2–18.

[74] Woolston C. Funders drawn to alternative metrics. *Nature.* 2014;516(7530): 147–147.

[75] McCambridge J, Hawkins B, Holden C. Industry use of evidence to influence alcohol policy: a case study of submissions to the 2008 Scottish government consultation. *PLoS Med.* 2013;10(4): e1001431.

[76] Errington TM, Mathur M, Soderberg CK, et al. Investigating the replicability of preclinical cancer biology. *Elife.* 2021;10: e71601.

[77] Nakagawa S, Lagisz M, O'Dea RE, et al. The orchard plot: cultivating a forest plot for use in ecology, evolution, and beyond. *Res Synth Methods.* 2021;12(1): 4–12.