Flexible synthesis can deliver more tailored and timely evidence for research and policy

PNAS OPINION

Flexible synthesis can deliver more tailored and timely evidence for research and policy

Philip A. Martin[a,b,c,1], Alec P. Christie[a,b,d], Gorm E. Shackelford[a,b], Amelia S. C. Hood[a,e], Shengyu Wang[f,g], Bo Li[h,i], William Morgan[a,b], Madeleine Lee[a,b], David C. Aldridge[a,b], and William J. Sutherland[a,b]

For the purposes of evidence-based policymaking, research reviews are often redundant, not specific enough to address a given issue, or become dated too quickly. Such problems undermine their use in decision-making. There's a better approach. Image credit: Shutterstock/Tupungato.

The evidence synthesis and summarization that's conducted through systematic reviews has revolutionized decision-making. However, synthesis faces three serious problems. First, reviews often do not cover the issues or geographies needed for specific decisions (1). Second, duplication of reviews on similar topics hinders efficient use of research resources (1). Third, reviews rapidly become out of date (2). These problems undermine evidence use in decision-making. There are ways to overcome these challenges by providing customizable and transparent evidence synthesis that is easily updatable. Such tools would be of great use to policymakers looking to inform their decisions with current, reliable evidence.

Major efforts have been made to improve the robustness of scientific evidence (3), but we need further progress to make evidence more relevant to practitioners and policymakers (4). Even when evidence about an intervention of interest has been summarized, the relevance of any conclusions to a decision-maker may be limited if a review includes studies from contexts that differ from those of the decision-makers (3). Transferability is key here—the extent to which an intervention's effectiveness would be the same in a different context (5). Myriad factors may influence transferability, including: (i) settings, locations or environments in which interventions are carried out; (ii) the target population of a study; (iii) intervention delivery, such as its timing or the experience of the person delivering it; and (iv) how outcomes are measured. Differences in settings, target populations,

The authors declare no competing interest.

Any opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and have not been endorsed by the National Academy of Sciences.

[1]To whom correspondence may be addressed. Email: phil.martin.research@gmail.com.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2221911120/-/DCSupplemental.

Published June 14, 2023.

and intervention delivery can cause real differences in outcomes, whereas apparent differences in outcomes may disappear if study methodologies are uniform (6).

Differences in context matter. This is particularly true where the impacts of interventions differ considerably. Systematic review and meta-analysis can be valuable for informing decisions in such fields, but generalizing about intervention effectiveness may be difficult (7). Further analyses can help explain this variation, but the analyses that authors consider interesting (and, therefore, the analysis that they conduct) may not be the analysis that decision-makers desire.

> **Providing access to the highly relevant and up-to-date evidence suited to practitioners and policymakers requires a new method of evidence synthesis that builds on advances in systematic review.**

One way to deal with varying contexts is to work with decision-makers to coproduce new primary studies or systematic reviews that fit a specific context (8). However, this requires considerable investment for each decision, resulting in studies that are not cost-effective or scalable, and the evidence produced is still of limited relevance to other contexts. We think a more promising solution for delivering relevant evidence may be: (i) compiling evidence for multiple interventions across a broad topic; (ii) determining the relevance of studies to contexts requested by decision-makers; and (iii) assessing the effectiveness of interventions for the context of interest.

In addition to the problem of relevance, there is sometimes duplication of syntheses on a given topic, potentially wasting time and money (1). Furthermore, reviews can become out of date quickly due to the rapidly increasing number of research papers published every year, potentially resulting in incorrect estimates of intervention effectiveness (9). These challenges were put into stark relief at the start of the COVID-19 pandemic, when the rush to synthesize evidence produced nearly 9,000 systematic reviews (10). However, many of these syntheses repeated similar syntheses, and, due to the massive number of clinical trials being published, almost all of these are now out of date (10). The problem of syntheses becoming out of date is common in other fields, too. For example, repeating the same systematic review in internal medicine within 2 years of publication led to substantial changes in apparent treatment effects for a quarter of studies (2). One solution to this problem is to have an online platform providing details of existing and ongoing syntheses that can be dynamically updated to create "living reviews."

Timely Evidence

Providing access to the highly relevant and up-to-date evidence suited to practitioners and policymakers requires a new method of evidence synthesis that builds on advances in systematic review. Such a method should allow users to explicitly define their context in terms of the setting, target population, intervention delivery, and outcome measurements (as detailed in *SI Appendix*, Table S1), allowing users to navigate existing evidence. This method should allow decision-makers to address a wide range of questions, while allowing for flexibility relating to different contextual factors. Below, we present an implementation of "dynamic meta-analysis"—a novel method combining systematic review, meta-analysis, and an interactive web app. The key difference between dynamic and traditional meta-analysis is that dynamic meta-analysis allows users to filter and weight evidence depending on their interests to produce tailor-made analyses (see Fig. 1 for an overview). Currently, online tools typically only summarize evidence on interventions in a static and qualitative way (e.g., https://www.conservationevidence.com/ or https://www.wiseinterventions.org/) that lack flexibility.

We have developed an interactive tool based on this idea called Metadataset (https://www.metadataset.com/) (11). The database underpinning the tool contains data from primary studies on agricultural interventions (seeking to increase target crop yields or reduce their negative environmental impacts) and invasive plant management (seeking to reduce the abundance of problematic plants). These were compiled by using standard systematic review methodologies. We are expanding the database to include more data on invasive species management, which currently include 374 studies across 31 (mostly plant) invasive species.

Fig. 1 shows the path that a user might take to better understand the topic. This example shows the impacts of different types of invasive plant management, with a particular focus on controlling an invasive plant species in the genus *Spartina* (cordgrass). This path is represented by the black line and arrows in Fig. 1. Meta-analysis can produce different effect sizes, but here, we used percentage change of the outcome measures, derived from the log response ratio. In the meta-analyses in Fig. 1, after looking at the impacts of management on multiple outcomes (Fig. 1, box 1), the user elected to examine only the effect on the abundance of invasive plants (Fig. 1, box 2), then only the effect on *Spartina* species abundance (Fig. 1, box 3), and then the specific effect of chemical control by comparing different herbicides' effects on *Spartina* species abundance (box 4). The final analysis (Fig. 1, box 4) shows that *Spartina* abundance was reduced by 55% (across 11 studies with 78 comparisons) when glyphosate herbicide was used, but was not significantly reduced when other herbicides were used. This is just one of many routes that users could take to analyze outcomes, enabling them to choose routes of greatest interest to them. Another user may wish to analyze the management impacts on carbon storage or effects on native species, or filter to look at studies in North America.

Building on previous work (12), this approach was used as part of a review in January 2023, in response to Chinese Government plans to control *Spartina* across China by 2025 (13). Metadataset was used to run several distinct analyses to inform the management of *Spartina* using both the global dataset and a subset for China, suggesting that integrated controls are likely to be most effective for reducing its abundance, while chemical control also had significant control efficiency, but precise effects depended on herbicide type (14). However, standalone physical interventions only appeared to be effective in the short-term (14). The rapid

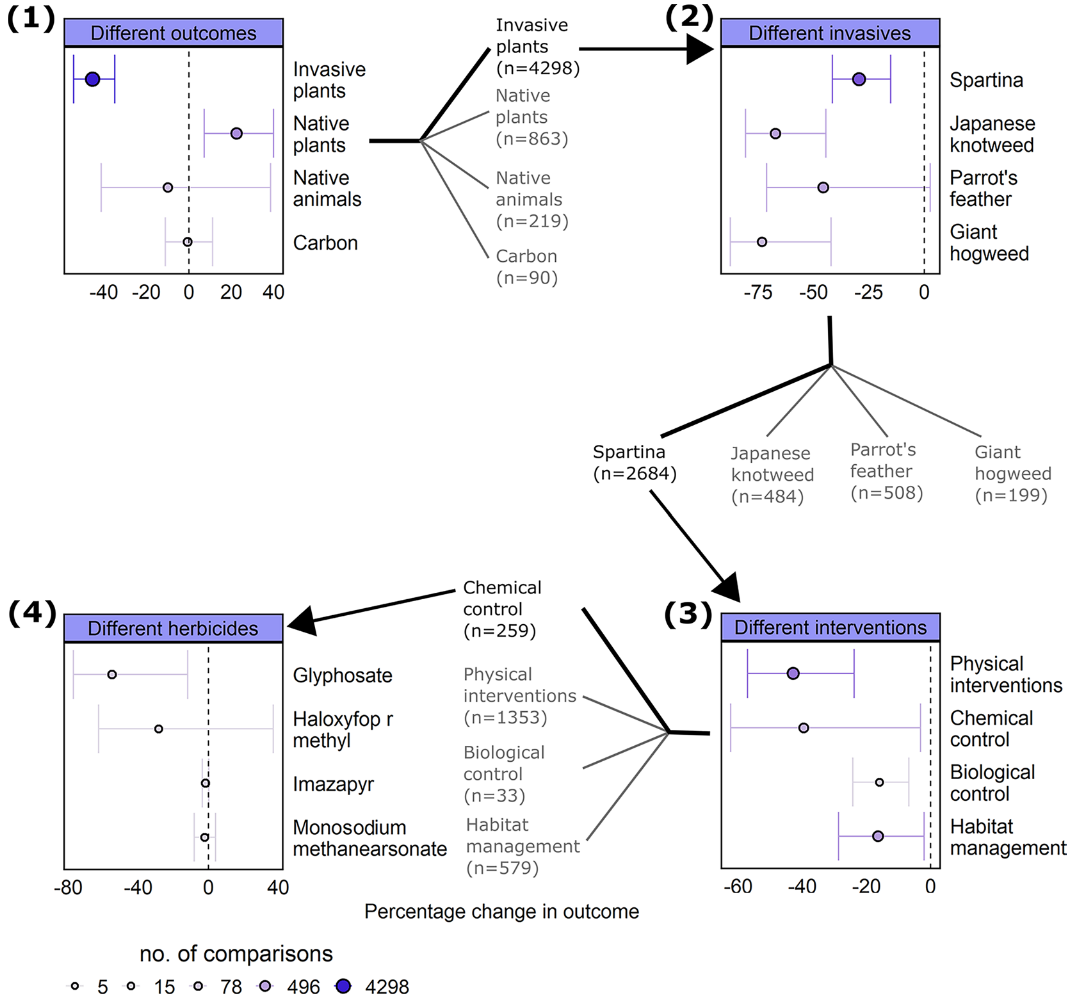

Fig. 1. **A typical workflow of an analysis on Metadataset. Each box represents a separate meta-analysis, and the lines between them represent the filtering process of selecting different populations, species, and interventions. Point size and color indicate the number of comparisons per estimate. Labels on the *y* axis refer to comparisons made at each level of meta-analysis. Branches between boxes represent potential ways of filtering the data, with the chosen route given in black. *n* shows the number of pairwise comparisons for each branch. In box 4, the meta-analysis reports data on the effectiveness of different methods of chemical control for *Spartina* on the abundance of the *Spartina*, but could equally show different methods of biological control for Parrot's feather, etc.**

meta-analyses provided the evidence base for discussions with a panel of experts on *Spartina* management, who checked the results and supplemented the review with key contextual considerations for management options, delivering evidence-based advice within a short policy window (14).

Beyond allowing users to filter out studies they regard as irrelevant, the system also allows users to recalibrate study weights so that those with higher relevance contribute more to meta-estimates. For example, a user may be interested in evidence from a particular geographic region or species, or may wish to only look at studies that used the most robust methodologies. Thus, users can identify studies with the most relevant features, with results that are likely to transfer relatively well to their own context. Doing this dynamically using an online interactive tool allows users to customize the analysis to their specific context, rather than being constrained by a static systematic review.

Not only do syntheses vary in their relevance to decision-makers, but they can quickly become out of date. One solution is using "living reviews," in which the evidence base is searched and regularly updated as studies are published. This approach is rapidly gaining acceptance in healthcare and has great potential in other fields. Metadataset can facilitate the production of living reviews, as data can be added as they are published—this platform therefore represents both a living and customizable synthesis; hence, dynamic meta-analyses are a unique type of living review. In addition, by providing a repository for the data underlying systematic reviews of interventions, our tool clearly identifies which topics have been covered or are actively being covered. Scaling up our approach could provide a centralized, open-access repository that would allow knowledge gaps to be assessed, aiding scoping of future synthesis.

Dynamic meta-analysis presents an exciting new direction for evidence synthesis and decision-making. It provides practitioners and decision-makers with quick and easy access to relevant and up-to-date evidence for mission-driven disciplines. Dynamic meta-analysis may be particularly useful for complex interventions delivered through multifaceted programs, since their results are perceived to be highly context-specific (15). Although our recent focus has been agriculture and nature conservation, we believe that dynamic meta-analysis would be useful in other mission-driven fields, such as public health, development, and education. One potentially fruitful case in point: assessing antibullying campaigns in schools, the results of which may vary widely depending on their frequency, duration, and ability to mobilize bystanders to bullying (16). Dynamic meta-analysis would allow a decision-maker to explore how the effectiveness of antibullying campaigns varies due to different contextual factors in a user-friendly, accessible format.

Realizing the Potential

Although dynamic meta-analysis is a powerful tool, there clearly is much work to be done. In the short-term, we need to: (i) expand our database to cover more topics relating to biodiversity conservation, thereby allowing more powerful analyses on a wider range of topics; and (ii) refine the tool through user testing to increase usability. There are also several barriers that must be overcome. One is the long-term maintenance of the tool, a barrier shared by living systematic reviews. To overcome this, collaboration—particularly international collaboration—with other researchers and decision-makers can facilitate the longevity of projects such as ours by providing access to diverse and long-term funding streams. We also envision governance of the tool being determined by strategic and technical boards made up of individuals from collaborating organizations. Automated searches combined with natural-language processing may also substantially reduce the resources required to update the evidence base and focus funding on the long-term functioning of the tool.

Another problem is how to ensure the trustworthiness and transparency of the data underpinning the tool. We see two main ways to do this: (i) the production of protocols detailing how data were compiled and (ii) publishing peer-reviewed papers describing the database when significant updates are made. Similarly, there may be concerns about the inappropriate use of analyses to find statistical significance for a particular result, also known as “p-hacking” or “cherry picking.” This could result from individuals selecting evidence that supports a specific agenda or searching for significant results for publication. To guard against this, our tool allows users to provide a URL link to the analysis that they carried out, thus allowing for other users to see exactly how the analysis was performed. These other users could then repeat the analysis, modifying settings as they see fit, to examine whether reported analyses are robust to other evidence, a major advantage over static meta-analysis.

Dynamic meta-analysis will not be useful in all circumstances. For example, systematic reviews that are codesigned with decision-makers who will later make use of the generated evidence can be extremely useful for informing expensive, large-scale policy decisions, where the costs of mistakes are high (17). However, most decisions are relatively small, local decisions, made by a small team (17). In these cases, dynamic meta-analysis has the potential to provide quantitative answers for problems where context is important, decisions need to be made relatively rapidly, and costs must be kept low. We also recognize that not all fields, particularly those where research is not translated into practice, will find dynamic meta-analysis useful.

We envision a future in which dynamic meta-analysis can be used to address questions when answers are needed rapidly or where resource constraints preclude targeted systematic review. To achieve this, organizations that operate at the science–policy interface should train and hire scientists who are able to interpret meta-analyses and critically appraise scientific evidence to act as intermediaries between researchers and decision-makers. If well-implemented, this step could aid society in overcoming many pressing global-scale challenges that require effective interventions. Although these challenges are global in nature, their solutions often must be delivered at national or subnational scales. Synthesis that generates locally relevant evidence is therefore an important step. But whatever the scale, better, more tailored, timely, and reliable evidence will give researchers and policymakers a better chance to solve critical problems in the years to come.

Author affiliations: [a]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom; [b]BioRISC (Biosecurity Research Initiative at St Catharine’s), St Catharine’s College, Cambridge CB2 1RL, United Kingdom; [c]Basque Centre for Climate Change (BC3), Leioa 48940, Spain; [d]Downing College, Cambridge CB2 1DQ, United Kingdom; [e]Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AR, United Kingdom; [f]National Observations and Research Station of Wetland Ecosystems of the Yangtze Estuary, Institute of Biodiversity Science, School of Life Sciences, Fudan University, Shanghai 200438, China; [g]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Coastal Ecosystems Research Station of the Yangtze River Estuary, Institute of Biodiversity Science, Fudan University, Shanghai 200438, China; [h]Yunnan Key Laboratory of Plant Reproductive Adaptation and Evolutionary Ecology, Institute of Biodiversity, School of Ecology and Environmental Science, Yunnan University, Kunming 650504, China; and [i]Centre for Invasion Biology, Institute of Biodiversity, School of Ecology and Environmental Science, Yunnan University, Kunming 650504, China

Author contributions: P.A.M., A.P.C., G.E.S., B.L., and W.J.S. designed research; P.A.M., A.P.C., G.E.S., A.S.C.H., S.W., B.L., W.M., and M.L. performed research; P.A.M. analyzed data; P.A.M., A.P.C., G.E.S., A.S.C.H., D.C.A., and W.J.S. wrote the paper.

1. J. P. A. Ioannidis, The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* **94**, 485–514 (2016).
2. K. G. Shojania *et al.*, How quickly do systematic reviews go out of date? A survival analysis. *Ann. Intern. Med.* **147**, 224–233 (2007).
3. P. M. Rothwell, External validity of randomised controlled trials: “To whom do the results of this trial apply?” *Lancet* **365**, 82–93 (2005).
4. K. Oliver, S. Innvar, T. Lorenc, J. Woodman, J. Thomas, A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Serv. Res.* **14**, 2 (2014).
5. P. R. A. Baker, J. J. Shipp, S. H. Wellings, N. Priest, D. P. Francis, Assessment of applicability and transferability of evidence-based antenatal interventions to the Australian indigenous setting. *Health Promot. Int.* **27**, 208–219 (2012).
6. J. A. Catford, J. R. U. Wilson, P. Pyšek, P. E. Hulme, R. P. Duncan, Addressing context dependence in ecology. *Trends Ecol. Evol.* **37**, 158–170 (2022).
7. D. Kneale, J. Thomas, A. O’Mara-Eves, R. Wiggins, How can additional secondary data analysis of observational data enhance the generalisability of meta-analytic evidence for local public health decision making? *Res. Synth. Methods* **10**, 44–56 (2019).
8. N. R. Haddaway *et al.*, A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environmental Evidence* **6**, 11 (2017).
9. J. Koricheva, E. Kulinskaya, Temporal instability of evidence base: A threat to policy making? *Trends Ecol. Evol.* **34**, 895–902 (2019).
10. H. Pearson, How COVID broke the evidence pipeline. *Nature* **593**, 182–185 (2021).
11. G. E. Shackelford *et al.*, Dynamic meta-analysis: A method of using global evidence for local decision making. *BMC Biol.* **19**, 33 (2021).
12. A. Roberts, A. S. Pullin, The effectiveness of management interventions for the control of *Spartina* species: A systematic review and meta-analysis. *Aquat. Conserv.* **18**, 592–618 (2008).
13. E. Stokstad, China battles alien weed at unprecedented scale. *Science* **379**, 972 (2023).
14. S. Reynolds Spartina invasive management–a review of the evidence. OSF Preprints [Preprint] (2023). https://osf.io/3bwzx/ (Accessed 1 March 2023).
15. D. Nilsson, G. Baxter, J. R. A. Butler, C. A. McAlpine, How do community-based conservation programs in developing countries change human behaviour? A realist synthesis. *Biol. Conserv.* **200**, 93–103 (2016).
16. E. Menesini, C. Salmivalli, Bullying in schools: The state of knowledge and effective interventions. *Psychol. Health Med.* **22**, 240–253 (2017).
17. W. J. Sutherland, H. Downey, W. F. Frick, P. Tinsley-Marshall, T. McPherson, Planning practical evidence-based decision making in conservation within time constraints: The Strategic Evidence Assessment Framework. *J. Nat. Conserv.* **60**, 125975 (2021).