Code-sharing policies are associated with increased reproducibility potential of ecological findings

Alfredo Sánchez-Tójar[1*], Aya Bezine[1], Marija Purgar[2,3], Antica Culina[2,4*]

[1] Department of Evolutionary Biology, Bielefeld University, Bielefeld, Germany; [2] Ruder Boskovic Institute, Zagreb, Croatia; [3] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA; [4] Netherlands Institute of Ecology, NIOO-KNAW, Wageningen, the Netherlands; * Corresponding authors: alfredo.tojar@gmail.com; Antica.Culina@irb.hr

ORCID: Alfredo Sánchez-Tójar: 0000-0002-2886-0649; Aya Bezine: 0009-0005-8202-152X; Marija Purgar: 0000-0002-7192-1486; Antica Culina: 0000-0003-2910-8085

Data and code availability statement

All data and code are available at the following GitHub repository (https://github.com/ASanchez-Tojar/code-sharing_policies_matter) and in Zenodo (https://doi.org/10.5281/zenodo.14357339).

Abstract

Software code (e.g., analytical code) is increasingly recognised as an important research output, as it improves transparency, collaboration, and research credibility. Many scientific journals have introduced code-sharing policies; however, surveys show alarmingly low compliance with these policies. In this study, we expand on a recent survey of ecological journals with code-sharing policies by investigating sharing practices in a comparable set of ecological journals without code-sharing policies. Our aims were to estimate code- and data-sharing rates, assess key reproducibility-boosting features like the reporting of software versioning, and compare reproducibility potential between journals with and without a code-sharing policy. We reviewed a random sample of 314 articles published between 2015-2019 across 12 ecological journals without a code-sharing policy. Only 15 articles (4.8%) provided analytical code, with the percentage nearly tripling over time (2015-2016: 2.5%, 2018-2019: 7.0%). Data-sharing was higher than code-sharing (2015-2016: 31.0%, 2018-2019: 43.3%), yet only 8 articles (2.5%) shared both code and data. Compared with a comparative sample of 346 articles from 14 ecological journals with a code-sharing policy, journals without code-sharing policies showed 5.6 times lower code-sharing, 2.1 times lower data-sharing, and 8.1 times lower reproducibility potential. Despite these differences, key reproducibility-boosting features between the two types of journals were similar. About 90% of all articles reported the analytical software used; however, for journals with and without a code-sharing policy, software version was often missing (49.8% and 36.1% of articles, respectively), and only proprietary (i.e., non-free) software was used in 16.7% and 23.5% of articles, respectively. Our study suggests that journals with code-sharing policies have greater reproducibility potential than those without. Code-sharing policies are likely a necessary but insufficient key step toward increasing reproducibility. Journals should prioritize adopting explicit, easy-to-find and strict code-sharing policies to facilitate researcher compliance as well as implement mechanisms such as checklists to ensure compliance.

Keywords: replicability, reliability, robustness, generalizability, verification, replication, FAIR, checklist

Introduction

Sharing software code is essential for robust, reproducible and impactful science (Peng 2011; Borregaard and Hart 2016; Lewis et al. 2018; Cole et al. 2024). Software code is used for processing and analysing data, creating figures, and even producing fully executable articles (Mislan et al. 2016; Lasser 2020), and its complexity is increasing (Touchon and McCoy 2016; Feng et al. 2020). Code helps with understanding and critically evaluating data analysis, and importantly, can be used and extended by others, allowing faster scientific progress (Cadwallader et al. 2022). The computational reproducibility of scientific findings (i.e., using the same code on the same data to reproduce the same results; Benureau and Rougier 2018), a seemingly simple, but in practice difficult-to-achieve feature in modern science (e.g., Campbell et al. 2023; Kambouris et al. 2024), greatly improves when analytical code is available (Laurinavichyute et al. 2022).

Code availability has been slowly increasing in ecology (Maitner et al. 2024; Sperandii et al. 2024) and other fields (Cao et al. 2023; but see Serghiou et al. 2022), likely as a consequence of several changes. First, software and software code are becoming recognised as essential research output (DORA: https://sfdora.org/read/; ReSA: https://www.researchsoft.org/; Jay et al. 2021). Second, training and guidelines on reproducible code and software management are more available to researchers (Donoho et al. 2008; McKiernan 2017; Kohrs et al. 2023). Third, funders and journals have been slowly but steadily introducing code-sharing policies. For example, the percentage of journals with code-sharing policies increased rapidly for a subset of 96 ecological journals, from 15% in 2015 (Mislan et al., 2016) to 75% in 2020 (Culina et al. 2020). Recently, a larger survey of 275 journals in ecology and evolution found that 72% mandate or encourage code-sharing as of 2024 (Ivimey-Cook et al. *In prep.*). While the evidence that the mere existence of journal code-sharing policies likely increases code availability is accumulating (Cadwallader et al. 2022; Fišar et al. 2024; Ivimey-Cook et al. *In prep.*), policy compliance remains alarmingly low. For example, only 27% of articles published between 2015 and 2019 in the subset of 96 ecological journals with code-sharing policies shared their code (Culina et al. 2020), showing that policies are only partially efficient if they are not enforced. In addition, policies that do not specify and require best-sharing practices likely lead to low code reusability, and ultimately low reproducibility of scientific findings.

Code-sharing itself does not necessarily translate into code that is easy to understand, adapt and reuse. Multiple technical challenges to code reuse range from dependencies on the original researcher's computational environment such as the operating system and libraries used, to inadequate documentation on how to install, run, and use the code (Boettiger et al. 2015). Code can also easily *rot* after software updates are released, leading to changes in the functionality, compatibility and, ultimately, the reproducibility of the results (Hinsen 2019). Although container technology such as Docker, which packages the software and its dependencies into a standardized environment, has been suggested as a solution to improving portability and reproducibility (Boettiger et al. 2015; Grüning et al. 2018; Essawy et al. 2020; Trisovic et al. 2022), its adoption remains low. At the minimum, the software and packages used for the analyses should be stated and appropriately referenced, and the version(s) used clearly stated in the manuscript and/or as part of stand-alone documentation (e.g., README, or inline comments; Benureau and Rougier 2018; Jenkins et al. 2023; Ivimey-Cook et al. 2023). In addition, code should ideally be written using free (i.e., non-proprietary) and open-source software (also known as FOSS; Ostermann and Granell 2016) such as the free and open-source R programming language (R Core Team 2023) that is widely used in ecology (Lai et al. 2019; Culina et al. 2020; Kambouris et al. 2024). Further, code should be shared in a permanent repository (e.g., Zenodo) and assigned with an open and permissive licence and a persistent identifier such as a DOI (Krafczyk et al. 2021; Kim et al. 2022; Jenkins et al. 2023). This is particularly important given the far-from-ideal rates of link persistence found for scientific code in fields such as astrophysics (Allen et al. 2018).

In this work, we study whether implementing code-sharing policies leads to higher rates of code-sharing. In addition, we explore the reporting of features associated with higher long-term reproducibility in journals with and without code-sharing policies. We assess the code-sharing and reporting features of 314 articles published in 12 ecological journals without code-sharing policies and compare them with those from a comparable sample of 346 articles published in 14 ecological journals with code-sharing policies. We predict that ecological journals without code-sharing policies will have lower rates of sharing compared to journals with code-sharing policies. However, we do not have a clear expectation on whether the reporting of features associated with higher long-term reproducibility such as the software used, its versioning and accessibility (free or not), and the location where code is shared will differ between both sets of journals. This is because many code-sharing policies are not explicit (Ivimey-Cook et al. *In prep.*), and thus they might not explicitly prompt the authors to follow best practices, whereas authors who share their code in the absence of code-sharing policies might be primed to follow best practices. Finally, we anticipate that code availability and the reporting of features associated with higher long-term reproducibility will both increase over time, regardless of the existence of code-sharing policies, given recent changes in scientific attitudes and norms, and the rise of open science (Cao et al. 2023).

Methods

Our study design closely matches that of Culina et al. (2020) who surveyed 14 ecological journals that had a code-sharing policy from at least 2015 to 2019. In a follow-up study here, we aimed to identify 14 comparable ecological journals without a code-sharing policy for the same period (i.e., 2015-2019). For that, we used the set of 96 ecological journals originally assessed by Mislan et al. (2016) and subsequently reassessed by Culina et al. (2020), and identified 12 journals without a code-sharing policy as of 2020. This was done by carefully reading the author guidelines and open research policies of these journals compiled by Culina et al. (2020). While initially, we identified 24 potentially eligible journals (i.e., without a code-sharing policy), we later removed from the list two review journals (‘Trends in Ecology and Evolution’, and ‘Annual Review of Ecology, Evolution, and Systematics’), nine journals that mentioned code as part of their data-sharing policy (‘Aquatic Microbial Ecology’, ‘Behavioral Ecology and Sociobiology’, ‘Ecology and Evolution’, ‘Global Change Biology’, ‘Journal of Soil and Water Conservation’, ‘Marine Ecology Progress Series’, ‘Microbial Ecology’, ‘Oryx’, and ‘Paleobiology’), and one journal that had been discontinued (‘Journal of the North American Benthological Society’). We judged the remaining 12 journals eligible (i.e., no code-sharing policy by March 2020; see Table S1 in Culina et al. 2020), as they did not mention programming code or other terms that could be interpreted as such (e.g., script, research artefacts) in their author guidelines: ‘Basic and Applied Ecology’, ‘Behavioral Ecology’, ‘Ecosystems’, ‘Freshwater Science’, ‘Frontiers in Ecology and the Environment’, ‘International Journal of Sustainable Development and World Ecology’, ‘Journal of Plant Ecology’, ‘Landscape Ecology’, ‘Oecologia’, ‘Oikos’, ‘Polar Research’, and ‘Wildlife Research’. Note that since the initial screening in March 2020, some of these journals might have adopted code-sharing policies; however, this would not affect our study as here we focused on articles published between 2015 and 2019.

We performed a search in Web of Science Core Collection (databases covered: Science Citation Index Expanded (SCI-EXPANDED) since 1945, Social Sciences Citation Index (SSCI) since 1956, Arts & Humanities Citation Index (AHCI) since 1975, Emerging Sources Citation Index (ESCI) since 2017) in February 2022, and extracted all the records published in those 12 journals during the same two distinct temporal periods as Culina et al. (2020): (i) from the 1st of June of 2015 to the 9th of December 2016 (N = 2499 records), and (ii) from the 1st of January 2018 to the 21st of May 2019 (N = 2275 records). We then took a random sample of 200 articles from each of these two periods (N = 400 in total) using the function 'sample()' in R v.4.3.1 (R Core Team, 2023). We screened their titles and abstracts for eligibility using the software Rayyan (Ouzzani et al. 2016). To meet our inclusion criteria, an article had to conduct a statistical analysis, develop and run a mathematical model, or conduct simulations. Following Culina et al. (2020), we excluded reviews, opinions, commentaries and purely bioinformatics studies. In addition, we excluded two articles from the 2018-2019 subset that performed landscape analyses because we lacked the expertise to understand the analyses and software used. Each article was screened by two reviewers (AB, AST) and conflicts among them (~5%) were resolved collectively. In total, 314 nonmolecular articles passed the title-and-abstract screening and their full-text was read in detail for data extraction. The screening process is presented in the PRISMA diagram (Figure S1; O'Dea et al. 2021).

Data extraction for each article was conducted by two reviewers (AB and either AST, AC, or MP) to increase the reproducibility and reliability of the data extraction process. Any conflicts were resolved by involving a third reviewer and are marked and explained in the provided data (see 'data and code availability statement'). For each article, we recorded (i) bibliographic information (title, authors, journal, publication year), (ii) the type of analyses conducted since our interest was only on articles performing statistical analyses and/or simulations, (iii) whether code and data (if used) had been shared (levels: yes, no, partially), (iv) for instances of shared code, we recorded where it was shared (levels: repository, supplementary material, website), and the name of the repository (if any used), and (iv) several additional key reproducibility-boosting features (i.e., software and additional package(s)/extension(s) (hereafter referred to as "package(s)") used, number of software and package(s) for which version was provided, and whether the software used was free (i.e., non-proprietary; levels: yes, no, partially)).

To test whether reproducibility potential is higher in journals with vs without code-sharing policies, we revisited, updated, and extended the dataset used for the analyses presented by Culina et al. (2020). Specifically, for the 346 nonmolecular articles included in Culina et al. (2020), we extracted the package(s) used and the number of software and package(s) for which version was provided. We also checked already collected variables of interest from Culina et al (2020) for any inconsistencies.

Post-hoc decisions we took when processing our data were: (1) whenever packages or extensions were not reported, we assigned the number of packages as 0, even if the software used may not actually have any packages or extensions. We did this because it was not possible for us to find out information about the existence of packages or extensions for all the software reported; (2) for articles that shared some or all of their data only within figures (e.g., in a scatterplot), we assigned them as not sharing their data; (3) we searched for software and package versions not only within the text but also in the reference list of the corresponding article; (4) in some rare cases when the article did not report the software used but we could infer it from the packages or extensions reported, we assigned the software as "Not Stated".

Results

Code- and data-sharing

We investigated a total of 314 nonmolecular articles that performed statistical analyses or simulations and were published between 2015 and 2019 (2015-2016: 157 articles, 2018-2019: 157 articles) in 12 ecological journals without a code-sharing policy as of March 2020. In these 12 journals, the statistical analysis or simulation code underlying the research findings was shared in only 15 of 314 articles (4.8%). Those 15 articles were accompanied by either seemingly all (10 articles, 3.2%) or some (5 articles, 1.6%) of the code. The overall percentage of code shared increased by about threefold over the two periods (2.5% versus 7.0%, in 2015–2016 and 2018–2019, respectively; Figure 1a). At the journal level, the percentage of articles where code was shared ranged between 0% and 8.7% (median = 1.2%, mean =

3.1%; Table 1), indicating that not sharing code is a general phenomenon across ecological journals without a code-sharing policy. Of those 15 articles that shared code, 12 (80%) provided it as part of the article's supplementary material, 1 (6.7%) at a website, and only 2 (13.3%) in a repository (i.e., Dryad).

Table 1. Code- and data-sharing for 314 nonmolecular articles that conducted statistical analysis or simulations published between 2015 and 2019 in 12 ecological journals without a code-sharing policy.

Journal	**Total number of articles sampled [using data]**	**Number of articles providing code (%)**	**Number of articles providing data (%)**
Basic and Applied Ecology	12 [12]	0 (0.0%)	2 (16.7%)
Behavioral Ecology	44 [44]	3 (6.8%)	25 (56.8%)
Ecosystems	23 [23]	2 (8.7%)	10 (43.5%)
Freshwater Science	16 [16]	0 (0.0%)	1 (6.2%)
Frontiers in Ecology and the Environment	4 [4]	0 (0.0%)	2 (50.0%)
International Journal of Sustainable Development and World Ecology	6 [6]	0 (0.0%)	1 (16.7%)
Journal of Plant Ecology	20 [20]	0 (0.0%)	4 (20.0%)
Landscape Ecology	44 [44]	1 (2.3%)	21 (47.7%)
Oecologia	79 [79]	5 (6.3%)	19 (24.1%)
Oikos	42 [40]	3 (7.1%)	25 (62.5%)
Polar Research	7 [7]	0 (0.0%)	4 (57.1%)
Wildlife Research	17 [17]	1 (5.9%)	2 (11.8%)

In the 12 journals without a code-sharing policy, data were shared in 116 of 312 nonmolecular articles that used data (37.2%). These articles were accompanied by either seemingly all (75 articles, 24.0%) or some (41 articles, 13.1%) of the data and the overall percentage of data shared increased by about 40% over the 5-year period studied (31.0% versus 43.3%, in 2015–2016 and 2018–2019, respectively; Figure 1b). Furthermore, at the journal level, the percentage of articles where data were shared ranged between 6.2% and 62.5% (median = 33.8%, mean = 34.4%; Table 1), suggesting large differences in data-sharing across the 12 ecological journals without a code-sharing policy.

Altogether, only 8 (2.5%) articles had seemingly shared both all data (if any used) and all code, meaning that the potential for computational reproducibility in the 12 ecological journals without any code-sharing policy surveyed in our study could be as low, and likely lower than 2.5%. This percentage is 8.2 times smaller than the corresponding percentage found in journals with a code-sharing policy (20.8%; Culina et al. 2020).

Features boosting long-term reproducibility in journals with and without a code-sharing policy

Our survey showed that 11.8% of articles (N = 37) published in journals without a code-sharing policy did not state the analytical software used (Figure 2a), a value that is only slightly larger than the 10.1% (N = 35) found for articles published in journals with a code-sharing policy (Culina et al. 2020; Figure 2b). For those stating the statistical software used, 36.1% (N = 100) of articles published in journals without a code-sharing policy did not report the version of all software used (Figure 2c), whereas that percentage was 49.8% (N = 155) for articles published in journals with a code-sharing policy (Figure 2d). The mean number of analytical software used was 1.27 (median = 1.00, range = 1 to 6) in journals without a code-sharing policy and 1.81 (median = 1.00, range = 1 to 14) in journals with a code-sharing policy. The reporting of software versioning remained slightly higher for journals without a code-sharing policy than those with when expressed as the average percentage of software with version per article (without policy: median = 100%, mean = 67.5%, range: 0 to 100%; with policy: median = 100%, mean = 59.6%, range: 0 to 100%).

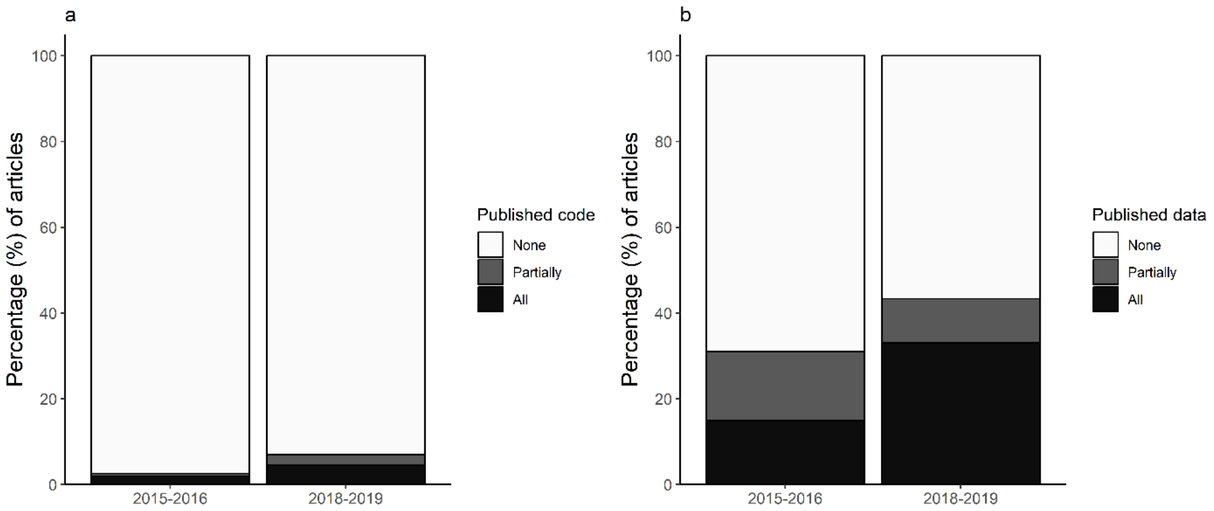

Figure 1. Code- and data-sharing are uncommon in 12 ecological journals without a code-sharing policy. Percentage of nonmolecular articles surveyed that provided code (a) or data (b) for each of the periods studied (2015–2016: 157 articles, 2018–2019: 157 articles).

For articles stating to have used additional packages, 67.6% (N = 96) of articles published in journals without a code-sharing policy did not provide the version of all packages used (Figure 2e), whereas that percentage was 80.5% (N = 165) for articles published in journals with a code-sharing policy (Figure 2f). The mean number of packages used was 2.30 (median = 2.00, range = 1 to 10) in journals without a code-sharing policy and 2.41 (median = 2.00, range = 1 to 14) in journals with a code-sharing policy. The reporting of package versioning remained slightly higher for journals without a code-sharing policy than those with when expressed as the average percentage of software with version per article (without policy: median = 33.3%, mean = 45.1%, range: 0 to 100%; with policy: median = 0%, mean = 30.8%, range: 0 to 100%).

For articles stating the statistical software used, 23.5% (N = 65) of articles published in journals without a code-sharing policy used exclusively non-free (i.e., proprietary) software (Figure 2g) compared to 16.7% (N = 52) of articles published in journals with a code-sharing policy (Figure 2h).

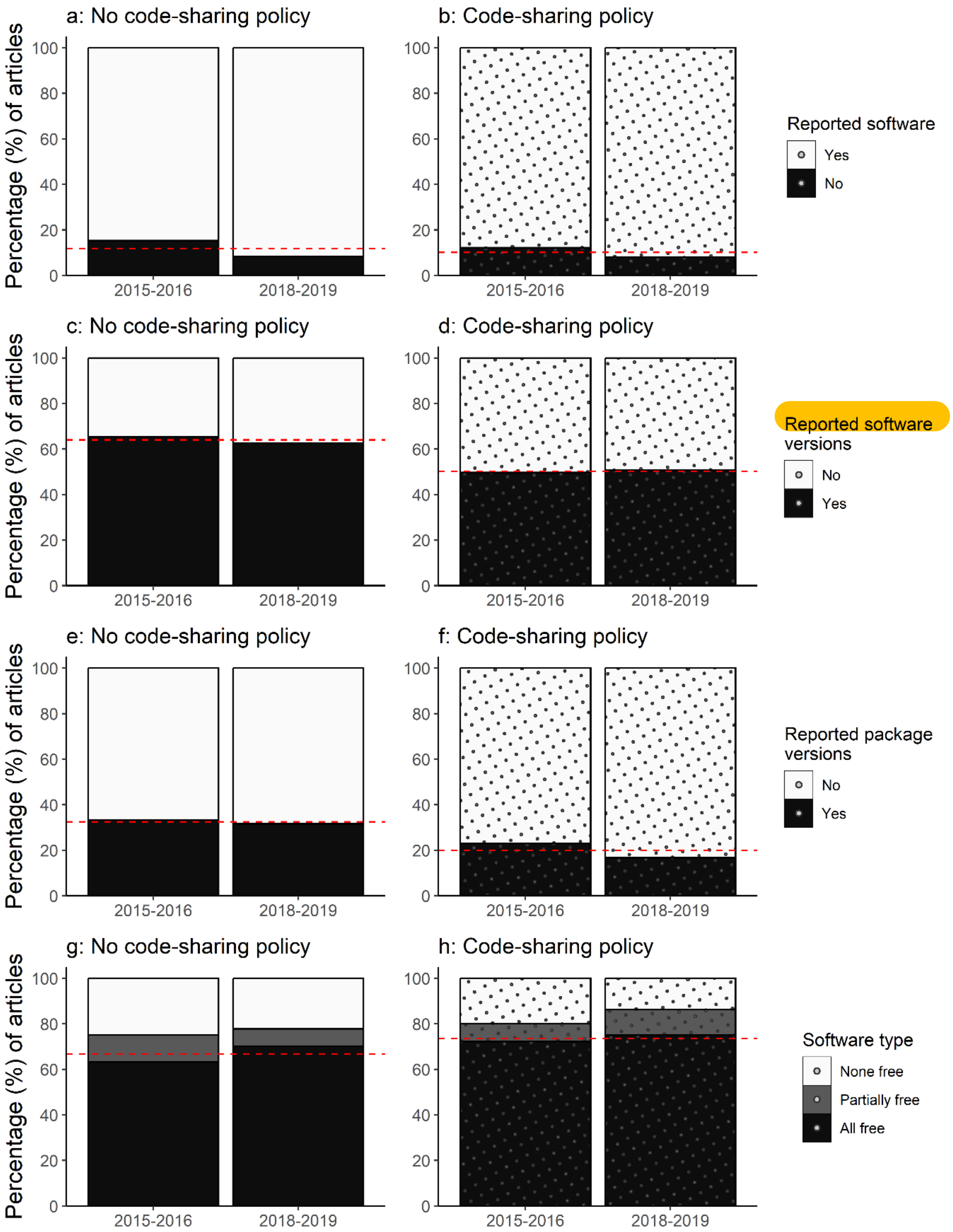

Figure 2. Features boosting long-term reproducibility in journals with (b, d, f, h; dotted fill) and without a code-sharing policy (a, c, e, g; non-dotted fill). The red dashed line corresponds to the mean for the category coloured in black (i.e., No software reported, Software versions reported, Package versions reported, and Free software, respectively).

Discussion

Our results show that code-sharing is almost non-existent (5%) for nonmolecular articles published in ecological journals without a code-sharing policy, a figure that is about six times lower than a comparative sample from journals with a code-sharing policy. Data availability faired better, with about one-third of articles published in ecological journals without a code-sharing policy sharing data, which corresponds to about half the rate observed in journals with a code-sharing policy. These low sharing rates lead to an extremely low reproducibility potential (less than 3%) of results published in journals without a code-sharing policy. Importantly, this is likely an overestimate since we also found that key reproducibility features (e.g., software name or versioning) are mostly lacking. Overall, our results confirm previous surveys in ecology and other fields: code-sharing is low, and simply implementing a code-sharing policy likely increases code-sharing, but not to the desired level. Below we place our results within and across fields, and discuss code-sharing and the importance of explicit policies. We also provide suggestions for journals on how to improve code-sharing and the (long-term) reproducibility of scientific findings (Box 1).

Open science practices are on the rise. When asked, most scientists agreed with the general open science norms and values decades ago (Anderson et al. 2007), but only recently we are starting to see more evidence of scientists not only agreeing but adhering to such norms and values. For example, a recent survey in the social sciences found that the percentage of scientists who self-reported to have used open science practices increased from 49% in 2010 to 87% in 2020 (Ferguson et al. 2023; see also Borycz et al. 2023). Meta-research studies have confirmed that several transparency indicators, including, but not limited to, data- and code-sharing are on the rise in ecology (Evans 2016; Culina et al. 2020; Roche et al. 2022a) and other fields (Heumüller et al. 2020; Cao et al. 2023; Colavizza et al. 2024; Sharma et al. 2024). Our current survey detected similar trends in ecological journals without a code-sharing policy, with code-sharing tripling from 2015-2016 (2.5%) to 2018-2019 (7.0%). Our results also support the observations from previous meta-research studies on authors being more likely to share data than code in ecology (Culina et al. 2020) and other fields (Bellomo et al. 2024; Sharma et al. 2024). Researchers may perceive greater risks and fewer benefits associated with sharing code compared to data, including unfamiliarity with best sharing practices, insecurity about code quality, fears of misuse or unsolicited appropriation of ideas, and excess preparation costs (Cadwallader & Hrynaszkiewicz, 2022; Gomes et al., 2022), coupled with the lack of incentives for code sharing. This discrepancy might also be in part due to journal policies often having a stronger emphasis on data- than code-sharing (Page et al. 2022; Ivimey-Cook et al. *In prep.*) and is likely less evident in (sub)disciplines that heavily rely on computational methods, such as computational biology and software engineering (Heumüller et al. 2020; Cadwallader et al. 2022).

Importantly, our results suggest that journals likely have a central role in increasing code-sharing rates: code-sharing was higher among nonmolecular articles published in journals with a code-sharing policy (27%) compared to those published in journals without a code-sharing policy (4.8%). A recent survey of meta-analyses in ecology and evolutionary biology detected similar patterns (21.2% and 9.1%, respectively, Kambouris et al. 2024). Though indirectly, previous studies have also suggested a link between the introduction of code-sharing policies and a subsequent increase in code availability. For example, code-sharing jumped from 53% in 2019 and 61% in 2020 to 87% in 2022 after the introduction of a mandatory code-sharing policy by PLOS Computational Biology (Cadwallader et al. 2022). Similarly, the percentage of initial submissions providing a link to data and/or code increased from 16.9% in 2021 to 42.6% in 2023 after Ecology Letters changed their sharing policy from simply providing a statement to mandating (and enforcing) providing a link to data and code (Ivimey-Cook et al. *In prep.*; for other examples, in ecology and beyond see Evans 2016; Hamilton et al. 2023; Ellis et al. 2024; Bellomo et al. 2024; Sperandii et al. 2024). Regardless of whether journals have a code-sharing policy or not, we also detected trends of increase in code availability over time. This is likely caused by other factors, such as changes in norms, better training, and better support in code-writing and sharing.

While having a policy helps in increasing code-sharing, it is certainly not enough without enforcing compliance (Culina et al. 2020). Our previous survey of 14 ecological journals with a code-sharing policy study indicated that the strictness of the policy did not affect code availability since the percentage of articles sharing code was similar between journals with encouraged (mean and range: 29.7% [14-50%], 3 journals), mandatory (23.0% [22-38%], 5 journals), and encouraged/mandatory (24.3% [7-53%], 6 journals) policies (Culina et al. 2020). A recent survey in biomedical research found more promising rates, with up to 50% of articles published in 8 journals with a code-sharing policy making code availably and the likelihood of code-sharing being double in journals with mandatory compared to encouraged policies (Sharma et al. 2024). Overall, despite low compliance, which has been linked to factors such as difficult-to-find or unclearly written sharing policies (Christian et al. 2020), these examples suggest that even under low policy enforcement, policy interventions can shift research practices towards greater openness. Indeed, implementing a code-sharing policy is a positive step forward even when the resources for enforcing such a policy and reviewing code (e.g., by adopting data and code editors) are not yet available.

We found that features boosting long-term reproducibility, such as using free software and reporting its version, were similar between journals with and without a code-sharing policy, which suggests that although code-sharing policies seem to increase code availability, they might not increase software reporting without being more explicit about best practices. We found that versions of the statistical software and packages were often missing, and about a tenth of the articles did not even state the software used. Reporting software and package versions is important for several reasons. First, they can help in understanding and solving technical issues related to software dependencies, which are one of the most often encountered factors hindering computational reproducibility (Laurinavichyute et al. 2022; Kellner et al. 2024; Samuel and Mietchen 2024). Different versions of software and/or packages can lead to inconsistencies in results and even to code rot, which occurs when code relies on specific versions of software or packages that are no longer available or have undergone significant changes (e.g., deprecated functions), rendering the code incompatible with current operating systems (Boettiger 2015; Laurinavichyute et al. 2022). Second, software reporting standards are key for computational reproducibility (i.e., obtaining the same results using the same input data and code) but also for analytical reproducibility (i.e., obtaining the same results writing *fresh code* using the provided written methodological descriptions when data but not code are available; Kambouris et al. 2024), and thus, should be prioritized by authors and journals alike (Box 1). Last, about one-fifth of articles used exclusively non-free (i.e., proprietary) software. Reproducibility is hindered when code relies on proprietary software that requires licenses or subscriptions. Proprietary software restricts access to its source code and is inaccessible to researchers who cannot afford it, ultimately limiting independent verification and building upon the original research (Ostermann and Granell 2016; Benureau and Rougier 2018; Konkol et al. 2019; Laurinavichyute et al. 2022). Ideally, the code used for a study should be peer-reviewed to ensure its completeness, reusability and reproducibility prior to manuscript acceptance (Ivimey-Cook et al. 2023). Before code review becomes a norm, authors, reviewers and editors should ensure that the minimum requirements needed for reproducibility are met, which can be facilitated by the use of checklists and by policies explicitly linking to best practices.

BOX 1. How can journals increase code availability? Here is a list of suggestions for journals sorted by the ease of implementation. For more information, journals should consider contacting the journal liaison officer of the Society for Open, Reliable and Transparent Ecology and Evolutionary Biology (SORTEE; https://www.sortee.org/).

- Introduce a code-sharing policy: this can range from simply mandating a code availability statement (Hamilton et al. 2023; Sharma et al. 2024) to encouraging or, in the best case, mandating code-sharing, ideally coupled with policy enforcement (Ivimey-Cook et al. *In prep.*). Policies should be clearly written, explicit and easy to find, and ideally shared among journals within and/or among publishers (Christian et al. 2020).

- Implement a reproducibility checklist: this should integrate a minimal list of code-sharing best practices such as the use of persistent identifiers like DOIs which ensure long-term accessibility and proper attribution (Gewin 2016; Trisovic et al. 2022) or ensuring all software and their versioning is provided. Journals should also offer clear guidelines (and support) for authors on how to share code, report software and adhere to reproducibility standards.

- Review and verify code: ask authors to share their code upon first submission to allow reviewers to have access and review the code. Encourage reviewers to use code (and data) during their reviews (Ivimey-Cook et al. *In prep.*). Consider officially integrating code review as part of the editorial process by adding data and code editors to ensure code functionality and adherence to standards (Krafczyk et al 2021).

Our study design has several limitations. Despite that we matched journals in time and from a seemingly representative list of ecological journals, journals with a code-sharing policy are more likely to have a data-sharing policy too (Ivimey-Cook et al. *In prep.*), which may increase code-sharing simply by increasing the visibility of sharing in general. However, it is fair to assume that some of the 12 journals without a code-sharing policy studied here did not have a data-sharing policy between 2015 and 2019, which may partially account for lower code-sharing as a by-product. In addition, the journals with and without a code-sharing policy may have differed in other transparency indicators or predictors of computational reproducibility such as the existence of reporting checklists, differences in prestige or the type of research published. Despite those potential limitations, our study adds to the mounting evidence that journal policies are an important stepping stone to increasing code availability. Finally, a potentially important factor for increasing data and code sharing not explored in our study is the funding source, with evidence suggesting that research funded by competitive grants tends to have higher code- and data-sharing rates, presumably due to those funding bodies often having mandates or strong recommendations for sharing as part of their grant conditions (Tan et al. 2024). Thus, here we not only call on journals to introduce code-sharing policies but also on funders to make a stronger push for mandating data- and code-sharing regardless of whether they currently have or not the mechanisms necessary to enforce those policies.

Conclusions

In sum, our study adds to the mounting evidence showing that code-sharing policies increase code availability, which ultimately increases the reproducibility potential of scientific findings. Specifically, our study suggests that based on code and data availability, computational reproducibility potential is about 8 times lower in ecological journals without a code-sharing policy (2.5%) compared to those with one (21%). Importantly, however, those should be considered ceiling values since we also found that software reporting needs improvement to allow reproducibility, and previous studies have found that open code (Obels et al. 2020; Laurinavichyute et al. 2022; Henderson et al. 2024) and data are often incomplete and difficult to use due to poor documentation (Roche et al. 2015; Roche et al. 2022b). The perceived costs and benefits of sharing code and data have been studied, dissected and discussed in detail elsewhere (Soeharjono and Roche 2021; Gomes et al. 2022; Borycz et al. 2023; Nguyen et al. 2023). Low sharing and reporting are key factors increasing research waste in ecology (Purgar et al. 2022) and other fields (Chalmers and Glasziou 2009), and as such, more efforts are needed to reduce research waste (see more suggestions in Grainger et al. 2020; Buxton et al. 2021; Purgar et al. 2024). Here, we particularly call on all journals and funders to introduce data- and code-sharing policies, even if they do not currently have the resources or mechanisms necessary to enforce them.

Authors contributions

Alfredo Sánchez-Tójar: conceptualisation (equal); data curation (lead); formal analysis (lead); investigation (equal); methodology (lead); project administration (equal); software (lead); supervision (equal); visualization (equal); writing – original draft (lead); writing – review and editing (equal). **Aya Bezine**: data curation (equal); investigation (equal); writing – review and editing (equal). **Marija Purgar**: data curation (equal); investigation (equal); validation (equal); writing – review and editing (equal). **Antica Culina**: conceptualisation (equal); data curation (equal); investigation (equal); project administration (equal); supervision (equal); visualization (equal); validation (equal); writing – original draft (equal); writing – review and editing (equal).

Acknowledgements

We would like to thank Ilona van den Berg and Simon Evans for their original contribution to generating the dataset of articles from journals with a code-sharing policy reanalyzed here.

Funding

This study was funded by the Corona-Unterstützung support program (Bielefeld University) granted to Alfredo Sánchez-Tójar, and by the Croatian Science Foundation under project number IP-2022-10-2872 to Antica Culina.

Competing interest

Alfredo Sánchez-Tójar, Marija Purgar and Antica Culina are officers at the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE). Aya Bezine declares no competing interests.

References

Allen, Alice, Peter J. Teuben, and P. Wesley Ryan. 2018. 'Schroedinger's Code: A Preliminary Study on Research Source Code Availability and Link Persistence in Astrophysics'. *The Astrophysical Journal Supplement Series* 236 (1): 10. https://doi.org/10.3847/1538-4365/aab764.

Anderson, Melissa S., Brian C. Martinson, and Raymond De Vries. 2007. 'Normative Dissonance in Science: Results from a National Survey of U.S. Scientists'. *Journal of Empirical Research on Human Research Ethics* 2 (4): 3–14. https://doi.org/10.1525/jer.2007.2.4.3.

Bellomo, Rosa Katia, Emmanuel A. Zavalis, and John P. A. Ioannidis. 2024. 'Assessment of Transparency Indicators in Space Medicine'. Edited by Tadej Debevec. *PLOS ONE* 19 (4): e0300701. https://doi.org/10.1371/journal.pone.0300701.

Benureau, Fabien C. Y., and Nicolas P. Rougier. 2018. 'Re-Run, Repeat, Reproduce, Reuse, Replicate: Transforming Code into Scientific Contributions'. *Frontiers in Neuroinformatics* 11 (January):69. https://doi.org/10.3389/fninf.2017.00069.

Boettiger, Carl. 2014. 'An Introduction to Docker for Reproducible Research'. arXiv. https://doi.org/10.48550/arXiv.1410.0846.

Borregaard, Michael Krabbe, and Edmund M. Hart. 2016. 'Towards a More Reproducible Ecology'. *Ecography* 39 (4): 349–53. https://doi.org/10.1111/ecog.02493.

Borycz, Joshua, Robert Olendorf, Alison Specht, Bruce Grant, Kevin Crowston, Carol Tenopir, Suzie Allard, Natalie M. Rice, Rachael Hu, and Robert J. Sandusky. 2023. 'Perceived Benefits of Open Data Are Improving but Scientists Still Lack Resources, Skills, and Rewards'. *Humanities and Social Sciences Communications* 10 (1): 339. https://doi.org/10.1057/s41599-023-01831-7.

Buxton, Rachel T., Elizabeth A. Nyboer, Karine E. Pigeon, Graham D. Raby, Trina Rytwinski, Austin J. Gallagher, Richard Schuster, et al. 2021. 'Avoiding Wasted Research Resources in Conservation Science'. *Conservation Science and Practice* 3 (2): e329. https://doi.org/10.1111/csp2.329.

Cadwallader, Lauren, and Iain Hrynaszkiewicz. 2022. 'A Survey of Researchers' Code Sharing and Code Reuse Practices, and Assessment of Interactive Notebook Prototypes'. *PeerJ* 10 (August):e13933. https://doi.org/10.7717/peerj.13933.

Cadwallader, Lauren, Feilim Mac Gabhann, Jason Papin, and Virginia E. Pitzer. 2022. 'Advancing Code Sharing in the Computational Biology Community'. *PLOS Computational Biology* 18 (6): e1010193. https://doi.org/10.1371/journal.pcbi.1010193.

Campbell, Tristan, Kingsley W. Dixon, and Rebecca N. Handcock. 2023. 'Restoration and Replication: A Case Study on the Value of Computational Reproducibility Assessment'. *Restoration Ecology* 31 (8): e13968. https://doi.org/10.1111/rec.13968.

Cao, Hancheng, Jesse Dodge, Kyle Lo, Daniel A McFarland, and Lucy Lu Wang. 2023. 'The Rise of Open Science: Tracking the Evolution and Perceived Value of Data and Methods Link-Sharing Practices'. arXiv. https://doi.org/10.48550/arXiv.2310.03193.

Chalmers, Iain, and Paul Glasziou. 2009. 'Avoidable Waste in the Production and Reporting of Research Evidence'. *The Lancet* 374 (9683): 86–89. https://doi.org/10.1016/S0140-6736(09)60329-9.

Christian, Thu-Mai, Amanda Gooch, Todd Vision, and Elizabeth Hull. 2020. 'Journal Data Policies: Exploring How the Understanding of Editors and Authors Corresponds to the Policies Themselves'. Edited by Cassidy Rose Sugimoto. *PLOS ONE* 15 (3): e0230281. https://doi.org/10.1371/journal.pone.0230281.

Colavizza, Giovanni, Lauren Cadwallader, Marcel LaFlamme, Grégory Dozot, Stéphane Lecorney, Daniel Rappo, and Iain Hrynaszkiewicz. 2024. 'An Analysis of the Effects of Sharing Research Data, Code, and Preprints on Citations'. Edited by Yongli Tang. *PLOS ONE* 19 (10): e0311493. https://doi.org/10.1371/journal.pone.0311493.

Cole, Nicki Lisa, Eva Kormann, Thomas Klebel, Simon Apartis, and Tony Ross-Hellauer. 2024. 'The Societal Impact of Open Science: A Scoping Review'. *Royal Society Open Science* 11 (6): 240286. https://doi.org/10.1098/rsos.240286.

Culina, Antica, Ilona Van Den Berg, Simon Evans, and Alfredo Sánchez-Tójar. 2020. 'Low Availability of Code in Ecology: A Call for Urgent Action'. *PLOS Biology* 18 (7): e3000763. https://doi.org/10.1371/journal.pbio.3000763.

Donoho, David L., Arian Maleki, Inam Ur Rahman, Morteza Shahram, and Victoria Stodden. 2009. 'Reproducible Research in Computational Harmonic Analysis'. *Computing in Science & Engineering* 11 (1): 8–18. https://doi.org/10.1109/MCSE.2009.15.

Ellis, David A., John Towse, Olivia Brown, Alicia Cork, Brittany I. Davidson, Sophie Devereux, Joanne Hinds, et al. 2024. 'Assessing Computational Reproducibility in Behavior Research Methods'. *Behavior Research Methods* 56 (8): 8745–60. https://doi.org/10.3758/s13428-024-02501-5.

Essawy, Bakinam T., Jonathan L. Goodall, Daniel Voce, Mohamed M. Morsy, Jeffrey M. Sadler, Young Don Choi, David G. Tarboton, and Tanu Malik. 2020. 'A Taxonomy for Reproducible and Replicable Research in Environmental Modelling'. *Environmental Modelling & Software* 134 (December):104753. https://doi.org/10.1016/j.envsoft.2020.104753.

European Organization For Nuclear Research. 2013. 'Zenodo'. Zenodo. 2013. https://www.zenodo.org/.

Evans, Simon Robin. 2016. 'Gauging the Purported Costs of Public Data Archiving for Long-Term Population Studies'. *PLOS Biology* 14 (4): e1002432. https://doi.org/10.1371/journal.pbio.1002432.

Feng, Xiao, Huijie Qiao, and Brian J Enquist. 2020. 'Doubling Demands in Programming Skills Call for Ecoinformatics Education'. *Frontiers in Ecology and the Environment* 18 (3): 123–24. https://doi.org/10.1002/fee.2179.

Ferguson, Joel, Rebecca Littman, Garret Christensen, Elizabeth Levy Paluck, Nicholas Swanson, Zenan Wang, Edward Miguel, David Birke, and John-Henry Pezzuto. 2023. ‘Survey of Open Science Practices and Attitudes in the Social Sciences’. *Nature Communications* 14 (1): 5401. https://doi.org/10.1038/s41467-023-41111-1.

Fišar, Miloš, Ben Greiner, Christoph Huber, Elena Katok, Ali I. Ozkes, and the Management Science Reproducibility Collaboration. 2024. ‘Reproducibility in *Management Science*’. *Management Science* 70 (3): 1343–56. https://doi.org/10.1287/mnsc.2023.03556.

Gewin, Virginia. 2016. ‘Data Sharing: An Open Mind on Open Data’. *Nature* 529 (7584): 117–19. https://doi.org/10.1038/nj7584-117a.

Gomes, Dylan G. E., Patrice Pottier, Robert Crystal-Ornelas, Emma J. Hudgins, Vivienne Foroughirad, Luna L. Sánchez-Reyes, Rachel Turba, et al. 2022. ‘Why Don’t We Share Data and Code? Perceived Barriers and Benefits to Public Archiving Practices’. *Proceedings of the Royal Society B: Biological Sciences* 289 (1987): 20221113. https://doi.org/10.1098/rspb.2022.1113.

Grainger, Matthew J., Friederike C. Bolam, Gavin B. Stewart, and Erlend B. Nilsen. 2020. ‘Evidence Synthesis for Tackling Research Waste’. *Nature Ecology & Evolution* 4 (4): 495–97. https://doi.org/10.1038/s41559-020-1141-6.

Grüning, Björn, John Chilton, Johannes Köster, Ryan Dale, Nicola Soranzo, Marius Van Den Beek, Jeremy Goecks, Rolf Backofen, Anton Nekrutenko, and James Taylor. 2018. ‘Practical Computational Reproducibility in the Life Sciences’. *Cell Systems* 6 (6): 631–35. https://doi.org/10.1016/j.cels.2018.03.014.

Hamilton, Daniel G., Kyungwan Hong, Hannah Fraser, Anisa Rowhani-Farid, Fiona Fidler, and Matthew J. Page. 2023. ‘Prevalence and Predictors of Data and Code Sharing in the Medical and Health Sciences: Systematic Review with Meta-Analysis of Individual Participant Data’. *BMJ* 382 (July):e075767. https://doi.org/10.1136/bmj-2023-075767.

Henderson, Alec S., Roslyn I. Hickson, Morgan Furlong, Emma S. McBryde, and Michael T. Meehan. 2024. ‘Reproducibility of COVID-Era Infectious Disease Models’. *Epidemics* 46 (March):100743. https://doi.org/10.1016/j.epidem.2024.100743.

Heumüller, Robert, Sebastian Nielebock, Jacob Krüger, and Frank Ortmeier. 2020. ‘Publish or Perish, but Do Not Forget Your Software Artifacts’. *Empirical Software Engineering* 25 (6): 4585–4616. https://doi.org/10.1007/s10664-020-09851-6.

Hinsen, Konrad. 2019. ‘Dealing With Software Collapse’. *Computing in Science & Engineering* 21 (3): 104–8. https://doi.org/10.1109/MCSE.2019.2900945.

Ivimey-Cook, Edward R., Joel L. Pick, Kevin R. Bairos-Novak, Antica Culina, Elliot Gould, Matthew Grainger, Benjamin M. Marshall, et al. 2023. ‘Implementing Code Review in the Scientific Workflow: Insights from Ecology and Evolutionary Biology’. *Journal of Evolutionary Biology* 36 (10): 1347–56. https://doi.org/10.1111/jeb.14230.

Ivimey-Cook, Edward, Alfredo Sánchez-Tójar, Antica Culina, Dominique Roche, Ilias Berberi, Rafaela A. Almeida, Bawan Amin, et al. *In preparation*. ‘From Unclear Policies to Practise: Slow Progress Towards Data- and Code-Sharing in Ecology and Evolution’. Preregistration available at: https://osf.io/zxurh.

Jay, Caroline, Robert Haines, and Daniel S. Katz. 2021. ‘Software Must Be Recognised as an Important Output of Scholarly Research’. *International Journal of Digital Curation* 16 (1): 6. https://doi.org/10.2218/ijdc.v16i1.745.

Jenkins, Gareth B., Andrew P. Beckerman, Céline Bellard, Ana Benítez-López, Aaron M. Ellison, Christopher G. Foote, Andrew L. Hufton, et al. 2023. ‘Reproducibility in Ecology and Evolution: Minimum Standards for Data and Code’. *Ecology and Evolution* 13 (5): e9961. https://doi.org/10.1002/ece3.9961.

Kambouris, Steven, David P. Wilkinson, Eden T. Smith, and Fiona Fidler. 2024. ‘Computationally Reproducing Results from Meta-Analyses in Ecology and Evolutionary Biology Using Shared Code and Data’. Edited by Elias Kaiser. *PLOS ONE* 19 (3): e0300333. https://doi.org/10.1371/journal.pone.0300333.

Kellner, Kenneth F., Jeffrey W. Doser, and Jerrold L. Belant. 2024. ‘Functional R Code Is Rare in Species Distribution and Abundance Papers’. *Ecology*, November, e4475. https://doi.org/10.1002/ecy.4475.

Kim, Albert Y., Valentine Herrmann, Ross Barreto, Brianna Calkins, Erika Gonzalez-Akre, Daniel J. Johnson, Jennifer A. Jordan, et al. 2022. ‘Implementing GitHub Actions Continuous Integration to Reduce Error Rates in Ecological Data Collection’. *Methods in Ecology and Evolution* 13 (11): 2572–85. https://doi.org/10.1111/2041-210X.13982.

Kohrs, Friederike E, Susann Auer, Alexandra Bannach-Brown, Susann Fiedler, Tamarinde Laura Haven, Verena Heise, Constance Holman, et al. 2023. ‘Eleven Strategies for Making Reproducible Research and Open Science Training the Norm at Research Institutions’. Edited by Mone Zaidi. *eLife* 12 (November):e89736. https://doi.org/10.7554/eLife.89736.

Konkol, Markus, Christian Kray, and Max Pfeiffer. 2019. ‘Computational Reproducibility in Geoscientific Papers: Insights from a Series of Studies with Geoscientists and a Reproduction Study’. *International Journal of Geographical Information Science* 33 (2): 408–29. https://doi.org/10.1080/13658816.2018.1508687.

Krafczyk, M. S., A. Shi, A. Bhaskar, D. Marinov, and V. Stodden. 2021. ‘Learning from Reproducing Computational Results: Introducing Three Principles and the *Reproduction Package*’. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379 (2197): rsta.2020.0069, 20200069. https://doi.org/10.1098/rsta.2020.0069.

Lai, Jiangshan, Christopher J. Lortie, Robert A. Muenchen, Jian Yang, and Keping Ma. 2019. ‘Evaluating the Popularity of R in Ecology’. *Ecosphere* 10 (1): e02567. https://doi.org/10.1002/ecs2.2567.

Lasser, Jana. 2020. ‘Creating an Executable Paper Is a Journey through Open Science’. *Communications Physics* 3 (1): 143. https://doi.org/10.1038/s42005-020-00403-4.

Laurinavichyute, Anna, Himanshu Yadav, and Shravan Vasishth. 2022. ‘Share the Code, Not Just the Data: A Case Study of the Reproducibility of Articles Published in the Journal of Memory and Language under the Open Data Policy’. *Journal of Memory and Language* 125 (August):104332. https://doi.org/10.1016/j.jml.2022.104332.

Lewis, Keith P., Eric Vander Wal, and David A. Fifield. 2018. ‘Wildlife Biology, Big Data, and Reproducible Research’. *Wildlife Society Bulletin* 42 (1): 172–79. https://doi.org/10.1002/wsb.847.

Maitner, Brian, Paul Efren Santos Andrade, Luna Lei, Jamie Kass, Hannah L. Owens, George C. G. Barbosa, Brad Boyle, et al. 2024. ‘Code Sharing in Ecology and Evolution Increases Citation Rates but Remains Uncommon’. *Ecology and Evolution* 14 (8): e70030. https://doi.org/10.1002/ece3.70030.

McKiernan, Erin C. 2017. ‘Imagining the “Open” University: Sharing Scholarship to Improve Research and Education’. *PLOS Biology* 15 (10): e1002614. https://doi.org/10.1371/journal.pbio.1002614.

Mislan, K.A.S., Jeffrey M. Heer, and Ethan P. White. 2016. ‘Elevating The Status of Code in Ecology’. *Trends in Ecology & Evolution* 31 (1): 4–7. https://doi.org/10.1016/j.tree.2015.11.006.

Nguyen, Phi-Yen, Joanne E. McKenzie, Daniel G. Hamilton, David Moher, Peter Tugwell, Fiona M. Fidler, Neal R. Haddaway, et al. 2023. ‘Systematic Reviewers’ Perspectives on Sharing Review Data, Analytic Code, and Other Materials: A Survey’. *Cochrane Evidence Synthesis and Methods* 1 (2): e12008. https://doi.org/10.1002/cesm.12008.

Obels, Pepijn, Daniël Lakens, Nicholas A. Coles, Jaroslav Gottfried, and Seth A. Green. 2020. ‘Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology’. *Advances in Methods and Practices in Psychological Science* 3 (2): 229–37. https://doi.org/10.1177/2515245920918872.

O’Dea, Rose E., Malgorzata Lagisz, Michael D. Jennions, Julia Koricheva, Daniel W.A. Noble, Timothy H. Parker, Jessica Gurevitch, et al. 2021. ‘Preferred Reporting Items for Systematic Reviews and Meta-analyses in Ecology and Evolutionary Biology: A PRISMA Extension’. *Biological Reviews* 96 (5): 1695–1722. https://doi.org/10.1111/brv.12721.

Ostermann, Frank O., and Carlos Granell. 2017. ‘Advancing Science with VGI: Reproducibility and Replicability of Recent Studies Using VGI’. *Transactions in GIS* 21 (2): 224–37. https://doi.org/10.1111/tgis.12195.

Ouzzani, Mourad, Hossam Hammady, Zbys Fedorowicz, and Ahmed Elmagarmid. 2016. ‘Rayyan—a Web and Mobile App for Systematic Reviews’. *Systematic Reviews* 5 (1): 210. https://doi.org/10.1186/s13643-016-0384-4.

Page, Matthew J., Phi-Yen Nguyen, Daniel G. Hamilton, Neal R. Haddaway, Raju Kanukula, David Moher, and Joanne E. McKenzie. 2022. ‘Data and Code Availability Statements in Systematic Reviews of Interventions Were Often Missing or Inaccurate: A Content Analysis’. *Journal of Clinical Epidemiology* 147 (July):1–10. https://doi.org/10.1016/j.jclinepi.2022.03.003.

Peng, Roger D. 2011. ‘Reproducible Research in Computational Science’. *Science* 334 (6060): 1226–27. https://doi.org/10.1126/science.1213847.

Purgar, Marija, Paul Glasziou, Tin Klanjscek, Shinichi Nakagawa, and Antica Culina. 2024. ‘Supporting Study Registration to Reduce Research Waste’. *Nature Ecology & Evolution* 8 (8): 1391–99. https://doi.org/10.1038/s41559-024-02433-5.

Purgar, Marija, Tin Klanjscek, and Antica Culina. 2022. ‘Quantifying Research Waste in Ecology’. *Nature Ecology & Evolution* 6 (9): 1390–97. https://doi.org/10.1038/s41559-022-01820-0.

R Core Team. 2023. ‘R: A Language and Environment for Statistical Computing.’ Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Roche, Dominique G., Graham D. Raby, Tommy Norin, Rasmus Ern, Hanna Scheuffele, Michael Skeeles, Rachael Morgan, et al. 2022a. ‘Paths towards Greater Consensus Building in Experimental Biology’. *Journal of Experimental Biology* 225 (Suppl_1): jeb243559. https://doi.org/10.1242/jeb.243559.

Roche, Dominique G., Ilias Berberi, Fares Dhane, Félix Lauzon, Sandrine Soeharjono, Roslyn Dakin, and Sandra A. Binning. 2022b. ‘Slow Improvement to the Archiving Quality of Open Datasets Shared by Researchers in Ecology and Evolution’. *Proceedings of the Royal Society B: Biological Sciences* 289 (1975): 20212780. https://doi.org/10.1098/rspb.2021.2780.

Roche, Dominique G., Loeske E. B. Kruuk, Robert Lanfear, and Sandra A. Binning. 2015. ‘Public Data Archiving in Ecology and Evolution: How Well Are We Doing?’ *PLOS Biology* 13 (11): e1002295. https://doi.org/10.1371/journal.pbio.1002295.

Samuel, Sheeba, and Daniel Mietchen. 2024. ‘Computational Reproducibility of Jupyter Notebooks from Biomedical Publications’. *GigaScience* 13 (January):giad113. https://doi.org/10.1093/gigascience/giad113.

Serghiou, Stylianos, Despina G. Contopoulos-Ioannidis, Kevin W. Boyack, Nico Riedel, Joshua D. Wallach, and John P. A. Ioannidis. 2021. ‘Assessment of Transparency Indicators across the Biomedical Literature: How Open Is Open?’ Edited by Lisa Bero. *PLOS Biology* 19 (3): e3001107. https://doi.org/10.1371/journal.pbio.3001107.

Sharma, Nitesh Kumar, Ram Ayyala, Dhrithi Deshpande, Yesha Patel, Viorel Munteanu, Dumitru Ciorba, Viorel Bostan, et al. 2024. ‘Analytical Code Sharing Practices in Biomedical Research’. *PeerJ Computer Science* 10 (June):e2066. https://doi.org/10.7717/peerj-cs.2066.

Soeharjono, Sandrine, and Dominique G Roche. 2021. ‘Reported Individual Costs and Benefits of Sharing Open Data among Canadian Academic Faculty in Ecology and Evolution’. *BioScience* 71 (7): 750–56. https://doi.org/10.1093/biosci/biab024.

Sperandii, Marta Gaia, Manuele Bazzichetto, Glenda Mendieta-Leiva, Sebastian Schmidtlein, Michael Bott, Renato A. Ferreira de Lima, Valério D. Pillar, Jodi N. Price, Viktoria Wagner, and Milan Chytrý. 2024. 'Towards More Reproducibility in Vegetation Research'. *Journal of Vegetation Science* 35 (1): e13224. https://doi.org/10.1111/jvs.13224.

Tan, Aidan C., Angela C. Webster, Sol Libesman, Zijing Yang, Rani R. Chand, Weber Liu, Talia Palacios, Kylie E. Hunter, and Anna Lene Seidler. 2024. 'Data Sharing Policies across Health Research Globally: CROSS -sectional Meta-research Study'. *Research Synthesis Methods* 15 (6): 1060–71. https://doi.org/10.1002/jrsm.1757.

Touchon, Justin C., and Michael W. McCoy. 2016. 'The Mismatch between Current Statistical Practice and Doctoral Training in Ecology'. *Ecosphere* 7 (8): e01394. https://doi.org/10.1002/ecs2.1394.

Trisovic, Ana, Matthew K. Lau, Thomas Pasquier, and Mercè Crosas. 2022. 'A Large-Scale Study on Research Code Quality and Execution'. *Scientific Data* 9 (1): 60. https://doi.org/10.1038/s41597-022-01143-6.

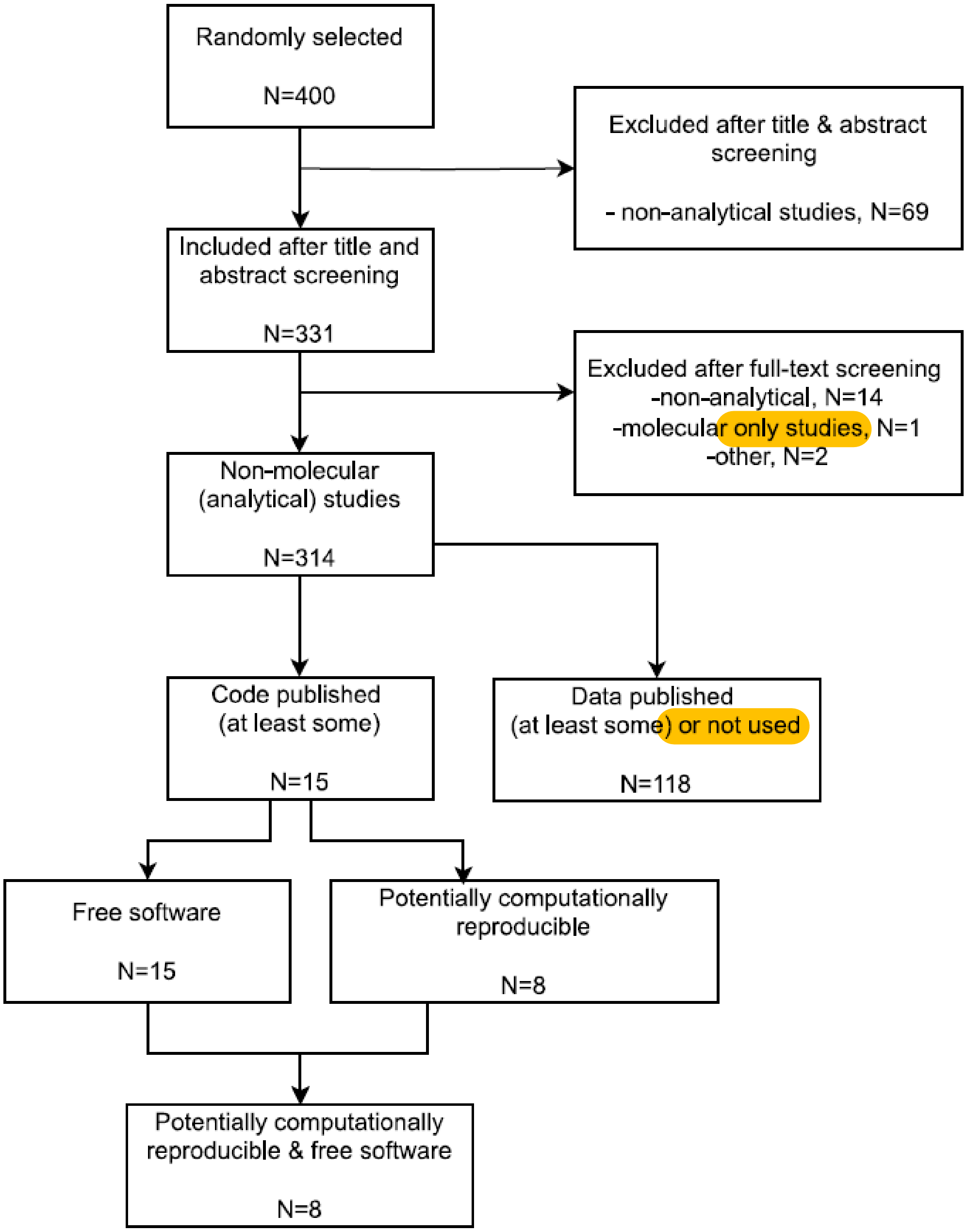

Figure S1. PRISMA diagram detailing the screening procedure and final number of articles included.