FEATURE

Clarity over complexity: Statistical reporting that resonates

Stephen R. Midway* | Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, Louisiania, USA
Daniel J. Daugherty | Texas Parks and Wildlife Department, Heart of the Hills Fisheries Science Center, Mountain Home, Texas, USA
*Corresponding author: Stephen R. Midway. Email: smidway@lsu.edu.

Image conceived by Stephen R. Midway and rendered using AI.

https://doi.org/10.1093/fshmag/vuaf043

ABSTRACT

Effective statistical reporting is essential for the credibility, reproducibility, and advancement of scientific research. Despite existing guidelines, many researchers—especially early career researchers—feel underprepared to handle increasingly complex statistical methods. This gap is notable in fields like fisheries science, where advanced quantitative tools are often required for critical decision making. Unclear statistical reporting—whether due to selective reporting, insufficient sample sizes, or inadequate model understanding—undermines the integrity and use of research findings. This article reviews common issues in statistical reporting and offers practical recommendations for clarity, transparency, and objectivity. We organize topics into general concerns, common issues, and small errors to provide a relative magnitude to the nature of the issue. Topics include selecting appropriate models, avoiding significance bias, addressing challenges with small sample sizes, and ensuring reproducibility, among others. We advocate for clear documentation of methods, effective use of visuals, and incorporation of supplemental materials, such as data and code, to facilitate understanding and replication. Rather than prescribing absolutes, we encourage researchers to embrace practices that prioritize clarity and reader comprehension. By adopting these recommended practices, scientists can ensure their work is not only accessible but also capable of advancing knowledge and informing policy—often the primary goal of scientific reporting.

INTRODUCTION

Does science really need another paper on reporting statistics? Guidance already exists, often buried in discipline-specific journals (Davis & Kay, 2023). Despite available advice, Barraquand et al. (2014) surveyed over 900 early career ecologists and found a clear self-perceived lack of quantitative training: 75% were unsatisfied with their understanding of mathematical models and felt that in their degree programs, the level of mathematics was inadequate; 90% wanted more mathematics classes for ecologists; and 95% wanted more statistical training. Research on and management of natural resources often involves fairly advanced statistics; examples include capture–recapture models to estimate populations (Williams et al., 2002), multivariate statistics to reduce complex data sets to a manageable number of variables (Barraquand et al., 2014; Legendre & Legendre, 2012), and increasingly popular non-frequentist Bayesian approaches (Dorazio, 2016). To meet these needs, an ever-increasing bounty of freely available software and analytical packages has made quantitative approaches more accessible to all, such as the open-source statistical programming language R (Barraquand et al., 2014; R Core Team, 2025). Thus, underprepared scientists, combined with increasingly complex statistical methods (Low-Décarie et al., 2014), suggest that reporting standards need continual attention. This is particularly the case for fisheries journals, as guidance is generally limited to symposium proceedings (Hunter, 1990) and textbooks (Guy & Brown, 2007) with limited accessibility.

Statistics is the language we use to tell our scientific stories. The values we report help create the strength of our arguments when presenting and interpreting our findings. The clearer that our statistical methods and results are described, the greater the likelihood that a reader will understand and confidently take further action, such as citing the work, building upon it, or implementing policy (Davis & Kay, 2023). The recent development and proliferation of statistical methods has undoubtedly opened doors for scientists and helped usher in a boom in published articles over the past few decades. However, statistical complexity has accompanied the increase in statistical usage (Low-Décarie et al., 2014), often with deleterious effects. In one direct measure of negative outcomes from statistical complexity, Fawcett and Higginson (2012) found that "28% fewer citations overall for each additional equation per page." Unfortunately, most scientific communication issues are not simply solved by removing a few equations.

Many of the issues with statistical usage are likely the subtle and embedded habits of poor reporting or innocent ignorance regarding what readers want to know. Unfortunately, the opposite is also true—statistical opacity can result in even the most important of scientific issues being ignored because they are not understood. The field of fisheries science can be a quantitative field. Many natural resource professionals have advanced statistical training because they study populations and environments upon which major decisions rely on the results of statistical models. Fisheries science finds itself at the center of the complexity problem—many of us are more comfortable around numbers than our non-fish colleagues, yet that means the bar is raised for us to clearly explain what we are doing. With this elevated expectation in mind, we review some of the common problems with statistical reporting and ultimately arrive at some best practices and recommendations. Throughout this article, we provide recommendations over absolutes, because the reality of statistics is that there is judgment associated with their use. Some journals have called for an end to *P*-values (Woolston, 2015), while other journals have simply provided reporting guidance (Michel et al., 2020; Wootton & Craig, 2011). Yet, many journals remain agnostic in their guidance. We are not so dogmatic as to tell anyone exactly what they must do—other than to report clearly. We prefer to highlight effective options that scientists have and then trust their judgment in the correct application. The ultimate motivation to clarify your statistical reporting is so that your work is understandable, discoverable, and amplified. Who wouldn't want that?

GENERAL PROBLEMS WITH STATISTICAL REPORTING

Despite its importance, statistical reporting often suffers from shortcomings that undermine the integrity and reliability of the findings. In later sections, we will discuss specific technical problems, but here, we overview some of the general issues. In some cases, these general problems set the stage for technical errors to arise. In many cases, there is no quick fix for a general problem, but being aware of them can certainly help (for a summary and checklist of issues discussed throughout this document, see Figure 1). Generally speaking, we recommend that when designing a study, you explicitly consider any analytical methods that could be used. Armed with a solid question, anticipate what model you might need and think about

Statistical Reporting Checklist

	GENERAL CONCERNS
☐	Use the most **appropriate model** for your question and data. Be able to justify the model.
☐	Report **non-significant results**, too, because they are information.
☐	Understand your **sample size** (e.g., do a power analysis).
☐	Use a **model you understand**—if you don't understand it, don't use it.
☐	Make sure a reader can **reproduce your analyses** if they want to.
☐	Do not assume your reader has your level of technical expertise: **explain everything** clearly.
	COMMON ISSUES
☐	**Predictor variables**: consider what they represent, how to interpret them, that they are uncorrelated with others, and whether they have been transformed or standardized.
☐	**Sample size**: collect more data, clarify sample sizes, and consider non-parametric techniques.
☐	**Significance testing**: numerous issues with *P*-values can be addressed with effect sizes, reporting language, or even using other estimation (e.g., Bayesian).
☐	**Model selection**: Only select from models that you would actually want; consider cross-validation.
	SMALL ERRORS
☐	Provide **clear questions or hypotheses** for your reader.
☐	Always use the **simplest methods** possible.
☐	**Report all assumptions** for your reader to know.
☐	Equations almost always are more clear than text describing a model; **use equations**.
☐	**Use statistical notation** and not computer code (for equations).
☐	Understand and **justify any multiple comparisons test** you used.
☐	**Uncertainty**: Know it, report it, and don't hide it.
☐	**Consistently report significant digits**—remember, values cannot be more precise than the instruments.
☐	If your work is complex, **use subheadings to organize** and help your reader.
☐	Put important (but not that important) stuff in **supplements**. Don't delete it.
☐	**Include data and code** to enable reproducibility.
☐	Cognitively **effective visuals** will help promote understanding of your work.

Figure 1. A summary and checklist of statistical reporting issues for authors to consider.

its assumptions and requirements. Many statistical issues are unforced errors because the model is approached post hoc, after the data are fixed, in-hand, or otherwise cannot be changed.

The right model for the right question

The first general problem that exists can be the hardest one to address, and that is which statistical method (test, model, etc.) is the most appropriate for a given data set and question. The problem also presupposes that you have thoughtfully developed and refined a study question (or hypothesis), which we cannot overstate the importance of. The model you select can often be subjective because, for some data sets and questions, there may be multiple appropriate approaches (see Gould et al., 2025 for an interesting experiment on this topic). This fact should be remembered by both authors and reviewers during peer review when it is common for a reviewer to recommend what they would have done instead of a recommendation based on whether the method in question was objectively appropriate. The issue of using an appropriate statistical test is not something we can easily diagnose here, other than recommending that authors continually revisit the study questions because, ultimately, any model is just a routine to provide you an answer to your question. The model will always provide an answer—the key is making sure the model is answering the question you asked. Understanding which model you need often comes from years of studying and working with statistical models. However,

some studies have attempted to tackle this overarching challenge; for example, Tredennick et al. (2021) provides a useful reference for both understanding and directly comparing a larger number of models (that users may otherwise learn individually and, therefore, without the context of other options).

Selective reporting

A second general issue can be the selective reporting of statistical tests and results, which can lead to publication bias, where studies with significant findings are more likely to be published than those with null or nonsignificant results. Significance bias can distort the overall body of scientific literature, skewing perceptions of the prevalence and magnitude of effects (Ioannidis, 2005; Kimmel et al., 2023). Because the intent of this article is to provide usable advice for individual author teams, we will not pursue the topic of nonsignificant results further, other than to mention that it is a statistically imposed bias that shapes what we see in the published literature. Yet results of no difference or nonsignificance may still hold biological or ecological importance and sharing them may prevent other scientists from repeating nonsignificant experiments and studies. Nonsignificant results may also be of value to meta-analyses or other holistic evaluations of a topic.

Small sample size

Inadequate sample sizes and lack of statistical power are pervasive issues in statistical reporting, reducing the reliability and generalizability of results. Often, small sample sizes are a reality that cannot be overcome, particularly after the project has been completed. One way to address statistical power is to conduct a power analysis before or during project development in order to get an estimate of the recommended sample size (Thomas & Juanes, 1996). Unfortunately, power analyses are not often done (Jennions, 2003). Every scientist is limited in some way, whether in the purchase of equipment, facility space, days in the field, or numerous other logistics that shape how we can design a study. Unfortunately, some scientists may need to ask themselves whether their work should undergo peer review if the sample size is simply not enough to have confidence in the results. The risk of small sample sizes is twofold. Small sample sizes not only make it challenging to detect genuine effects but also increase the risk of spurious findings and false conclusions (Button et al., 2013).

Model understanding

Most scientists have at their fingertips an expansive and impressive set of statistical tools. Rarely (if ever) have scientists felt so well equipped to play the role of statistician and analyze their data. However, as the complexity and precision of statistical tools have increased, the reality is that most fishery scientists do not have full command of all of the statistical options they have access to. For example, a popular statistical software in fisheries science (and beyond) is program R (R Core Team, 2025), and while R has advanced the quantitative abilities of many scientists, its canned functions can be a black box to others. In other words, simply because you are able to run a model does not mean you should if you do not understand what the model does, what the model assumptions are, and how to interpret the model results. The quantitatively enriched environment we find ourselves in can, however, be looked at as an opportunity for scientists to take the time they need to learn about a model that may be of interest to them before they try to publish it. Warren Buffet famously said of investing, "Never invest in a business you cannot understand." We adopt that sentiment for statistics and strongly recommend, "Never use a model you cannot understand." Fortunately, a variety of good quantitative texts for fisheries scientists and ecologists exist (Guy & Brown, 2007; Haddon, 2020; Ogle, 2015).

Reproducibility

Although several individual mistakes are discussed in later sections, a common aggregate effect of many reporting issues means that a study cannot be replicated. The ability to replicate a piece of science is needed for the ideas within to be evaluated, developed, and contribute to the collective knowledge. Rather than discussing the merits of replication, we encourage authors to consider reproducibility, which is often defined as the reader's ability to reproduce the analyses that were done in a scientific study (Peng & Hicks, 2021). For many kinds of research in fisheries, replication is not possible, and reproduction is the only option. Peng (2009) encourages the idea that reproducible research is a minimum standard for replication. So, while replication may be a measure of how much an independent scientist could repeat the study that you did (in its entirety), reproducibility is an analytically focused version where a reader may even have the data and code such that they could independently run the same analysis. Not every author team will have the ability to share their data or code, yet we think that the spirit of reproducibility remains a good idea to keep in mind as you are documenting your methods and results. Many journals have also adopted this philosophy, as data and code requirements for manuscript submissions are increasingly common. Frankly, in today's world of open science (e.g., publicly available data and open access journal articles), should we trust results that we don't think we could actually reproduce?

The curse of knowledge

Pinker (2014) is behind the idea of the "curse of knowledge," which he describes as "…the failure to understand that other people don't know what we know." Although this idea was developed in the context of writing, a statistical curse of knowledge happens anytime we are so familiar with our own statistics that we forget to report details because we assume the reader knows as much as we do. Many statistical methods beyond basic regression fall vulnerable to the curse of knowledge. If you pick up the latest issue of your favorite fisheries journal, do you know all there is to know about Bayesian statistics, eigenvectors, and autoregressive moving averages? We are not advocating that every method comes with a primer, but rather to keep in mind that the more technical you report, the more likely you are to lose readers. Do not assume the reader knows what you know about your data and the decisions you made to analyze it. Please tell us how you did what you did.

COMMON ISSUES IN REPORTING STATISTICAL ANALYSES

The previous section was designed to describe some of the more systemic issues in statistical usage and help understand the

conditions that often lead to technical problems in statistical reporting. In this section, we identify several smaller and technical statistical reporting issues and offer some ways to address them. The general problems are those that you want to be aware of and think about in the long term; these common issues are ones you can directly revise with often immediate benefit to your study and for your reader.

Predictor variables

Other than model type, the predictors (e.g., independent variables) we use in our models generally occupy a lot of our consideration. For most models, it is recommended that predictors have a hypothesized mechanism that you are testing for an effect on the response (dependent) variable. Another way to think about this is to make sure that you are prepared to interpret any of the effects (both their direction and magnitude) that come out of any of your predictors. It is also advised to consider any correlation(s) between predictor variables such that they are not correlated to a degree that could cause problems in the estimation (Dormann et al., 2013). Finally, consider if your variables need a certain error distribution (McCullagh & Nelder, 1989; Warton et al., 2016), error structure (Bolker et al., 2009), transformations (Gelman & Nolan, 2017), or standardization (Gelman, 2008).

Small sample size

Sample size was previously discussed as a general concern because when sample size simply cannot be large enough, there is a greater question of whether something should even be published. However, many studies are published with small sample sizes, and so we need to know how to handle any associated challenges. Small sample sizes increase the risk of spurious findings and false conclusions, compromising the validity and generalizability of statistical analyses (Lakens, 2022). Some useful approaches to dealing with small sample sizes can be to prevent them in the first place via power analysis, collect more data (if possible), make clear your sample sizes by reporting them (and possibly discussing subsequent limitations elsewhere in your article), and perhaps adding some type of simulation analysis to your study where you have the ability to understand your question with a larger, simulated data set (Peck, 2004). Simulated data sets have their own shortcomings, but sample size is rarely a limiting factor.

While not a small sample size issue, per se, here is a reasonable place to include the suggestion of considering nonparametric statistics when appropriate. Parametric statistical tests are given most of our attention in statistics—often for good reason, because of their robustness to violations of normality (Midway & White, 2025). When you can assume distributional information about your data, you can benefit from distributional inference from parametric models. However, despite the relative underuse of nonparametric statistics (Leech & Onwuegbuzie, 2002), they are commonly available, alleviate parametric assumptions, and often provide adequate statistical power—particularly in cases with interval and ratio scale data (Leech & Onwuegbuzie, 2002). Most common parametric tests have a nonparametric counterpart, and looking into nonparametric alternatives may reveal a variety of new tests for analyzing data. Several good resources on nonparametric methods are available (e.g., Hollander et al., 2013; Potvin & Roff, 1993).

Significant testing

Concerns with null hypothesis significance testing (NHST), often discussed through *P*-values, is perhaps one of the largest challenges in current statistics. Some disciplines and journals have gone so far as to discourage or ban *P*-values (Woolston, 2015), yet we are focusing on how information can be reported and not whether entire statistical domains should be avoided. But the challenges associated with *P*-values are very real, and a large number of studies can be referenced to understand both the problems that exist and how people have suggested dealing with them (Chen et al., 2023; Stephens et al., 2007; Wasserstein & Lazar, 2016). One common problem is known as *P*-hacking and data dredging, in which researchers conduct multiple statistical tests until a significant result is found (Head et al., 2015). This practice inflates the likelihood of Type I errors (i.e., falsely rejecting the null) and compromises the integrity of statistical inference. *P*-values are also widely used in hypothesis testing to assess the evidence against a null hypothesis. However, *P*-values are often misinterpreted as measures of effect size or the probability of the null hypothesis being true (see Goodman, 2008; Greenland et al., 2016 for common misconceptions about *P*-values). Reliance on arbitrary significance thresholds (e.g., $\alpha = 0.05$) can lead to dichotomous thinking and misrepresentation of statistical evidence (Amrhein et al., 2019). Everyone needs to remember that statistical significance does not necessarily equate to biological or ecological significance. For example, consider a study of fish growth between two lakes, where fish in the first lake were found to grow at an average rate of 5 cm per year, while fish in the second lake grew at an average rate of 6 cm per year. With a large enough sample, this 1-cm difference could be statistically significant. However, such a small difference likely has no effect on the species' biology or how they are managed.

Numerous alternatives have been proposed to move away from *P*-values (Stephens et al., 2007). A simple alternative to reporting only *P*-values would be to (also) report effect sizes. Effect sizes quantify the magnitude of relationships or differences observed in statistical analyses, providing valuable information about the practical significance of findings. Nakagawa and Cuthill (2007) and Stephens et al. (2007) provide accessible discussions of effect sizes for non-statisticians. While a *P*-value will decrease with sample size (and yes, you can achieve significance through increased sampling; see Weber et al., 2018), an effect size is a more straightforward measure of how much your predictor matters. Small effect sizes may still have statistical significance, and such a condition can be important to know because statistical significance alone is often interpreted as all-around importance.

One of the more extreme alternatives to NHST is to adopt a different statistical paradigm, such as Bayesian estimation. While Bayesian estimation certainly addresses many of the issues surrounding the use of *P*-values, a move to Bayesian estimation may represent a substantial undertaking in everything from technical expertise to philosophical positions about understanding the nature of uncertainty. For those curious about Bayesian methods, Doll and Lauer (2014) provide a fishery science example of an analysis done by both Bayesian and frequentist methods. Kruschke (2021) recently provided an excellent overview of how to report Bayesian methods.

As the debate about NHST has continued, many acknowledge that *P*-values and their underlying estimation remain a permanent fixture. Recent voices in this debate have focused their attention on language and how using different terms might lessen some of the concerns with traditional significance. Sterne and Smith (2001) have suggested ignoring the 0.05 significance threshold and simply reporting the *P*-value. Dushoff et al. (2019) propose a language of "clarity" over "significance" and Muff et al. (2022) adopt a linguistic framework of "evidence" over significance. Although we do not recommend one approach over another, authors (and reviewers) should be aware of these ongoing discussions of how we present our results and that your choice can still have an impact on your reader.

Model selection

Model selection is often provided as an alternative to NHST (Anderson et al., 2000), exemplified by techniques like the Akaike information criterion (AIC), and is essential in statistical analysis but is not without its challenges. A primary concern is the potential for overfitting, wherein complex models fit the data too closely, leading to poor generalization and applicability to other (new) data (Burnham & Anderson, 2004). Conversely, simpler models may underfit the data, failing to capture important patterns. Comparing multiple models increases the likelihood of Type I errors, particularly if adjustments for multiple comparisons are not made (Midway et al., 2020). Scientists must carefully consider model complexity, validate their chosen models, and account for potential biases to ensure the reliability and robustness of model selection procedures in statistical analysis.

When conducting model selection, scientists should mitigate the above problems first by limiting the number of models that are compared. Overly simple or complex models may even be excluded from being evaluated. Often, an all-subsets model selection is not recommended because it will include models that do not warrant consideration. Limiting the number of models considered also reduces the likelihood of Type I errors. Simply concluding that "model X was the best fit of our candidate models" does not mean that model X was any good to begin with. Another recommendation is to understand what different model selection routines do. For example, most information criterion (AIC, Bayesian information criterion, etc.) balance goodness of fit and model complexity (also referred to as parsimony), in which goodness of fit is calculated through the negative log-likelihood. However, a model selection routine like cross-validation will evaluate a model based on how it performs—often through prediction—on another data set. For these reasons, something like an AIC might be better for understanding one data set and less appropriate for prediction (Burnham & Anderson, 1998), while cross-validation might be better for selecting a model that is needed for exclusively predictive purposes (Yates et al., 2023).

SMALL ERRORS AND MISCELLANEOUS CONSIDERATIONS

Question and hypothesis

Make sure your study question or hypothesis (1) is clearly stated before the methods, (2) is paired with a statistical analysis (e.g., model) that will answer the question, and (3) is answered in the Results section (and discussed in the Discussion section). Although we do not get into the specifics of research questions here, many good papers and book chapters are available for reference. Betts et al. (2021) provides thoughts on research hypotheses, Brown and Guy (2007) cover a useful research framework for questions, and Hand (1994) gives an in-depth look at why research questions succeed or fail.

Keep it simple

Just because you have access to high-octane statistical methods does not mean you need them. Simpler methods mean a lower likelihood of mistakes in execution and a larger audience familiar with your approach (Murtaugh, 2007). As the statistician Sir David Cox said, "Begin with very simple methods. If possible, end with simple methods."

Report your assumptions

All models have assumptions. In fact, you can usually get by without perfectly satisfying all of them. But make sure to both report the statistical assumptions under which you are working, and to follow up with any analyses to address the assumptions. For example, normally distributed residuals are an assumption of linear regression (Midway & White, 2025), which you should mention along with any description of the residuals from the regression model.

Consider an equation

Despite the finding that increasing the number of equations resulted in fewer citations (Fawcett & Higginson, 2012), many studies have no equations at all. Outside of extremely simple models (e.g., simple linear regression), consider an equation to clarify your model. Authors will often use a paragraph or more to explain what can be confirmed in one line of notation.

Use statistical notation

If you include an equation (see above point), report the equation using statistical notation. In other words, do not include an equation using code, syntax, or any other statistical or programming language. The model glm($y \sim x$, family = binomial) may make sense to you, but many people do not use R. An additional risk of using code is that over time programs change, and you want your science to age using the universal language of statistical notation—just ask the people who used Bio-Medical Data Package, which is no longer available. Edwards and Auger-Méthé (2019) provide an overview of how to use statistical notation.

Multiple comparisons

You may have used Tukey's honestly significant difference test because it is a common post hoc multiple comparisons test. There are many multiple comparisons tests out there and most are designed for specific situations. See Midway et al. (2020) for an overview of options for multiple comparison tests so that your selection is justified.

Know uncertainty

Measures of uncertainty are too numerous to cover here, but as with any metric you use and report, be sure you know what

it is used for. One common example worth mentioning is standard deviation vs. standard error, which are often confused. Standard deviation is a descriptive statistic (or parameter of the normal distribution), whereas standard error is an inferential statistic that tends to decrease with increasing sample size.

Significant digits

You can't have more accuracy in an estimate than you had in the measurement. Statistical software now often reports a lot of decimal places, usually beyond the precision of the measurement. Make sure that values are not more precise than your measurement and the values are consistent throughout the manuscript.

Subheadings help organize

Depending on the journal format, you may be able to use subheadings to your advantage. If your Methods section includes data collection, data manipulation, the model, and multiple comparisons, consider subheadings or at least four separate paragraphs to help distinguish the main steps and sequence of your methods. Doing so will help ensure that you provide an appropriate level of detail within each element as well as help the reader understand your approach. The sequence you adopt can also be mimicked (if possible) in the Results section so the Methods and Results are organizationally parallel and reduce cognitive friction with a reader who might be expecting certain content.

Use supplements

The digital age we live in has made at least one thing easier: supplements. When documenting methods or results, there are often pieces that are important but might not warrant inclusion in the main document (especially as journals require shorter submissions with higher page charges). Supplements remain the perfect compromise to including information in perpetuity without overloading the main document. Use them. Thanks in advance!

Include data and code

Although not always possible, it can be very helpful to your reader to include your data and/or code as supplementary files. Best practices for managing data are out there (Borghi et al., 2018; Broman & Woo, 2018; Rüegg et al., 2014) along with code (Filazzola & Lortie, 2022).

Effective visuals

Figures are usually reserved for the Results section, where they can greatly enhance and amplify your findings. In addition to making these visuals clear and effective (Midway, 2020), consider a diagram or other visual for statistical methods—particularly if your method is complex, has multiple steps, or can otherwise be confusing. There are important cognitive aspects to figure creation, and considering them can yield great understanding (Midway et al., 2023). Trushenski et al. (2019) provides an excellent example methods diagram that clearly outlines an otherwise complex procedure of experiments.

Meta-analyses

Most of what we discuss in this article pertains to original research but be aware that if you are doing a meta-analysis, there may be some differences in how you use and report your statistics. Fortunately, there are a few papers out there that discuss this topic and provide guidance (Gurevitch & Hedges, 1999; O'Dea et al., 2021).

CONCLUSION

Statistical reporting can be challenging because there is no clear finish line. Nearly all studies could benefit from some improvement. Additionally, statistical methods are continually evolving, and it can be daunting—especially for early career researchers—to figure out where to start. What is clear to one person may not be clear to another. Sometimes authors are more knowledgeable about a complex statistical analysis than are the reviewers or editors in peer review. Despite these persistent challenges—which may only increase in the future—best practices are emerging, and the topic of statistical reporting continues to be recognized (Hardwicke et al., 2023; Popovic et al., 2024; Zuur & Ieno, 2016).

When thinking about what can be done, a first step is for authors to be critical of their written methods, so that they are not writing in isolation and under the assumption that others know what they did. Review Figure 1 and other references we have included. We encourage authors to have their statistical reporting reviewed (before submitting to a journal)—something that can happen without reviewing the entire manuscript. Find a colleague who is not a coauthor and ask them to read your Methods and Results for clarity. People who are not part of the study are often excellent judges of what makes sense to an outsider. Artificial intelligence (generative chatbots, such as ChatGPT) also creates the ability to review writing and, when prompted, can provide feedback on clarity, reproducibility, or any other aspect you may have concerns about. Once submitted, there is a responsibility of journal editorial teams and peer reviewers to make sure they understand what is written, as this stage is often the last chance to catch errors before the permanence of publication. Some errors will still make it through, but we have tools to reduce them. The call remains for each and every author to make sure that what they write is clear and reproducible, and to make sure that all analytical actions are in service of a solid question. The only way we are going to help each other learn what we have done is to clearly explain it.

DATA AVAILABILITY

No data was recorded, created, or used in this study.

ETHICS STATEMENT

No animals or human subjects were used in this study.

FUNDING

The authors received no funding in support of this study.

CONFLICTS OF INTEREST

None declared.

ACKNOWLEDGMENTS

The authors thank countless colleagues, students, and others with whom they have had meaningful conversations and shared ideas about how to communicate statistics effectively.

REFERENCES

Amrhein, V., Trafimow, D., & Greenland, S. (2019). Inferential statistics as descriptive statistics: There is no replication crisis if we don't expect replication. *The American Statistician*, *73*, 262–270. https://doi.org/10.1080/00031305.2018.1543137

Anderson, D. R., Burnham, K. P., & Thompson, W. L. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *The Journal of Wildlife Management*, *64*, 912–923. https://doi.org/10.2307/3803199

Barraquand, F., Ezard, T. H., Jorgensen, P. S., Zimmerman, N., Chamberlain, S., Salguero-Gomez, R., Curran, T. J., & Poisot, T. (2014). Lack of quantitative training among early-career ecologists: A survey of the problem and potential solutions. *PeerJ*, *2*, Article e285. https://doi.org/10.7717/Peerj.285

Betts, M. G., Hadley, A. S., Frey, D. W., Frey, S. J. K., Gannon, D., Harris, S. H., Kim, H., Kormann, U. G., Leimberger, K., Moriarty, K., Northrup, J. M., Phalan, B., Rousseau, J. S., Stokely, T. D., Valente, J. J., Wolf, C., & Zarrate-Charry, D. (2021). When are hypotheses useful in ecology and evolution? *Ecology and Evolution*, *11*, 5762–5776. https://doi.org/10.1002/Ece3.7365

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H., & White, J. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*, 127–135. https://doi.org/10.1016/J.Tree.2008.10.008

Borghi, J., Abrams, S., Lowenberg, D., Simms, S., & Chodacki, J. (2018). Support your data: A research data management guide for researchers. *Research Ideas and Outcomes*, *4*, Article e26439. https://doi.org/10.3897/Rio.4.E26439

Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician*, *72*, 2–10. https://doi.org/10.1080/00031305.2017.1375989

Brown, M. L., & Guy, C. S. (2007). Science and statistics in fisheries research. In C. S. Guy, & M. L. Brown (Eds.), *Analysis and interpretation of freshwater fisheries data* (pp. 1–29). American Fisheries Society.

Burnham, K. P., & Anderson, D. R. (1998). Practical use of the information-theoretic approach. In *Model selection and inference* (pp. 75–117). Springer.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference. *Sociological Methods & Research*, *33*, 261–304. https://doi.org/10.1177/0049124104268644

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews: Neuroscience*, *14*, 365–376. https://doi.org/10.1038/Nrn3475

Chen, O. Y., Bodelet, J. S., Saraiva, R. G., Phan, H., Di, J., Nagels, G., Schwantje, T., Cao, H., Gou, J., Reinen, J. M., Xiong, B., Zhi, B., Wang, X., & De Vos, M. (2023). The roles, challenges, and merits of the *P* value. *Patterns*, *4*, Article 100878. https://doi.org/10.1016/J.Patter.2023.100878

Davis, A. J., & Kay, S. (2023). Writing statistical methods for ecologists. *Ecosphere*, *14*, Article e4539. https://doi.org/10.1002/Ecs2.4539

Doll, J. C., & Lauer, T. E. (2014). Comparing Bayesian and frequentist methods of fisheries models: Hierarchical catch curves. *Journal of Great Lakes Research*, *40*, 41–48. https://doi.org/10.1016/J.Jglr.2014.07.006

Dorazio, R. M. (2016). Bayesian data analysis in population ecology: Motivations, methods, and benefits. *Population Ecology*, *58*, 31–44. https://doi.org/10.1007/S10144-015-0503-4

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 27–46. https://doi.org/10.1111/J.1600-0587.2012.07348.X

Dushoff, J., Kain, M. P., & Bolker, B. M. (2019). I can see clearly now: Reinterpreting statistical significance. *Methods in Ecology and Evolution*, *10*, 756–759. https://doi.org/10.1111/2041-210X.13159

Edwards, A. M., & Auger-Méthé, M. (2019). Some guidance on using mathematical notation in ecology. *Methods in Ecology and Evolution*, *10*, 92–99. https://doi.org/10.1111/2041-210X.13105

Fawcett, T. W., & Higginson, A. D. (2012). Heavy use of equations impedes communication among biologists. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11735–11739. https://doi.org/10.1073/Pnas.1205259109

Filazzola, A., & Lortie, C. J. (2022). A call for clean code to effectively communicate science. *Methods in Ecology and Evolution*, *13*, 2119–2128. https://doi.org/10.1111/2041-210x.13961

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, *27*, 2865–2873. https://doi.org/10.1002/Sim.3107

Gelman, A., & Nolan, D. (2017). *Teaching statistics: A bag of tricks* (2nd Ed.). Oxford University Press.

Goodman, S. (2008). A dirty dozen: Twelve *P*-value misconceptions. *Seminars in Hematology*, *45*, 135–140. https://doi.org/10.1053/J.Seminhematol.2008.04.003

Gould, E., Fraser, H., Parker, T., Nakagawa, S., Griffith, S., Vesk, P., Fidler, F., Abbey-Lee, R., Abbott, J., Aguirre, L., Alcaraz, C., Altschul, D., Arekar, K., Atkins, J., Atkinson, J., Barrett, M., Bell, K., Bello, S., Berauer, B., … Tompkins, E. (2025). Same data, different analysts: Variation in effect sizes due to analytical decisions in ecology and evolutionary biology. *BMC Biology*, *23*, Article 35. https://doi.org/10.1186/S12915-024-02101-X

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, *P* values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*, 337–350. https://doi.org/10.1007/S10654-016-0149-3

Gurevitch, J., & Hedges, L. V. (1999). Statistical issues in ecological meta-analyses. *Ecology*, *80*, 1142–1149. https://doi.org/10.1890/0012-9658(1999)080[1142:Siiema]2.0.Co;2

Guy, C. S., & Brown, M. L. (2007). *Analysis and interpretation of freshwater fisheries data*. American Fisheries Society.

Haddon, M. (2020). *Using R for modelling and quantitative methods in fisheries*. Chapman and Hall.

Hand, D. J. (1994). Deconstructing statistical questions. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, *157*, 317–356. https://doi.org/10.2307/2983526

Hardwicke, T. E., Salholz-Hillel, M., Malički, M., Szűcs, D., Bendixen, T., & Ioannidis, J. P. A. (2023). Statistical guidance to authors at top-ranked journals across scientific disciplines. *The American Statistician*, *77*, 239–247. https://doi.org/10.1080/00031305.2022.2143897

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of *P*-hacking in science. *PLOS Biology*, *13*, Article e1002106. https://doi.org/10.1371/Journal.Pbio.1002106

Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric statistical methods* (3rd ed.). Wiley.

Hunter, J. (1990). *Writing for fisheries journals*. American Fisheries Society.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, Article e124. https://doi.org/10.1371/Journal.Pmed.0020124

Jennions, M. D. (2003). A survey of the statistical power of research in behavioral ecology and animal behavior. *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology*, *14*, 438–445. https://doi.org/10.1093/Beheco/14.3.438

Kimmel, K., Avolio, M. L., & Ferraro, P. J. (2023). Empirical evidence of widespread exaggeration bias and selective reporting in ecology. *Nature Ecology & Evolution*, *7*, 1525–1536. https://doi.org/10.1038/S41559-023-02144-3

Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behaviour*, *5*, 1282–1291. https://doi.org/10.1038/S41562-021-01177-7

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, *8*, Article 33267. https://doi.org/10.1525/Collabra.33267

Leech, N. L., & Onwuegbuzie, A. J. (2002, November 6–8). A call for greater use of nonparametric statistics [paper presentation]. Annual Meeting of the Mid-South Educational Research Association, Chattanooga, Tennessee.

Legendre, P., & Legendre, L. (2012). *Numerical ecology*. Elsevier.

Low-Décarie, E., Chivers, C., & Granados, M. (2014). Rising complexity and falling explanatory power in ecology. *Frontiers in Ecology and the Environment*, *12*, 412–418. https://doi.org/10.1890/130230

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall.

Michel, M. C., Murphy, T. J., & Motulsky, H. J. (2020). New author guidelines for displaying data and reporting data analysis and statistical methods in experimental biology. *Molecular Pharmacology*, *97*, 49–60. https://doi.org/10.1124/Mol.119.118927

Midway, S. (2020). Principles of effective data visualization. *Patterns*, *1*, Article 100141. https://doi.org/10.1016/J.Patter.2020.100141

Midway, S., Brum, J. R., & Robertson, M. (2023). Show and tell: Approaches for effective figures. *Limnology and Oceanography Letters*, *8*, 213–219. https://doi.org/10.1002/Lol2.10288

Midway, S., Robertson, M., Flinn, S., & Kaller, M. (2020). Comparing multiple comparisons: Practical guidance for choosing the best multiple comparisons test. *PeerJ*, *8*, Article e10387. https://doi.org/10.7717/Peerj.10387

Midway, S. R., & White, J. W. (2025). Testing for normality in regression models: Mistakes abound (but may not matter). *Royal Society Open Science 12*, Article 241904. https://doi.org/10.1098/rsos.241904

Muff, S., Nilsen, E. B., O'Hara, R. B., & Nater, C. R. (2022). Rewriting results sections in the language of evidence. *Trends in Ecology & Evolution*, *37*, 203–210. https://doi.org/10.1016/J.Tree.2021.10.009

Murtaugh, P. A. (2007). Simplicity and complexity in ecological data analysis. *Ecology*, *88*, 56–62. https://doi.org/10.1890/0012-9658(2007)88[56:Sacied]2.0.Co;2

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society*, *82*, 591–605. https://doi.org/10.1111/J.1469-185x.2007.00027.X

O'Dea, R. E., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W. A., Parker, T. H., Gurevitch, J., Page, M. J., Stewart, G., Moher, D., & Nakagawa, S. (2021). Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: A prisma extension. *Biological Reviews of the Cambridge Philosophical Society*, *96*, 1695–1722. https://doi.org/10.1111/Brv.12721

Ogle, D. H. (2015). *Introductory fisheries analyses with R*. CRC Press.

Peck, S. L. (2004). Simulation as experiment: A philosophical reassessment for biological modeling. *Trends in Ecology & Evolution*, *19*, 530–534. https://doi.org/10.1016/J.Tree.2004.07.019

Peng, R. D. (2009). Reproducible research and *Biostatistics*. *Biostatistics*, *10*, 405–408. https://doi.org/10.1093/Biostatistics/Kxp014

Peng, R. D., & Hicks, S. C. (2021). Reproducible research: A retrospective. *Annual Reviews of Public Health*, *42*, 79–93. https://doi.org/10.1146/Annurev-Publhealth-012420-105110

Pinker, S. (2014). *The sense of style: The thinking person's guide to writing in the 21st century*. Penguin Books.

Popovic, G., Mason, T. J., Drobniak, S. M., Marques, T. A., Potts, J., Joo, R., Altwegg, R., Burns, C. C. I., McCarthy, M. A., Johnston, A., Nakagawa, S., McMillan, L., Devarajan, K., Taggart, P. L., Wunderlich, A., Mair, M. M., Martínez-Lanfranco, J. A., Lagisz, M., & Pottier, P. (2024). Four principles for improved statistical ecology. *Methods in Ecology and Evolution*, *15*, 266–281. https://doi.org/10.1111/2041-210x.14270

Potvin, C., & Roff, D. A. (1993). Distribution-free and robust statistical methods: Viable alternatives to parametric statistics. *Ecology*, *74*, 1617–1628. https://doi.org/10.2307/1939920

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation For Statistical Computing. https://www.R-Project.Org/

Rüegg, J., Gries, C., Bond-Lamberty, B., Bowen, G. J., Felzer, B. S., Mcintyre, N. E., Soranno, P. A., Vanderbilt, K. L., & Weathers, K. C. (2014). Completing the data life cycle: Using information management in macrosystems ecology research. *Frontiers in Ecology and the Environment*, *12*, 24–30. https://doi.org/10.1890/120375

Stephens, P. A., Buskirk, S. W., & Del Rio, C. M. (2007). Inference in ecology and evolution. *Trends in Ecology & Evolution*, *22*, 192–197. https://doi.org/10.1016/J.Tree.2006.12.003

Sterne, J. A., & Smith, G. D. (2001). Sifting the evidence—what's wrong with significance tests? *BMJ*, *322*, Article 226. https://doi.org/10.1136/Bmj.322.7280.226

Thomas, L. E. N., & Juanes, F. (1996). The importance of statistical power analysis: An example from *Animal Behaviour*. *Animal Behaviour*, *52*, 856–859. https://doi.org/10.1006/Anbe.1996.0232

Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, *102*, Article e03336. https://doi.org/10.1002/Ecy.3336

Trushenski, J. T., Larsen, D. A., Middleton, M. A., Jakaitis, M., Johnson, E. L., Kozfkay, C. C., & Kline, P. A. (2019). Search for the smoking gun: Identifying and addressing the causes of postrelease morbidity and mortality of hatchery-reared Snake River Sockeye Salmon smolts. *Transactions of the American Fisheries Society*, *148*, 875–895. https://doi.org/10.1002/Tafs.10193

Warton, D. I., Lyons, M., Stoklosa, J., Ives, A. R., & Schielzeth, H. (2016). Three points to consider when choosing a LM or GLM test for count data. *Methods in Ecology and Evolution*, *7*, 882–890. https://doi.org/10.1111/2041-210x.12552

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on *P*-values: Context, process, and purpose. *The American Statistician*, *70*, 129–133. https://doi.org/10.1080/00031305.2016.1154108

Weber, F., Hoang Do, J. P., Chung, S., Beier, K. T., Bikov, M., Saffari Doost, M., & Dan, Y. (2018). Regulation of REM and non-REM sleep by periaqueductal gabaergic neurons. *Nature Communications*, *9*, Article 354. https://doi.org/10.1038/S41467-017-02765-W

Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and management of animal populations*. Academic Press.

Woolston, C. (2015). Psychology journal bans *P* values. *Nature*, *519*, Article 9. https://doi.org/10.1038/519009f

Wootton, R. J., & Craig, J. F. (2011). Reporting statistical results. *Journal of Fish Biology*, *78*, 697–699. https://doi.org/10.1111/J.1095-8649.2011.02914.X

Yates, L. A., Aandahl, Z., Richards, S. A., & Brook, B. W. (2023). Cross validation for model selection: A review with examples from ecology. *Ecological Monographs*, *93*, Article e1557. https://doi.org/10.1002/Ecm.1557

Zuur, A. F., & Ieno, E. N. (2016). A protocol for conducting and presenting results of regression-type analyses. *Methods in Ecology and Evolution*, *7*, 636–645. https://doi.org/10.1111/2041-210X.12577