**Flexible synthesis to deliver bespoke evidence for decision-making**

Authors: Philip A. Martin1,2,3, Alec P. Christie1,2,4, Gorm E. Shackelford1,2, Amelia S. C. Hood1,5, Shengyu Wang6, Li Bo6, William Morgan1,2, Maddy Lee1,2, David Aldridge1,2, William J. Sutherland1,2

Affiliations:

1. Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK.

2. BioRISC (Biosecurity Research Initiative at St Catharine’s), St Catharine’s College, Cambridge, CB2 1RL, UK.

3. Basque Centre for Climate Change (BC3), Edificio sede no 1, planta 1, Parque científico UPV/EHU, Barrio Sarriena s/n, 48940, Leioa, Bizkaia, Spain.

4. Downing College, Regent Street, Cambridge, CB2 1DQ.

5. Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, Earley Gate, PO Box 237, Reading, RG6 6AR, UK.

6. National Observations and Research Station of Wetland Ecosystems of the

Yangtze Estuary, and Ministry of Education Key Laboratory for Biodiversity

Science and Ecological Engineering, Coastal Ecosystems Research Station

of the Yangtze River Estuary, Institute of Biodiversity Science, School of Life

Sciences, Fudan University, #2005 Songhu Road, Shanghai, 200438, China

‘How effective will this intervention be for my needs?’ This is the critical question for almost any decision in mission-driven fields, such as medicine, conservation biology, or education. However, this question is often difficult to answer [(*1*, *2*)](https://paperpile.com/c/u4TPYy/gPod+qenN). Scientific evidence in the form of primary studies or evidence synthesis can help to answer this question, but such evidence should ideally be robust, relevant, and up-to-date [(*3*)](https://paperpile.com/c/u4TPYy/m8Uq). Huge amounts of resources have been dedicated to improving the robustness of scientific evidence [(*2*)](https://paperpile.com/c/u4TPYy/qenN), but the importance of relevant evidence has received relatively little attention [(*3*)](https://paperpile.com/c/u4TPYy/m8Uq). Even when robust evidence exists for an intervention of interest much of it is of limited relevance to decision-makers' needs because it relates to contexts that differ from those of the decision-maker [(*1*, *2*)](https://paperpile.com/c/u4TPYy/qenN+gPod). This is compounded by the fact that the evidence used to make decisions is often out-of-date [(](https://paperpile.com/c/u4TPYy/fnE6)*[4](https://paperpile.com/c/u4TPYy/fnE6)*[)](https://paperpile.com/c/u4TPYy/fnE6). Without access to robust, relevant, up-to-date evidence policymakers and practitioners may not use scientific evidence to inform their decisions [(*5*, *6*)](https://paperpile.com/c/u4TPYy/8evE+Fp45), which risks harmful and inefficient outcomes.

It is common sense that interventions will differ in their effectiveness in vastly different contexts. However, even subtle differences can influence how well the effectiveness of one intervention translates to another context. This relates to transferability - the extent to which an intervention would have the same effect in a different context [(](https://paperpile.com/c/u4TPYy/m2Nx)*[7](https://paperpile.com/c/u4TPYy/m2Nx)*[)](https://paperpile.com/c/u4TPYy/m2Nx). There are myriad factors that may influence the transferability of intervention effects from one context to another, but broadly we think that these can be grouped into differences in (i) settings, (ii) target populations, (iii) how the intervention is delivered, and (iv) how outcomes are measured (see Table 1 for details of these). Differences in settings, target populations, and intervention delivery can cause real differences in outcomes, whereas differences in outcome measurement can cause apparent differences that may disappear if study methodologies were uniform [(*8*)](https://paperpile.com/c/u4TPYy/Vn9J).

Differences in context matter; particularly in fields that are typified by high levels of complexity, such as conservation biology [(*8*)](https://paperpile.com/c/u4TPYy/Vn9J), public health [(*9*)](https://paperpile.com/c/u4TPYy/2gKK), and education [(*10*)](https://paperpile.com/c/u4TPYy/XcBn). Evidence synthesis in the form of systematic review combined with meta-analysis can be valuable for informing decisions, but where intervention outcomes are highly context dependent it may be difficult to generalise about intervention effectiveness [(*9*)](https://paperpile.com/c/u4TPYy/2gKK). Subgroup analysis or meta-regression can be used to explain this variation, but a paucity of studies often precludes detailed analyses covering a range of different contexts [(*9*)](https://paperpile.com/c/u4TPYy/2gKK).

One solution to this lack of relevant evidence is to co-produce new primary studies or systematic reviews with decision-makers that fit a specific context exactly [(*11*, *12*)](https://paperpile.com/c/u4TPYy/TAfK+eSAa). However, this requires time and money, it may be impractical for some decisions, and it still produces evidence that is of limited relevance for people working in other contexts [(*13*)](https://paperpile.com/c/u4TPYy/dfBH). An alternative solution involves determining the relevance of existing sources of evidence with respect to the local context, followed by an assessment of how effectiveness may change from one context to the next [(*9*, *13*,](https://paperpile.com/c/u4TPYy/dfBH+tzkX+2gKK) *[14](https://paperpile.com/c/u4TPYy/dfBH+tzkX+2gKK)*[)](https://paperpile.com/c/u4TPYy/dfBH+tzkX+2gKK). We think that this latter option presents a promising route for delivering evidence.

Not only does evidence synthesis vary in its relevance to decision-makers, but it is also often out-of-date. This is because as new primary studies are carried out, the apparent effectiveness of interventions changes [(*15*)](https://paperpile.com/c/u4TPYy/06kd). For example, a review of systematic reviews in internal medicine found that nearly a quarter of them were out-of-date within two years of publication [(*16*)](https://paperpile.com/c/u4TPYy/qrIz). This problem has been brought into stark focus during the COVID-19 pandemic - as promising new treatments emerged and were tested, systematic reviews summarising their effectiveness were produced. However, these systematic reviews were often out-of-date before they were even published [(*17*)](https://paperpile.com/c/u4TPYy/6o94). One solution to this is the use of ‘living reviews’ in which the evidence base is searched and regularly updated as new studies are published. One example of this is the Australian Stroke Foundation, who have reduced the time between guideline updates from 7 years to 3 months as a result of using a living review approach [(*17*)](https://paperpile.com/c/u4TPYy/6o94). This approach is rapidly gaining acceptance in healthcare and we think that it has great potential in other mission driven fields.

To provide access to the highly-relevant and up to date evidence that practitioners and policymakers need, we believe there is a need for a new method of evidence synthesis that builds on advances made 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 Table 1) in order to facilitate comparisons with the existing evidence base. This system should allow decision-makers to address a wide range of questions while allowing for flexibility relating to different contextual factors.

Below we outline how providing a database compiled using systematic review methods combined with an interactive web app can provide a possible solution to the barriers relating to relevance and access.

**Table 1** - Different contextual factors that may influence how well intervention effects transfer from one context to another. This list is not exhaustive, see [(*2*)](https://paperpile.com/c/u4TPYy/qenN) for examples from clinical medicine, [(*1*)](https://paperpile.com/c/u4TPYy/gPod) for examples from biodiversity conservation

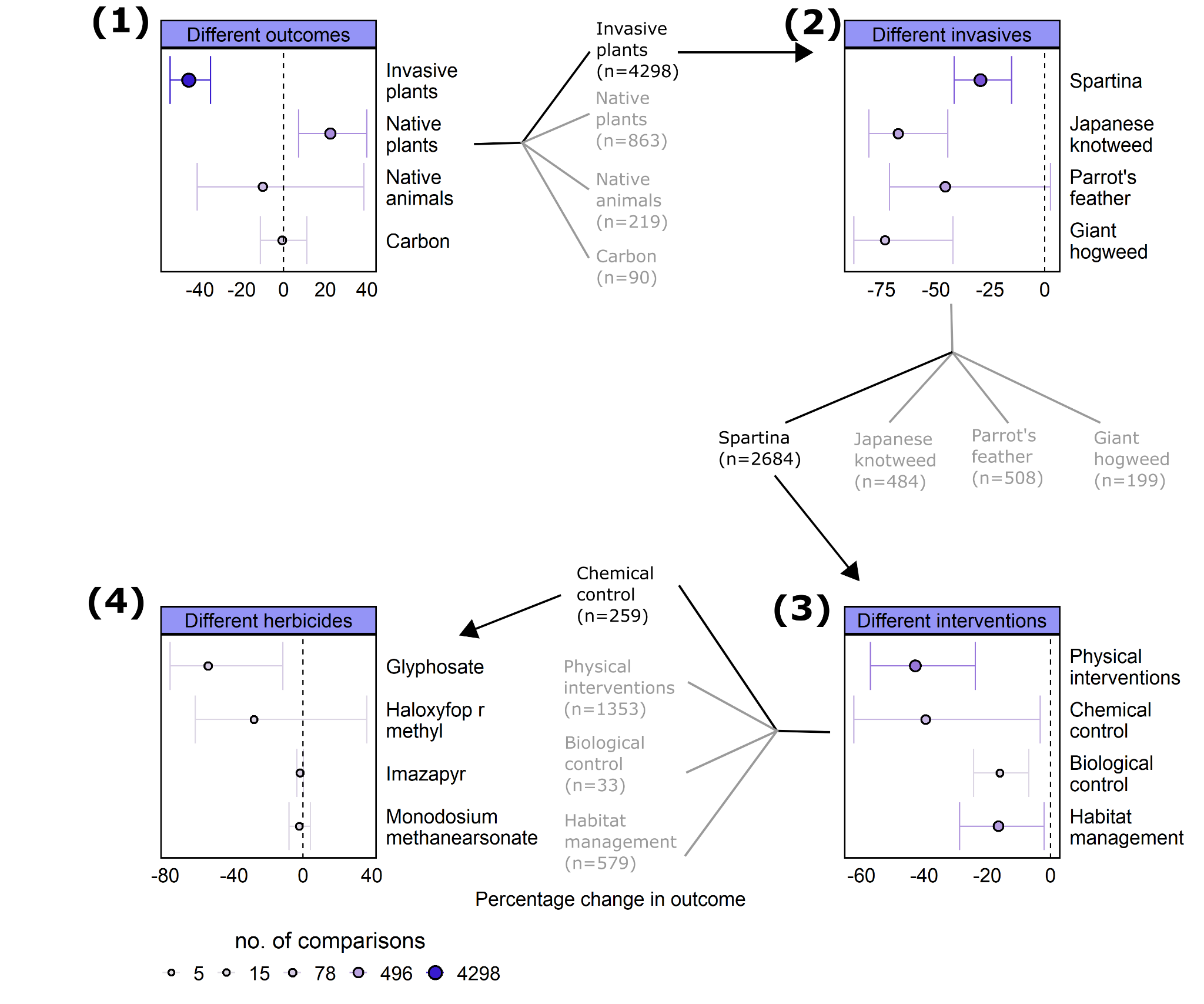
|  |  |  |  |
| --- | --- | --- | --- |
| **Contextual factor** | **Definition** | **Examples of these factors** | **Detailed examples** |
| Setting | The location where the intervention occurs and features relating to that location, excluding those relating to the target population. | * Differences in standard practice [(*2*)](https://paperpile.com/c/u4TPYy/qenN) * Cultural differences between countries/regions * Differences in biophysical conditions (e.g. climate) [(*1*)](https://paperpile.com/c/u4TPYy/gPod) * Differences in non-target ecological communities | Protected areas are more effective in reducing human pressure in economically developed regions than in less developed regions [(*18*)](https://paperpile.com/c/u4TPYy/Dorr) |
| Target population | The statistical population or sample on which the intervention aims to have an effect. | * Demographic differences [(*2*, *19*)](https://paperpile.com/c/u4TPYy/5y98+qenN) * Socioeconomic differences [(*2*, *19*)](https://paperpile.com/c/u4TPYy/5y98+qenN) * Differences in target species/taxonomic group [(*1*)](https://paperpile.com/c/u4TPYy/gPod) | Use of western, educated, industrialised, rich, and democratic subjects in studies of behaviour change interventions potentially limits the transferability of study results [(*19*)](https://paperpile.com/c/u4TPYy/5y98). |
| Intervention delivery |  | * Timing [(*2*)](https://paperpile.com/c/u4TPYy/qenN) * Intensity of intervention application [(*10*)](https://paperpile.com/c/u4TPYy/XcBn) * Frequency of application (i.e. is the intervention used once or multiple times) [(*10*)](https://paperpile.com/c/u4TPYy/XcBn) * Skill/experience of person applying intervention [(*6*)](https://paperpile.com/c/u4TPYy/Fp45) | School-wide anti-bullying interventions are generally effective, but their impacts vary widely [(*10*)](https://paperpile.com/c/u4TPYy/XcBn). The frequency and duration of programmes, and their ability to mobilise bystanders are seen as key determinants of effectiveness [(*10*)](https://paperpile.com/c/u4TPYy/XcBn). |
| Outcome measurement |  | * Metric measured [(*1*)](https://paperpile.com/c/u4TPYy/gPod) * Method used to measure outcome [(*20*)](https://paperpile.com/c/u4TPYy/m69B) * Timing of measurement after intervention * Spatial scale at which outcome is measured | Using only the abundance of target invasive species as an outcome measurement ignores the fact that native species often do not recover abundance/diversity following invasive management [(*21*, *22*)](https://paperpile.com/c/u4TPYy/blHy+QQ5o). |

## Providing an answer to ‘what works here?’/Identifying ‘what works here’

To address these barriers, we have developed a website and interactive tool called metadataset ([www.metadataset.com](http://www.metadataset.com)) to allow users to carry out ‘dynamic meta-analysis [(*23*)](https://paperpile.com/c/u4TPYy/ozus).’ The key difference between dynamic and traditional meta-analysis is that it allows users to filter and weight evidence depending on their interests to produce bespoke analyses (see Figure 1 for an overview). The database currently contains data from primary studies on agricultural interventions and invasive plant management and we hope to expand this further in the future. Figure 1 shows the path a user might take to better understand the impacts of different types of invasive plant management, with a particular focus on controlling an invasive plant species in the genus Spartina. This path is represented by the blue line in Figure 1a.

The user may first be interested in comparing the effect of invasive management on all outcomes to give a broad overview of impacts (first box Fig 1). Due to their interest in invasive plants, they may then wish to select only outcomes relating to invasive plants and compare across different invasive species ( second box Fig 1). After this the user may wish to compare the effects of different interventions on *Spartina* (third box Fig 1). Finally, having decided that they may be interested in using herbicide to control *Spartina* the user could then compare the impacts of different herbicides (fourth box Fig 1). This is just one route that could be taken in analysing outcomes, with users able to choose the route that interests them most (see caption of Figure 1 for more detail). Indeed, for an analysis of this sort a user, using our data on invasive species, could potentially decide to run 98 different meta-analyses.

In addition to allowing users to filter the outcomes and interventions that are of most interest to them, the system also allows users to either filter out studies that are not relevant to their context or recalibrate study weights so that those with higher relevance contribute more to meta-estimates [(*9*)](https://paperpile.com/c/u4TPYy/2gKK). 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 in a dynamic way using an online interactive tool means that users can customise the analysis to their specific context rather than relying on a static systematic review which may not address the context in which they are interested.



**Figure 1** - Typical workflow of an analysis on metadataset. Each box represents a separate meta-analysis and the lines between represent the filtering process of selecting different populations, species, and interventions. Point size and colour are indicative of the number of comparisons used for each estimate. Labels on 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 refers to the number of pairwise comparisons for each branch.

## Building an infrastructure for flexible evidence assessment

We think that dynamic meta-analysis presents an exciting new direction for evidence synthesis and decision-making - as demonstrated by our example above for the invasive plant species *Spartina.* This approach has the potential to help overcome barriers relating to relevance of evidence for almost any mission-driven discipline. We hypothesise that dynamic meta-analysis may be particularly useful for complex interventions delivered through multifaceted programs since the results of these studies are perceived to be very context specific. One such example is the implementation of anti-bullying campaigns in schools - the results of which appear to vary widely depending on their frequency and duration as well as their ability to mobilise bystanders to bullying [(Table1; *10*)](https://paperpile.com/c/u4TPYy/XcBn/?prefix=Table1%3B%20). Using dynamic meta-analysis would allow decision-makers to identify the potential effects of interventions for a context in a user-friendly, accessible format.

Although we think that dynamic meta-analysis is a powerful tool, we also acknowledge that there is much work to be done if its potential is to be fully realised. This work spans a variety of scales ranging from our research group and our collaborators, to policy-makers and funding bodies. In the short-term, our team and collaborators need to do two things: (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 our tool through user testing in order to make it easy to use. Outside of our immediate collaborators, we also encourage researchers in other mission-driven fields such as public health, development, and policing to explore systems for delivering similar tools. In the longer-term, in order to maximise the value of dynamic meta-analysis, and evidence synthesis in general, we think that there is a need for increased funding of roles in which people can act as intermediaries between research and policy, sometimes termed evidence bridges or knowledge-brokers [(*24*)](https://paperpile.com/c/u4TPYy/QYhC). These people could be trained how to use dynamic meta-analysis and related tools to produce robust analyses that can be used to inform practice and policy.

We think that dynamic meta-analysis is a powerful approach, but also recognise that it presents some problems. Of foremost concern is the inappropriate use of analyses to find statistical significance for a particular result, termed p-hacking. This could result from individuals with an agenda aiming to produce evidence that supports this agenda. However, we suspect that this is unlikely to present a major problem for our tool for three reasons. First, we think that many decision-makers are not seeking a significant result or to support a particular agenda but are simply trying to find an answer to a question. Second, our tool allows users to provide a link to the analysis that they carried out - allowing for other users to see exactly how the analysis was carried out. These other users could then repeat the analysis, modifying settings as they see fit, to examine whether reported analyses are robust. Third, if this approach becomes more common then the skills needed for analyses are likely to become more common within the practitioner community - leading to higher quality analyses.

We also acknowledge that dynamic meta-analysis is not necessarily useful in all circumstances. For example, we think that co-designed systematic reviews are useful for informing expensive, large-scale policy decisions where the costs of inaccurate estimates of intervention impacts are high [(*25*)](https://paperpile.com/c/u4TPYy/Rkqj). However, most decisions about interventions are probably not like these. Most decisions are relatively small, local decisions, made by a small team, where co-designed systematic reviews are not cost effective [(*25*)](https://paperpile.com/c/u4TPYy/Rkqj). In these cases in particular dynamic meta-analysis has the potential to provide rapid, quantitative answers, for problems where context is important and where decisions need to be made relatively rapidly, and costs need to be kept to a minimum [(*25*)](https://paperpile.com/c/u4TPYy/Rkqj).

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 co-produced systematic review. To achieve this we call on policymakers and funders to prioritise the training and hiring of people who act as intermediaries between researchers and policymakers. If well implemented, this step could aid society in overcoming many of the

pressing global-scale challenges that require the application of interventions.

Although these challenges occur at the global scale, their solutions must be delivered at national or subnational scales. We believe that synthesis which provides locally relevant evidence is one of the many steps needed to ensure this.

## References

1. [A. P. Christie, T. Amano, P. A. Martin, S. O. Petrovan, G. E. Shackelford, B. I. Simmons, R. K. Smith, D. R. Williams, C. F. R. Wordley, W. J. Sutherland, Poor availability of context-specific evidence hampers decision-making in conservation. *Biol. Conserv.* **248**, 108666 (2020).](http://paperpile.com/b/u4TPYy/gPod)

2. [P. M. Rothwell, External validity of randomised controlled trials: “To whom do the results of this trial apply?” *Lancet*. **365**, 82–93 (2005).](http://paperpile.com/b/u4TPYy/qenN)

3. [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).](http://paperpile.com/b/u4TPYy/m8Uq)

4. [E. M. Beller, J. K.-H. Chen, U. L.-H. Wang, P. P. Glasziou, Are systematic reviews up-to-date at the time of publication? *Syst. Rev.* **2**, 36 (2013).](http://paperpile.com/b/u4TPYy/fnE6)

5. [W. J. Sutherland, A. S. Pullin, P. M. Dolman, T. M. Knight, The need for evidence-based conservation. *Trends Ecol. Evol.* **19**, 305–308 (2004).](http://paperpile.com/b/u4TPYy/8evE)

6. [S. A. Avellar, J. Thomas, R. Kleinman, E. Sama-Miller, S. E. Woodruff, R. Coughlin, T. R. Westbrook, External Validity: The Next Step for Systematic Reviews? *Eval. Rev.* **41**, 283–325 (2017).](http://paperpile.com/b/u4TPYy/Fp45)

7. [S. Wang, J. R. Moss, J. E. Hiller, Applicability and transferability of interventions in evidence-based public health. *Health Promot. Int.* **21**, 76–83 (2006).](http://paperpile.com/b/u4TPYy/m2Nx)

8. [J. A. Catford, J. R. U. Wilson, P. Pyšek, P. E. Hulme, R. P. Duncan, Addressing context dependence in ecology. *Trends Ecol. Evol.* **0** (2021), doi:](http://paperpile.com/b/u4TPYy/Vn9J)[10.1016/j.tree.2021.09.007](http://dx.doi.org/10.1016/j.tree.2021.09.007)[.](http://paperpile.com/b/u4TPYy/Vn9J)

9. [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).](http://paperpile.com/b/u4TPYy/2gKK)

10. [E. Menesini, C. Salmivalli, Bullying in schools: the state of knowledge and effective interventions. *Psychol. Health Med.* **22**, 240–253 (2017).](http://paperpile.com/b/u4TPYy/XcBn)

11. [N. R. Haddaway, C. Kohl, N. Rebelo da Silva, J. Schiemann, A. Spök, R. Stewart, J. B. Sweet, R. Wilhelm, A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environmental Evidence*. **6**, 11 (2017).](http://paperpile.com/b/u4TPYy/TAfK)

12. [C. N. Cook, H. P. Possingham, R. A. Fuller, Contribution of systematic reviews to management decisions. *Conserv. Biol.* **27**, 902–915 (2013).](http://paperpile.com/b/u4TPYy/eSAa)

13. [W. J. Sutherland, G. Shackelford, D. C. Rose, Collaborating with communities: co-production or co-assessment? *Oryx*. **51**, 569–570 (2017).](http://paperpile.com/b/u4TPYy/dfBH)

14. [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).](http://paperpile.com/b/u4TPYy/tzkX)

15. [J. Koricheva, E. Kulinskaya, Temporal Instability of Evidence Base: A Threat to Policy Making? *Trends Ecol. Evol.* **34**, 895–902 (2019).](http://paperpile.com/b/u4TPYy/06kd)

16. [K. G. Shojania, M. Sampson, M. T. Ansari, J. Ji, S. Doucette, D. Moher, How quickly do systematic reviews go out of date? A survival analysis. *Ann. Intern. Med.* **147**, 224–233 (2007).](http://paperpile.com/b/u4TPYy/qrIz)

17. [J. Elliott, R. Lawrence, J. C. Minx, O. T. Oladapo, P. Ravaud, B. Tendal Jeppesen, J. Thomas, T. Turner, P. O. Vandvik, J. M. Grimshaw, Decision makers need constantly updated evidence synthesis. *Nature*. **600**, 383–385 (2021).](http://paperpile.com/b/u4TPYy/6o94)

18. [J. Geldmann, A. Manica, N. D. Burgess, L. Coad, A. Balmford, A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23209–23215 (2019).](http://paperpile.com/b/u4TPYy/Dorr)

19. [J. Henrich, S. J. Heine, A. Norenzayan, *Behav. Brain Sci.*, in press, doi:](http://paperpile.com/b/u4TPYy/5y98)[10.1017/S0140525X0999152X](http://dx.doi.org/10.1017/S0140525X0999152X)[.](http://paperpile.com/b/u4TPYy/5y98)

20. [A. P. Christie, T. Amano, P. A. Martin, G. E. Shackelford, B. I. Simmons, W. J. Sutherland, Simple study designs in ecology produce inaccurate estimates of biodiversity responses.](http://paperpile.com/b/u4TPYy/m69B) [J. Appl. Ecol](http://paperpile.com/b/u4TPYy/QQ5o)[. **56**, 2742–2754 (2019).](http://paperpile.com/b/u4TPYy/m69B)

21. [D. E. Pearson, Y. K. Ortega, J. B. Runyon, J. L. Butler, Secondary invasion: The bane of weed management. *Biol. Conserv.* **197**, 8–17 (2016).](http://paperpile.com/b/u4TPYy/blHy)

22. [K. M. Kettenring, C. R. Adams, Lessons learned from invasive plant control experiments: a systematic review and meta-analysis. *J. Appl. Ecol.* **48**, 970–979 (2011).](http://paperpile.com/b/u4TPYy/QQ5o)

23. [G. E. Shackelford, P. A. Martin, A. S. C. Hood, A. P. Christie, E. Kulinskaya, W. J. Sutherland, Dynamic meta-analysis: a method of using global evidence for local decision making. *BMC Biol.* **19**, 33 (2021).](http://paperpile.com/b/u4TPYy/ozus)

24. [A. N. Kadykalo, R. T. Buxton, P. Morrison, C. M. Anderson, H. Bickerton, C. M. Francis, A. C. Smith, L. Fahrig, Bridging research and practice in conservation. *Conserv. Biol.* (2021), doi:](http://paperpile.com/b/u4TPYy/QYhC)[10.1111/cobi.13732](http://dx.doi.org/10.1111/cobi.13732)[.](http://paperpile.com/b/u4TPYy/QYhC)

25. [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).](http://paperpile.com/b/u4TPYy/Rkqj)