Maximizing Legacy and Impact of Primary Research: A Call for Better Reporting of Results

Neal R. Haddaway

Received: 8 March 2014/Revised: 22 April 2014/Accepted: 2 May 2014/Published online: 29 May 2014

Abstract Much of the scientific literature in existence today is based on model systems and case studies, which help to split research into manageable blocks. The impact of this research can be greatly increased in meta-analyses that combine individual studies published over time to identify patterns across studies; patterns that may go undetected by smaller studies and that may not be the main subject of investigation. However, many potentially useful studies fail to provide sufficient data (typically means, true sample sizes, and measures of variability) to permit meta-analysis. Authors of primary research studies should provide these summary statistics as a minimum, and editors should require them to do so. By putting policies in place that require these summary statistics to be included, or even those that require raw data, editors and authors can maximize the legacy and impact of the research they publish beyond that of their initial target audience.

Keywords Publishing · Meta-analysis · Systematic review · Data reporting · Quantitative synthesis · Missing data

INTRODUCTION

Some 8323 scientific journals were listed in Journal Citation Reports in 2013, with tens of thousands more journals unlisted. The vast majority of these journals have been given impact factors in the lower end of the spectrum, giving a classic Poisson distribution with a median of approximately 0.5 and 1.0 (Thomson Reuters 2013). Thus, the majority of journals are typically more *applied* (i.e., focused on more practical subjects) than their counterparts at the far end of the spectrum, publishing research that targets specific audiences. Much of the research in these

publications uses model species and habitats or case studies to simplify more complex systems (e.g., Rantalainen et al. 2008). While these studies are often quite specific, they can inform wider analyses if, for example, used in a meta-analysis and systematic reviews (SRs) (Pullin and Knight 2001).

Meta-analyses are statistical methods that combine like studies to create a single study of far greater effective sample size than any of its constituent parts (Glass 1976). These analyses are used where individual studies disagree, or where individual studies are thought to be of insufficient power to identify significant effects. Meta-analyses are powerful tools to increase the value and impact of research. Meta-analysis has been widely used in recent decades in medicine to identify significant patterns in the evidence that may go undetected in individual studies (O'Rourke 2007). Analyzed together, the evidence provided by individual studies is more powerful than the sum of its individual analyses. Furthermore, meta-analyses allow us to examine the effect of modifying factors that may not have been considered in the original research. For example, while individual studies on the effect of drainage on greenhouse gas emissions from peatlands may each have been undertaken in sites with a specific mean annual rainfall and temperature, when studies are combined in a meta-analysis the effect of meteorology on the relationship between land management and emissions can be examined (also referred to as *sources of heterogeneity* and *effect modifiers*) (Haddaway et al. 2014).

Meta-analyses in the health sciences have identified significant positive effects of potentially life-saving therapies where individual studies have failed to find an effect. One example of the potential influence of meta-analyses on policy is demonstrated by the review of the use of streptokinase in the treatment of myocardial infarction

(commonly known as a heart attack). A meta-analysis that arranged and analyzed studies cumulatively through time over a 30-year period identified a statistically significant reduction in mortality resulting from the therapeutic use of streptokinase following myocardial infarction. This significant effect was clear in the cumulative meta-analysis after only 14 years of research, but streptokinase was not widely recommended until more than a decade later when two large-scale trials (mega trials) identified a significant effect (Lau et al. 1992). This striking example demonstrates the potentially preventable loss of life that results from missing patterns in the evidence identified through pooling studies.

Meta-analyses in medicine, and more recently in environmental management and conservation (Gurevitch et al. 1992), have been developed even further by the establishment of systematic review methodology (Pullin and Stewart 2006; Higgins and Green 2011). Systematic reviews aim to identify all available evidence for a specific question using a detailed, pre-defined methodology. This methodology aims to minimize various biases, such as publication bias and selection bias that may affect traditional reviews.

The power and utility of meta-analyses, however, is reduced significantly when primary research does not report sufficient data to allow full quantitative analyses. These studies with missing data must be excluded from the analysis despite being relevant and providing some informative results. Broadly speaking, primary research articles should report three key measures to facilitate their inclusion in a meta-analysis: mean *effect size*, *sample size*, and *measure of variability* (typically standard deviation, standard error, or confidence intervals). *Effect sizes* are summary statistics that estimate the magnitude of effect of a specific intervention (e.g., application of a pesticide) or exposure (e.g., soil water content). One form of effect sizes where studies report their results in the same units would be the raw mean difference, the control sample mean subtracted from the intervention sample mean, which represents the direct additional effect of the intervention in meaningful units. Other examples of effect sizes include correlation coefficients, risk ratios, and specific effect sizes designed for meta-analysis such as *Hedges g*. Different effect size types are suitable for different outcome measures and data types (Borenstein et al. 2011). *Measures of variability* indicate the uncertainty of effect size estimates and are used in meta-analyses to weight studies according to the variability in the data around the sample means, in order to give more weight to more precise studies. A range of possible variability measures can be used in meta-analyses as these are interchangeable. *Sample sizes* relate to the true sample size of the study and should not include pseudoreplicates. True replicates are those samples that are

measured at the same level as that at which the intervention is experienced: if treatments are delivered at the field level, then replicates are fields and NOT plots within fields.

Where quantitative data for the key measures described above are not presented in the text or tables of relevant studies, this information can often be extracted from figures of summary metrics or raw data (e.g., Tummers 2006). In some cases other data can be included in a meta-analysis. For example, meta-analysis can be performed on p values (Fisher 1932), but such analyses do not consider the magnitude or the direction of effect, and cannot investigate sources of heterogeneity, so should be restricted to use when other options for meta-analysis are exhausted (Jones 1995).

Where data on key measures are missing from some studies, for example variability measures, it may be appropriate to *impute* these values (see Harris et al. 2009). Imputing involves replacing a missing value with an appropriate substitute. It enables the inclusion of studies that would otherwise be excluded due to the lack of reported data, and thus mitigating the potential impact this would have on the power and bias of the pooled effect (Wiebe et al. 2006; Burgess et al. 2013). This may be generated, for example for variability measures, in one of a number of ways: it may be based on an understanding of the population being studied; from a mean variability identified from other studies included in the meta-analysis; or from the largest variance reported in other included studies in order to be more conservative. One final option is to perform multiple imputation using several methods and substituting some form of average where the data are missing. Imputing is often appropriate in medicine, where meta-analyses involve large numbers of studies and imputing of a small number of studies' variability is less influential on the overall analysis. Meta-analysis in environmental sciences, however, rarely involves large samples sizes, and large proportions of the evidence base may be missing data. Three recent systematic reviews highlight this problem. A recent systematic review of the impact of terrestrial protected areas on human well-being identified 281 outcome measures across 49 studies, but 82 percent of these studies reported measures with no variability (Pullin et al. 2013). Another review of the impact of land management on lowland peatland carbon greenhouse gas flux identified 33 of 111 studies that lacked measures of variability, precluding their inclusion in meta-analysis (Haddaway et al. 2014). In a systematic review of the impact of reindeer grazing on arctic and alpine vegetation, currently underway, 30 percent of the included articles were unable to be included in meta-analysis due to a lack of either variability (10 of 53 studies) or true sample size (6 of 53 studies) (Bernes et al. 2013). Despite the availability and use of imputing methods in the health care discipline, these

are not always feasible in the environment setting, and therefore there are even more imperative primary studies to report the variance data. Studied human populations are typically far less variable and more predictable than the range of studied populations included in meta-analyses in the environmental sciences (Haddaway et al. 2014). As a result, imputing in environmental sciences meta-analyses is rarely likely to be appropriate.

One other solution to the problem of missing values is to contact the authors of relevant studies with a request for supplemental data. Such requests are more successful with recently published manuscripts (Vines et al. 2014), where email addresses are supplied and are still functional. Email requests for data in meta-analyses typically have low success rates (e.g., Gibson et al. 2006), with only a small minority of contacted authors responding with usable data. The process should be encouraged where resources allow, since the increase in usable data is often valuable. For older research, however, such contact is often not feasible. This latter point raises concerns about a possible bias resulting from the presence of more usable data in meta-analyses from more recent research. Such bias should not be ignored, but little can be done to account for it.

In systematic reviews, study results can often still be synthesized narratively in the form of textual descriptions, tabulation, and the production of figures despite being lost from meta-analyses. However, such narrative syntheses are not as powerful as meta-analyses, which should be the main aim of a quantitative (aggregative) systematic review. Furthermore, if some studies are missing effect sizes and statistical results, little use can be made of their results.

Those with experience in meta-analysis and systematic review understand the value of well-reported summary data in primary research articles, and failing this, the provision of raw data. To ensure the legacy of primary research and maximize its value, however, it should be the priority of journal editors and manuscript authors to ensure that all primary researches report quantitative data either in summary or raw form. Summary data should be provided with measures of variability to ensure that it can be included in meta-analyses. Maximizing the use of existing evidence in meta-analyses may also potentially conserve resources that would otherwise be used for additional primary research, where answers already exist in the literature. This policy follows the recommendations made in the CONSORT Statement (BMJ 2010a, b) in medicine that call for better reporting of clinical trials.

Some journals have recently begun to demand the publication of raw data alongside manuscripts. The Public Library of Sciences (PLOS), for example, amended their data policy in December 2013 to state that “PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without

restriction.” Such a policy is a bold move in a competitive publishing market; the majority of other journals, particularly those that are not fully turning to Open Access, may find such a move difficult to implement. Summary data for treatment and control groups in the form of means, sample sizes, and variability measures are a far simpler, yet just as effective, requisite that will maximize the legacy and usability of primary research.

Systematic reviews and meta-analyses include research from a range of time periods, not solely more recent publications. As the publishing world advances and reporting of raw and summary data improve, the historic research that lacks sufficient data to permit meta-analysis could be made useful with the establishment of a universal database for the deposition of raw and summary data. Such a database could mirror the advances in independent post-publication peer review such as www.PubPeer.com. This project would require a significant effort to establish, maintain, and advertise.

Acknowledgments I thank Claes Bernes, Ruth Lewis, and two anonymous reviewers for their comments on a draft version of the manuscript.

REFERENCES

- Bernes, C., K.A. Bråthen, B.C. Forbes, A. Hofgaard, J. Moen, and J.D. Speed. 2013. What are the impacts of reindeer/caribou (*Rangifer tarandus* L.) on arctic and alpine vegetation? A systematic review protocol. *Environmental Evidence* 2: 6. doi:10.1186/2047-2382-2-6.
- BMJ. 2010a. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ* 340: 332. doi:10.1136/bmj.c332.
- BMJ. 2010b. The new CONSORT statement. *BMJ* 340: c1432. doi:10.1136/bmj.c1432.
- Borenstein, M., L.V. Hedges, J.P. Higgins, and H.R. Rothstein. 2011. *Introduction to meta-analysis*. Chichester: Wiley. ISBN 978-0-470-05724-7.
- Burgess, S., I.R. White, M. Resch-Rigon, and A.M. Wood. 2013. Combining multiple imputation and meta-analysis with individual participant data. *Statistics in Medicine* 32(26): 4499–4514. doi:10.1002/sim.5844.
- Fisher, R.A. 1932. *Statistical methods for research workers*. London: Oliver & Boyd. ISBN 0-05-002170-2.
- Gibson, C.A., B.W. Bailey, M.J. Carper, J.D. LeCheminant, E.P. Kirk, G. Huang, K. Drowatzky DuBose, and J.E. Donnelly. 2006. Author contacts for retrieval of data for a meta-analysis on exercise and diet restriction. *International Journal of Technology Assessment in Health Care* 22(02): 267–270. doi:10.1017/S0266462306051105.
- Glass, G.V. 1976. Primary, secondary, and meta-analysis of research. *Educational Researcher* 5(10): 3–8.
- Gurevitch, J., L.L. Morrow, A. Wallace, and J.S. Walsh. 1992. A meta-analysis of competition in field experiments. *The American Naturalist* 140(4): 539–572.
- Haddaway, N.R., A. Burden, C.D. Evans, J.R. Healey, D.L. Jones, S.E. Dalrymple, and A.S. Pullin. 2014. Evaluating effects of land management on greenhouse gas fluxes and carbon balances in

- boreo-temperate lowland peatland systems. *Environmental Evidence* 3: 5. doi:[10.1186/2047-2382-3-5](https://doi.org/10.1186/2047-2382-3-5).
- Harris, C., L.V. Hedges, and J.C. Valentine. 2009. *Handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation. ISBN 978-0-87154-163-5.
- Higgins, J.P.T., and S. Green. 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. Updated March 2011. Retrieved March 1, 2014, from www.cochrane-handbook.org.
- Jones, D. 1995. Meta-analysis: weighing the evidence. *Statistics in Medicine* 14: 137–149. doi:[10.1002/sim.4780140206](https://doi.org/10.1002/sim.4780140206).
- Lau, J., E.M. Antman, J. Jimenez-Silva, B. Kupelnick, F. Mosteller, and T.C. Chalmers. 1992. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine* 327(4): 248–254. doi:[10.1056/NEJM199207233270406](https://doi.org/10.1056/NEJM199207233270406).
- O'Rourke, K. 2007. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine* 100(12): 579–582.
- Pullin, A.S., and T.M. Knight. 2001. Effectiveness in conservation practice: Pointers from medicine and public health. *Conservation Biology* 15(1): 50–54. doi:[10.1111/j.1523-1739.2001.99499.x](https://doi.org/10.1111/j.1523-1739.2001.99499.x).
- Pullin, A.S., and G.B. Stewart. 2006. Guidelines for systematic review in conservation and environmental management. *Conservation Biology* 20(6): 1647–1656. doi:[10.1111/j.1523-1739.2006.00485.x](https://doi.org/10.1111/j.1523-1739.2006.00485.x).
- Pullin, A.S., M. Bangpan, S. Dalrymple, K. Dickson, N.R. Haddaway, J.R. Healey, H. Hauari, N. Hockley, et al. 2013. Human well-being impacts of terrestrial protected areas. *Environmental Evidence* 2: 19. doi:[10.1186/2047-2382-2-19](https://doi.org/10.1186/2047-2382-2-19).
- Rantalainen, M.L., J. Haimi, H. Fritze, T. Pennanen, and H. Setälä. 2008. Soil decomposer community as a model system in studying the effects of habitat fragmentation and habitat corridors. *Soil Biology & Biochemistry* 40(4): 853–863. doi:[10.1016/j.soilbio.2007.11.008](https://doi.org/10.1016/j.soilbio.2007.11.008).
- Thomson Reuters. 2013. 2012 Journal Citation Reports® Science Edition. Accessed October, 2013.
- Tummers, B. 2006. DataThief III. Retrieved March 1, 2014, from <http://datathief.org>.
- Vines, T.H., A.Y.K. Albert, R.L. Andrew, F. Débarre, D.G. Bock, M.T. Franklin, K.J. Gilbert, J.-S. Moore, et al. 2014. The availability of research data declines rapidly with article age. *Current Biology* 24: 94–97. doi:[10.1016/j.cub.2013.11.014](https://doi.org/10.1016/j.cub.2013.11.014).
- Wiebe, N., B. Vandermeer, R.W. Platt, T.P. Klassen, D. Moher, and N.J. Barrowman. 2006. A systematic review identifies a lack of standardization in methods for handling missing variance data. *Journal of Clinical Epidemiology* 59: 342–353. doi:[10.1016/j.jclinepi.2005.08.017](https://doi.org/10.1016/j.jclinepi.2005.08.017).

AUTHOR BIOGRAPHY

Neal R. Haddaway (✉) is a Postdoctoral Research Officer at the Centre for Evidence-Based Conservation, School of Environment, Natural Resources and Geography, Bangor University, Bangor, LL57 2UW, UK.
 Address: Centre for Evidence-Based Conservation, School of Environment, Natural Resources and Geography, Bangor University, Bangor LL57 2UW, UK.
 e-mail: neal_haddaway@hotmail.com