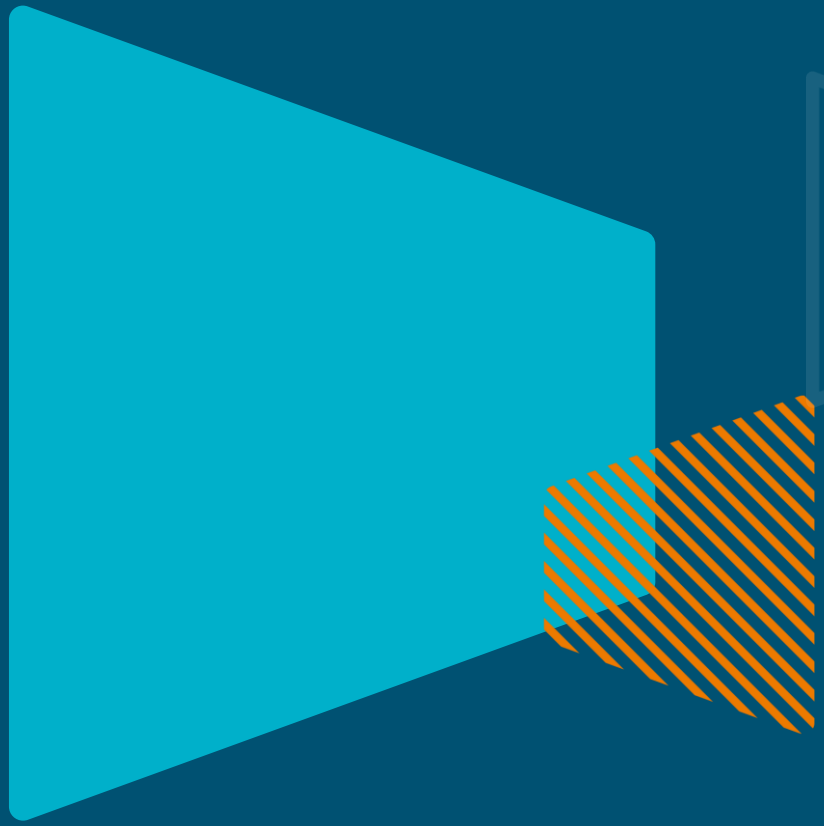


What librarians need to know about research integrity

Learn how to promote responsible research practices from experts in the field

SPRINGER NATURE GROUP

Speakers

Chris Graf, Research Integrity Director, Springer Nature

Allison Doerr, Chief Editor of *Nature Methods*, Nature Portfolio

Dominique Morneau-Brosnan, Chief Editor of *Nature Reviews Methods Primers*, Nature Portfolio

Research Integrity: Better together

Chris Graf
Research Integrity Director
Springer Nature

SPRINGER NATURE

Contents

Executive summary

What does research integrity mean to you? I'll ask you to consider that question, reflect on the complexities, and think about what kind of focus in research integrity might be most impactful.

Training in research integrity. I'll share two references and the evidence they provide for approaches to training in research integrity.

Resources from Springer Nature. I'll introduce two training resources from Springer Nature that you may find helpful as part of a comprehensive research integrity plan.

Closing thoughts. A research funder suggests all stakeholders have responsibility to ensure the publication system is conducive to (good) scholarship. Some things are better together.

Thank you!

What does research integrity mean to you?

SPRINGER NATURE

Research integrity

Enhancing
reproducibility

SPRINGER NATURE

8th World Conference on Research Integrity (Hybrid)

2-5 June 2024

Megaron Athens International Conference Centre (MAICC)

Athens - Greece

WCRI 2024: Abstract submission categories

https://wcri2024.org/abstracts-new/

Research environment. Healthy, supportive, inclusive workplace conducive to research integrity.

Training, supervision and mentoring.

Grant assessment, award, monitoring. Awarding grants and ensuring good grant governance.

Researcher assessment, evaluation, promotion. Researcher rewards and incentives.

Research procedures, materials and methods. How researchers design and carry out research.

Research ethics structures. Support for research ethics requirements.

Data practices and management. Data collection, storage, retention, archiving, and sharing.

Research collaboration.

Research publication and communication. Writing, reviewing, editing, and publishing.

Breaches of research integrity. Questionable practices, misconduct, breaches of integrity and its management.

Want research integrity? Stop the blame game

Helping every scientist to improve is more effective than ferreting out a few frauds.

Malcolm Macleod

Most scientists reading this probably assume that their research-integrity office has nothing to do with them. It deals with people who cheat, right? Well, it's not that simple: cheaters are relatively rare, but plenty of people produce imperfect, imprecise or uninterpretable results. If the quality of every scientist's work could be made just a little better, then the aggregate impact on research integrity would be enormous.

Nature 599, 533 (2021) https://doi.org/10.1038/d41586-021-03493-4

Training

SPRINGER NATURE

JOURNAL ARTICLE

Education and training policies for research integrity: Insights from a focus group study

Krishma Labib started her PhD at VUMC Amsterdam n January 2019. Her research is part of the European Commission Horizon 2020 funded Standard Operating Procedures for Research Integrity project SOPs4RI

> These results confirm the need for research institutions to develop a comprehensive RI plan that integrates RI education into the research endeavor

Labib, Evans, Roje et al. *Science and Public Policy*, 49;2:246–266 https://doi.org/10.1093/scipol/scab077

Needs and provision of research integrity training in Australian Institutions

Researchers value integrity training

Researchers value integrity training and would like to be offered more

> The results of the first national survey to investigate research integrity in Australia, a collaboration between the Australian Academy of Science and publisher Springer Nature, indicate broad support for mandatory research integrity training. The survey found that whilst 68% of respondents stated that their institution offered research integrity related training and 50% stated it was mandatory, 73% felt that such training should be mandatory for all those holding a research position.

Findings and data https://doi.org/10.6084/m9.figshare.19771759

Resources

SPRINGER NATURE

Authors

New: April 2023

Free

Research Integrity: An Introduction for Researchers

Research integrity is a key topic for everyone involved in science. However, it can present a bewildering array of topics, and early career researchers may receive little or no formal training in this area. How can you avoid common pitfalls and ensure your work is of the best possible standard? This course aims to give you an overview of the main areas in both research ethics and publication ethics.

We have designed this tutorial with early career researchers in mind, across all scholarly fields. Whether your work involves traditional lab work, field work or research that is literature or theory based, the principles of research ethics and publication ethics are still critical.

You will also have the opportunity to check your understanding with quiz questions as we go.

Access the course

English

Self-paced

45 minutes

Course contents

Research Integrity: An Introduction for Researchers

What is research integrity?

Conducting research with integrity

Study design and execution

Ethics approval

Informed consent

Trial registration

Animal research

Plants, geological samples, cell lines

Authorship

Data

Reporting guidelines

Avoiding plagiarism

Conflicts of interest

Citation manipulation

Duplicate submission

Predatory publishers

Post-publication changes

Paper mills

https://www.springernature.com/gp/authors/campaigns/research-integrity-course

SPRINGER NATURE

SPRINGER NATURE

Search EN

Editors | Editing a journal | Resources and Tools | Research Integrity | Courses

Free

< E-Learning courses

Course on Fundamentals of Peer Review

A strong understanding of all aspects of the peer review process is vital for all journal Editors. This free three-module course provides any Editor with an overview of the process in full.

Editor Resources

English

30 minutes per module

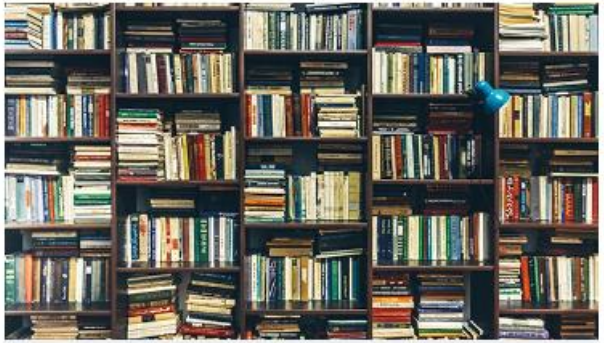
Module overview

Fundamentals of Peer Review: 3 modules that explain the basics of peer review.

Fundamentals of Peer Review 1 - Introduction to Peer Review

This module discusses the basics of the peer review process such as the different models, innovations in peer review, confidentiality and conflicts of interest.

Access the course ↗



Fundamentals of Peer Review 2 - Peer Reviewers

This module focuses on identifying and inviting reviewers, the characteristics of a good reviewer, author and submission fraud and reviewer fraud.

Access the course ↗

Fundamentals of Peer Review 3 – Reports and Decisions

This module covers the process of assessing the reports received, how a Journal Editor makes a decision and the indicators of good and bad peer review within a journal.

Access the course ↗

https://www.springernature.com/gp/editors/editor-courses/fundamentals-of-peer-review

Final thoughts

Alignment really matters

DFG: Academic Publishing as a Foundation and Area of Leverage for Research Assessment

> ensuring that the publication system develops in a way that is conducive to scholarship … enabling [academics] to avoid succumbing to misguided incentives

May 2022 https://doi.org/10.5281/zenodo.6538162

The German Research Foundation (DFG) awarded EUR3.6 billion funding in 2021

Thank you

SPRINGER NATURE

Allison Doerr
Chief Editor
Nature Methods

Promoting open science and research integrity at the Nature Research journals

Manuscript transparency is key to research integrity and reproducibility

What did you do and how did you do it?

→Methods section including unique materials, Supplementary Protocols, Reporting Summary, Code Availability Statement

What data supports your results?

→Data Availability Statement, Supplementary Information, accession codes

What previous work had been done?

→Introduction, References

Who did the work? Who funded the work? Do the authors stand to gain financially from publication?

→Author list, Acknowledgements, Ethics declaration

What are the limitations of the study?

→Discussion

Manuscript transparency is key to research integrity and reproducibility

What did you do and how did you do it?

→Methods reporting including unique materials, Supplementary Protocols, Reporting Summary, Code Availability Statement

What data supports your results?

→Data Availability Statement, Supplementary Information, accession codes

What previous work had been done?

→Introduction, References

Who did the work? Who funded the work? Do the authors stand to gain financially from publication?

→Author list, Acknowledgements, Ethics declaration

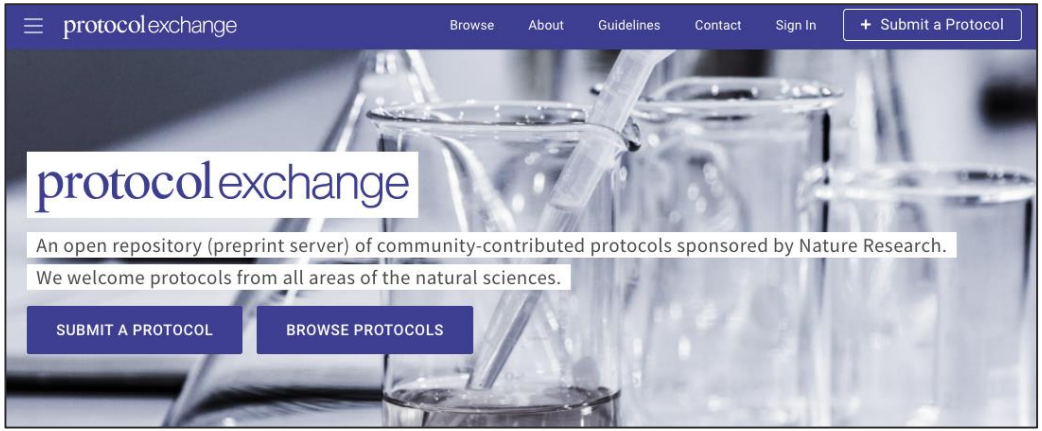
What are the limitations of the study?

→Discussion

Methods and protocols

Authors must provide detailed Methods sections

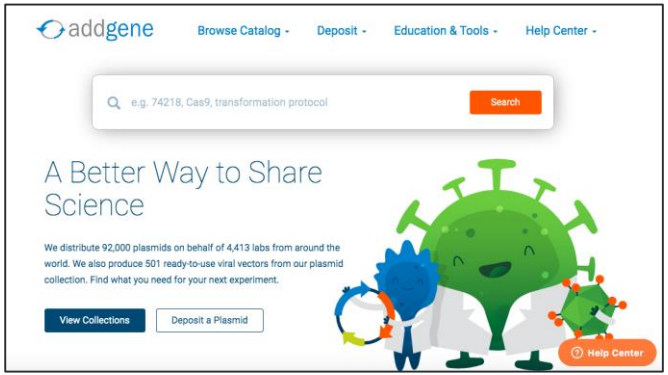
- No word count limits on Methods sections
- Supplementary Notes encouraged
- Step-by-step protocols encouraged, utilize protocol repositories, cite DOI in reference list



Unique materials

Authors must describe in full detail and agree to provide unique:

- Plasmids
- Antibodies
- Chemical compounds
- Cell lines
- Animal models



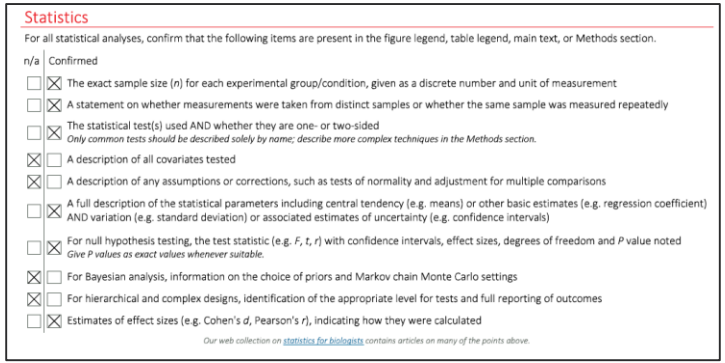
Nature Research journals require that:

- Plasmids, mutant strains and cell lines be deposited in public repositories and accession codes be provided
- Sources and catalog numbers for commercially available materials must be stated in the Methods
- Other unique reagents or materials be provided upon reasonable request for a reasonable fee

Nature Research Life Sciences Reporting Summary

Checklist

- Focuses on reporting basic, key elements to prevent them from being overlooked by authors and reviewers
- Experimental design
- Statistical information
- Unique reagents and materials
- Animal and human subject ethical guidelines
- Technique-specific modules (ChIP-seq, flow cytometry, MRI)
- Separate reporting tables for specific data types (crystallography, cryo-EM)
- Paper and checklist are evaluated by reviewers and in-house statistical experts

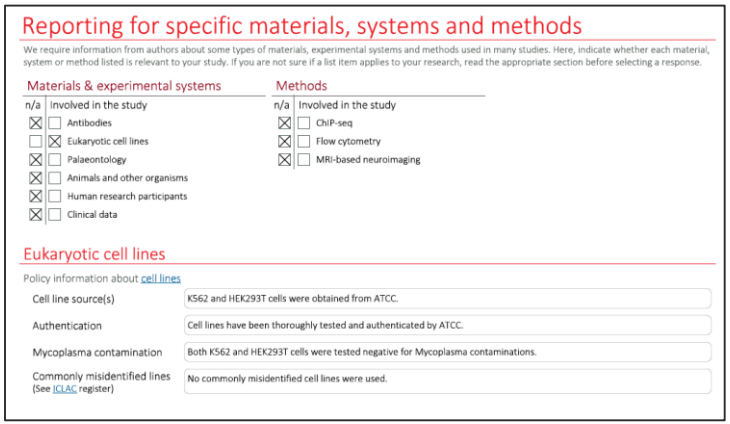


Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.



Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
☒	☐ Antibodies	☒	☐ ChIP-seq
☐	☒ Eukaryotic cell lines	☒	☐ Flow cytometry
☒	☐ Palaeontology	☒	☐ MRI-based neuroimaging
☒	☐ Animals and other organisms		
☒	☐ Human research participants		
☒	☐ Clinical data		

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	K562 and HEK293T cells were obtained from ATCC.
Authentication	Cell lines have been thoroughly tested and authenticated by ATCC.
Mycoplasma contamination	Both K562 and HEK293T cells were tested negative for Mycoplasma contaminations.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

SPRINGER NATURE GROUP

Code reporting

Most Nature Research journals require that:

- Code be made available if is central to the paper
- Code is provided for peer review, so that reviewers can test it
- Authors must fill in a code checklist

Depending on the paper and its claims, we might require all/some of the below:

- Mathematical description of the algorithm
- Source code
- Pseudocode
- Compiled software

Code Availability Statement

Authors and journals must take steps to ensure that code is reproducible, reusable, and remains available long-term!

Best practices dictate that authors:

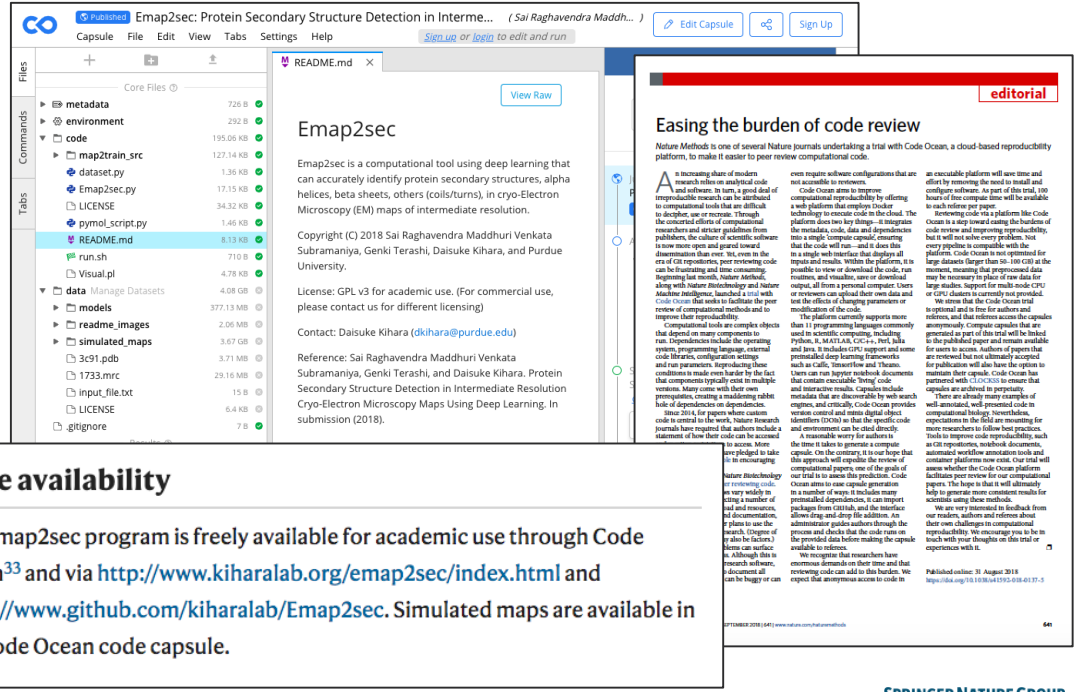
- Describe code availability and conditions of access in a Code Availability Statement
- Provide code via an established repository such as Github
- Mint a DOI (such as by Zenodo) and include it in the reference list
- Use versioning and continue to make the version used to generate results in the paper available
- Provide clear documentation for installation and use
- Provide a software license (ideally open source – any restrictions must be stated)
- Include sample data for others to test-run the software

Code Ocean

Code Ocean enables researchers to create a containerized version of their software, called a ‘compute capsule’

Benefits:

- All components required to re-run code are included in the compute capsule
- Helps authors comply with journal requirements
- Helps streamline code review process
- Provides easy and indefinite (digitally preserved) access to software



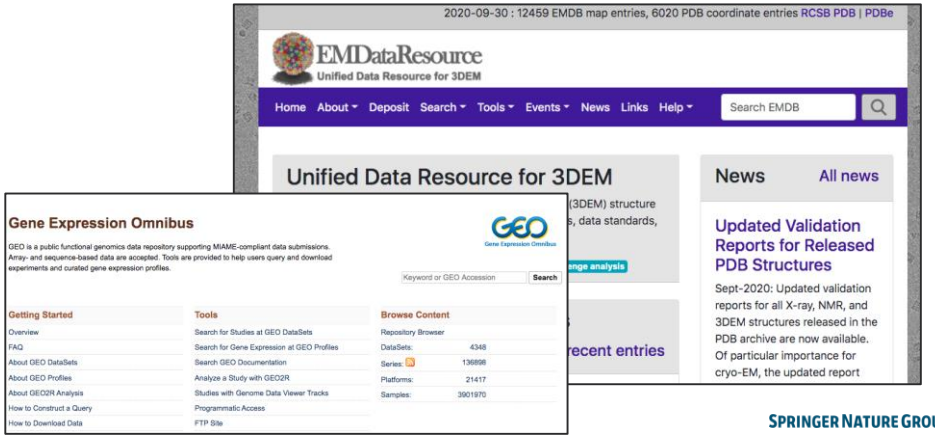
Data reporting

Nature Research promotes data sharing and data citation

- Many types of data are mandated to be shared via established repositories
- Authors encouraged to share non-mandated data via a repository if available
- Citation of dataset DOI in reference list is strongly encouraged
- Source data underlying graphs, gels/blots strongly encouraged

Best practices in data representation

- Avoid bar graphs
- Show full data distribution
- Avoid red/green contrast
- Show scale bars



Data Availability Statement

- How the data supporting the results reported in the article can be accessed
- Links to publicly available datasets that are analyzed/generated during study
- Details of source data
- States any restrictions on access
- We discourage “Data available upon request” statements

Data availability

All genomic data generated for this study are publicly available on the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE145695.

In Fig. 1, Scc1-calibrated ChIP–seq tracks from Hu et al.[22] were used for the cohesin pile-up SisterC heatmaps and ChIP–seq tracks. This dataset is available on GEO under accession number GSM1712309. The peaks were called on this dataset using MACS2. The pairwise cohesin interactions were compiled by listing all possible pairwise combinations of cohesin peak sites in the same chromosome, followed by separation based on distance between cohesin pairs (smaller than 10 kb, 10 to 20 kb, 20 to 35 kb and 35 to 50 kb). The cohesin sites in a 50-kb window around the chrIV were removed from the dataset. decay, Hi-C samples from cdc45 mutant cells his dataset is available on GEO under s data was processed identically to the Hi-C e sites of origin of replication were w.oridb.org/)[44]. Source data are provided with

Data availability

The human K562 XL-MS raw files (122 raw files (97 HILIC and 25 SCX fractions) from our recent proteome-wide human K562 XL-MS study[2]) analyzed in this study have been deposited to the ProteomeXchange Consortium via the PRIDE[40] partner repository with the dataset identifier PXD018771. Raw data from our PCA experiments are available from the corresponding author upon request. Protein sequences were obtained from the Uniprot database (https://www.uniprot.org/). Residue-level mapping was performed using data from the SIFTS database (https://www.ebi.ac.uk/pdbe/docs/sifts/index.html). Protein three-dimensional structures utilized in this study were obtained from the PDB (accession codes: 5GJQ, 1EUC, 1T9G, 5LNK, 1ZOY, 1NTM, 1V54, 5MY1, 5ADY, 5ME0, 2RDO, 2VRH, 4JK2, 4YLN, 4YLO, 4XO2, 4YFH and 4YF0). Source data are provided with this paper.

Transparency in peer review at Nature Research journals

Transparent peer review

- Authors choose whether to publish anonymized reviewer reports with the paper

Reviewer recognition

- Reviewers choose whether to be named on the published paper

Signed reviewer reports

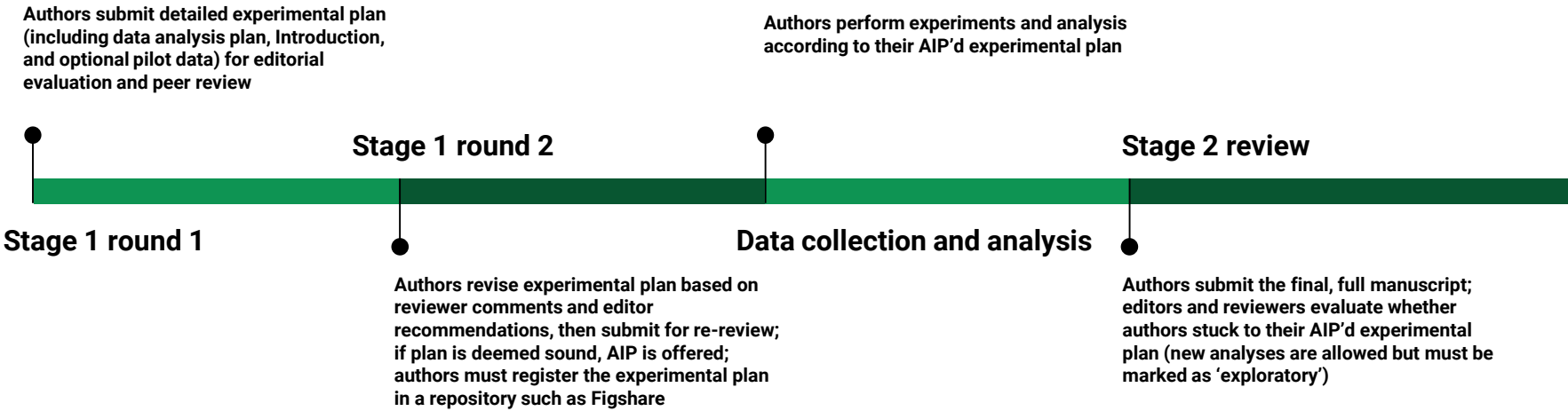
- Reviewers may sign their reports to reveal themselves to the authors (and potentially readers, if the authors select transparent peer review)

Double blind peer review

- Both authors and reviewers are anonymous

Registered Reports: a new format supporting research integrity

Registered Reports offer a format (being increasingly adopted by Nature Research journals) to improve the integrity of research by shifting peer review to the research plan, rather than the results.



Editorials: communicating with authors and readers

editorial

How editors edit

We shed some light on how the *Nature Methods* editorial team evaluates papers submitted to the journal.

The life of a professional scientific journal editor is exciting, challenging, and intellectually stimulating, but it requires a thick skin. We've been praised by some authors for our behind-the-scenes work, but we've also been called "paper pushers"—and worse. At conferences we are often approached by researchers curious about what it is that we do. We also receive comments such as, "Do you realize just how much power you have?" Editors should be viewed not as obstacles to publication, but as partners with the research community, tasked with curating, improving, and disseminating important, interesting, and high-quality work. Yet as we demand more transparency from our authors, we appreciate that we must also provide more insight into our own editorial processes, and so here we outline how the *Nature Methods* editorial team evaluates submitted papers.

Nature Methods has no external editorial board; all decisions are made by the team of in-house, full-time, PhD editors. To keep abreast of current trends and challenges, we travel to multiple conferences each year, visit institutes, and invite researchers to the office. We closely follow community cultures in enforcing standards for reporting and data deposition. While we don't discuss specific papers under consideration, we do share general scientific information with other editors at Nature Research. In short, we spend our days entrenched in the fields that we cover. But because we are not active scientists, we have no competing scientific or financial interests that might bias us toward or against a manuscript.

When a paper is submitted to *Nature Methods*, the chief editor assigns it to the editor with the most relevant expertise. The editor reads (yes, we read every paper that is submitted, and not just the abstract!), summarizes, and discusses the paper with at least one other editor with related expertise. Together, they make a decision to reject the paper or send it for peer review, and the lead editor communicates that decision to the author (usually within one week of submission).

We look for papers that demonstrate novel approaches to challenging and high-interest problems in basic biology research, enabling researchers to explore new corners of biology, or that solve nagging technical challenges. We exclude work that does not fall within this scope. Besides scope, novelty, potential interest, and practical value, we look at the meat of the paper: have the authors justified the advance over previous methods? Have they appropriately validated the method's performance? Have they done experiments to showcase cool new applications? If a paper is lacking in one of these areas, this does not necessarily spell its end; we may review a promising paper that needs more work, or we might suggest further experiments that could make it a stronger candidate for peer review. However, we send only about 12% of submitted papers out for peer review. This means that we reject many good-quality papers that just don't quite reach our bar, but we often recommend transfers to other journals in the Nature Research family that might be more suitable.

Many of the papers that we consider are multidisciplinary, so we solicit opinions not only from technical experts, but also from biologists who represent 'end users' of the technology. Authors' requests to exclude reviewers are always honored, as long as the list is reasonable. We are continually exploring new ways to improve peer review, such as our pilot project with Code Ocean to facilitate the review of code. We allow reviewers to reveal their identities to authors, should they so choose, and we also offer double-blind peer review as an author choice.

When a paper comes back from peer review, the editor summarizes the reviewers' comments and discusses her or his recommendation for revision, rejection with potential for resubmission, or complete rejection with at least one other team member. The chief editor is also involved in any recommendations for revision. Our decisions are never based on reviewer 'votes'. We read the reviewers' comments in light of our initial assessment of interest, novelty, validation, and application. We evaluate which concerns are crucial to address and which are not, and communicate this to the authors. Occasionally, we ask reviewers for further input on technical issues, or we may ask authors for a revision plan prior to making a decision. We strive to be open; if our expectations for a revision are unclear, we welcome further discussion by e-mail or phone.

We realize that reviewers can occasionally make mistakes or show signs of bias, and authors can address serious technical concerns that prompted us to reject an otherwise interesting paper. Therefore, we do consider appeals of rejected papers, though authors should very carefully consider the reasons the editor has stated for rejection and whether they can adequately address these concerns.

In the best-case scenario, the process from submission to acceptance will take just a few months. Occasionally papers go through multiple revisions over the course of a year or more, though we try to avoid more than two rounds of peer review. About half of the papers initially sent for peer review are published in *Nature Methods*. In the final stage just prior to acceptance, the editor carefully reads the paper and suggests wording changes to help clarify and focus the message or tone down claims that are not supported by the data, as well as ensures that our standards for reporting and for software, data, and materials availability are enforced.

To help recognize the dedication that the team puts into editing each paper, and to promote further transparency in our editorial processes, we will soon begin adding the editors' names to papers that we publish in *Nature Methods*.

We also do more than edit papers. We highlight interesting research published both in our own pages and in other journals with Research Highlights, News & Views, and Technology Features, as well as via social media. We aim to educate about and raise awareness of scientific and publishing issues through our Editorials and columns such as the Points of Significance. We commission Reviews and Perspectives of broad utility for scientists.

At the heart of our profession is an obligation to provide a useful service to the research community; we welcome feedback about how we can further improve. We take our roles with great responsibility and are proud to support and promote exciting and high-quality research. ❐

Published online: 30 January 2019
https://doi.org/10.1038/s41592-019-0324-z

editorial

The method comes first

A new method should be thoroughly tested, applied, described — and peer-reviewed — before biological discoveries generated using the method are published.

Which comes first, the method or the result? We think that most of our readers would agree that this is definitely not a 'chicken-or-the-egg' conundrum. It stands to reason that a new method should be carefully and thoroughly characterized and benchmarked — and its full description and these results peer-reviewed — before biological findings generated using this new method can be fully trusted.

As editors of a methods journal, we have observed many instances where this ideal chain of events has not been followed. Certainly it is not surprising that researchers who have discovered something novel and exciting using their new method would prioritize publishing these findings, especially if there is competition from other groups. Further, two groups may collaborate, one developing a method and the other applying the method to a biological question; these groups will have different priorities and may have papers ready for journal submission at different times.

Though we are aware of and sympathetic to these types of situations, we argue that publishing new biological findings generated using a novel method before the methods paper is accepted for publication in a peer-reviewed journal is detrimental to research.

In the most egregious examples, authors of a findings paper that uses an unpublished method or software tool will provide no details about the method and simply cite "manuscript in preparation." When reading a paper that has been peer-reviewed and undergone various editorial checks at a journal, a reader should be reasonably able to trust the results. But when the results hinge on a method that has not yet been vetted through peer review and communicated via publication, how can such findings be trusted? Even more worrying, how can the biological findings be reproduced by others? We urge peer reviewers to be on the lookout for this poor practice and flag it to the journal editor handling the paper.

Preprint servers allow authors to rapidly share unpublished work to the scientific community, something that we both support and encourage here at Nature Research. However, we argue that it is insufficient to cite a preprint reporting a method as evidence that the method has been properly validated. Our colleagues at *Nature Biotechnology*, for example, require that methods central to new results in a submitted manuscript be accepted for publication in a peer-reviewed journal before they will publish the manuscript, a stance we applaud. As they wrote in a 2017 Editorial, "peer-reviewed journals must ensure that the integration of minimally reviewed preprints into their papers does not compromise the reproducibility of the science they publish."

We strongly encourage researchers who want to publish two papers, one reporting a new method and the other a new finding, to prioritize writing up both. If it is not practical to publish the methods paper in a journal before submitting the findings paper, submission should at least be done concurrently. If both papers are submitted to the same journal, or to the same publisher, peer review and publication can often be coordinated. If the papers are submitted to different journals, the other paper should be provided to the editors (note that this is a requirement at *Nature Methods*). This allows the editors and the reviewers to understand how the method works and also to judge whether there is substantial overlap between the papers.

Even in cases where a methods and a findings paper have been simultaneously submitted to journals, peer review outcomes can be unpredictable. We advise authors to keep their editor informed about the status of the other paper and try to ensure that the methods paper is at least provisionally accepted (if not published) before the findings paper is published.

Authors should also be aware that if they describe a method in some detail in a paper where they report new biological findings, this may prevent them later being able to publish a dedicated methods paper in a journal (such as *Nature Methods*) where methodological novelty is an important editorial criterion. If we think a method is sufficiently exciting and important for us to potentially publish a paper focused on the method itself, we will occasionally consider it. But in such cases the methods paper must stand on its own: it must describe a new tool or an optimized workflow, or provide substantial additional characterization or validation data, or describe a novel application. In other words, there must be a good reason to justify publishing a dedicated methods paper following the initial report.

There are many examples of methods, tools and resources that have remained unpublished even for years. You might ask: why bother publishing a dedicated method paper at all? Methods are key to advancing scientific progress, and it's just as important for the method as for a novel finding, if not even more important, that the work go through a careful vetting process. At *Nature Methods*, we also uphold strict editorial standards regarding a method or tool's description (including making software code and unique materials available), its characterization and benchmarking in comparison to existing approaches (including making these data available), and a demonstration of general applicability. We think that these standards help improve the reliability and reproducibility of methods we publish, allowing readers to better trust new biological findings generated by such methods, as well as making the methods themselves more useful and practical for a broader audience. ❐

Published online: 30 November 2020
https://doi.org/10.1038/s41592-020-01017-y

editorial

What makes a *Nature Methods* paper

We explain what our editorial team looks for when considering a methods paper for publication.

Our editorial team members are often asked by potential authors about what we are looking for in a paper. Occasionally, authors express frustration that the editorial 'triage' or 'desk reject' stage, as it is referred to, feels subject to the whims of an editor or perhaps even their mood that day. However, we can assure you that we have long implemented a robust, multifactor editorial process for selecting papers to send out for peer review.

When a new manuscript arrives in our submission system, the chief editor assigns it to the team member whose expertise most closely matches the paper subject. More information about our editors and their scientific backgrounds is available on our website; we also invite you to read a summary of how we handle papers in our February 2019 Editorial.

No matter the field of research, there are some common elements that we always look for when judging methods paper submissions. To pass the editorial triage stage, a paper generally needs to check the boxes described below. (Note that here we focus on our Article and Brief Communication content types; other content types we publish are described in our August 2020 Editorial.)

Scope

Nature Methods' mission is to champion method and tool development research in the basic life sciences. Therefore, we consider methods papers with a primarily clinical, diagnostic or therapeutic focus to be out of scope. Methods papers in other fields such as chemistry or physics are also out of scope — unless the authors can make a strong case as to why the paper will make an impact on a broad life sciences community. If uncertain about whether a paper fits our scope, you are welcome to submit a presubmission inquiry via our submission system.

Interest

Though our broad scope covers all of the basic life sciences, there are certain major fields where we are particularly interested in publishing papers — for example, single cell analysis, genomics and transcriptomics, microscopy and imaging, structural biology, proteomics, metabolomics, genome engineering, stem cell biology, neuroscience and immunology. Within these fields, there are particular areas that represent the frontiers of methods development where we are most keen to receive papers. To learn more about what areas are piquing our interest, we invite you to read our recent Methods to Watch features, published in every January issue.

Novelty

Novelty is a key element of a *Nature Methods* paper. In a paper's introduction, we look for a clear explanation as to why a method or tool is a substantial advance over the state of the art. We also assess the paper in the context of the peer-reviewed literature to be confident that elements of the method or a very similar approach have not been previously reported by the authors or by another lab. That is not to say that we will not consider a strong paper in an area where there are already other published similar methods — we weigh the timeliness of the topic, the practical value of the method, the performance compared to other approaches, and whether the paper makes a strong case that the method will enable new applications (more on this below). We aim to publish a mix of papers with high conceptual novelty and high immediate practical value.

Method description

We want to publish methods that will be useable by others. It should therefore go without saying that method must be described in detail. The focus on the paper should be on the method and its characterization, not on new biological findings obtained. To enable method reuse, we often require that authors describing a complex experimental workflow provide a step-by-step protocol as a supplementary item (or better yet, deposit it in a protocol repository); authors must also describe their plan to distribute any unique experimental materials. To ensure that software tools are useable, we require a detailed description of the underlying algorithms, the code (ideally hosted in a code repository and assigned a DOI), a license and a user guide.

Validation and benchmarking

Strong validation of a method's performance is an essential ingredient of a *Nature Methods* paper. Whether an experimental approach or computational tool, authors should always take care to follow established field standards for best practices when validating its performance.

Experimental methods should be applied to at least one well-characterized system to demonstrate that the method produces expected results. Computational tools should be validated on a ground truth or gold standard dataset if available in the field. Simulated datasets, ideally with noise added to make the data more realistic, are also useful for validation, but we nearly always also want to see tests on real experimental datasets.

If similar methods have already been published, we also expect to see some performance benchmarking. This process can be somewhat fraught, as authors do not always have the technical expertise or access required to utilize different technologies or may not be sufficiently knowledgeable about the ins and outs of running another group's software. We rely on our expert reviewers to help us judge whether the new method has been appropriately compared to existing methods.

General applicability

We aim to publish methods that will be broadly applicable to life science researchers. Methods that are limited to studying a particular biological process may be too narrow in focus for our journal to consider. We also typically want to see data showing that good method performance is not just a one-off for a well-behaved system, but that it performs well with a diversity of systems or datasets. Just how many examples need to be shown to prove general applicability is field dependent, but two distinct applications is typically the minimum.

Challenging demonstration

Though our editorial focus is on the method itself, a 'killer application' can go a long way in showing readers why they should care about a method and perhaps consider adopting it in their own research. However, we are flexible about this – not every paper we publish has a killer application; it really depends on how well the paper has checked the other boxes. New biological results are not required for publication in *Nature Methods* — though it certainly doesn't hurt. Preliminary biological findings are often okay by us as long as conclusions are not overhyped and limitations are stated.

Editorials: communicating with authors and readers

Check for updates

editorial

What makes an author

Constructing a fair and accurate author list can be one of the most fraught aspects of manuscript publication. We provide some advice and resources for authors at all career levels.

The acknowledgement of scientific contributions in the form of manuscript authorship is vital at all stages of a researcher's career, from the well-established principal investigator applying for million-dollar grants to the undergraduate student applying to PhD programs. It's essential that authorship lists are constructed with utmost care.

The variety of authorship practices across the scientific literature, however, is vast. Different fields, different countries, even different labs have different norms. Some practices are troubling: lab technicians not included for their major contributions to a study because they are not on an academic track; contributors removed from author lists due to personal disputes; researchers who have not substantially contributed added to papers (in a misguided attempt to increase 'impact') without their consent; senior scientists taking advantage of power imbalance to undeservedly gain publications.

Even researchers with the best intentions can struggle with finalizing a fair and accurate author list. Here, we provide some best practice guidelines and explain how *Nature Methods* handles authorship issues.

First of all, community guidelines for authorship are available. Nature Portfolio's authorship policies are based on guidelines developed by McNutt et al. (*Proc. Natl Acad. Sci. USA* **115**, 2557–2560, 2018). Other guidelines in common use include those from the International Committee of Medical Journal Editors. As defined by Nature Portfolio, an author listed on a paper should have made a substantial contribution to the design of the work, the collection or analysis of data, the creation of a software tool, or the writing of the paper. This policy is meant to be broad and flexible, leaving "substantial contribution" up for quite a bit of room for interpretation.

In our view, job title or rank should never exclude a potential author. The lab technician or core facility scientist who developed a custom experimental workflow for the study should be included as an author. The first-year rotation student who spent several weeks collecting data should be included as an author. The software engineer who made substantial developments to an existing algorithm to analyze the data should be included as an author.

That said, not just any kind of assistance justifies authorship. People who provided routine services or basic technical help, contributed resources (such as by giving plasmids), proofread the manuscript, or gave general advice but did not otherwise significantly contribute to the scientific content of the paper should be thanked in the Acknowledgements. If previously published datasets or software tools are utilized in a new study without further development from their generators, there is no need to name them as authors. Even the person who secured funding need not necessarily be an author on a paper—they too ought to have scientifically contributed in a meaningful way. This is almost unheard of in lab-based science, where a principal investigator typically supervises the design of experiments and analysis of the resulting data, but it's relatively common practice in, for example, computer science for grad students to publish sole-author papers.

Though different research fields have different traditions, the custom in life sciences research is to name the person or people who did the bulk of the research first, followed by other contributors in descending order of the significance of their contributions, with the principal investigator(s) named at the very end of the list. Disputes often arise over who is named first on a paper. Most journals allow co-first-authorship designations to recognize cases of equal contribution, but one name must necessarily come first; the research community should take care to recognize these equal contributions. Those listed second should not feel that their contributions are minimized in any way.

Project managers should make defining authorship and authorship order a priority of a new study. Students and postdocs, collaborators, and service providers should speak up if authorship is not discussed early on. Setting clear parameters and communicating openly from the outset of a research study—in some cases even by signing formal authorship agreements—can go a long way toward preventing disputes and hurt feelings down the line.

All authors on a paper have a responsibility for at least part of its content. Nature Portfolio journals require author contribution statements, which in our view are crucial to clarify each author's role and responsibilities, to assign credit where it is due, to discourage the practice of including authors who did not significantly contribute to the study, and to assign accountability in (rare) cases of misconduct. The corresponding author, the main point of contact with a journal, has extra responsibilities. They are tasked with communicating with all coauthors at the submission, revision and final acceptance stages, including ensuring that all are satisfied with the manuscript text and content. The corresponding author must also check that all coauthors agree with changes to the author list, that any competing interests are declared, and that the paper complies with all of the journal's policies regarding data, materials and code sharing. Note that the journal corresponding author need not be the same person as the corresponding author(s) listed on the published paper, who take responsibility for post-publication inquiries.

We encourage our authors to speak up to let us know when best practices for authorship are not being followed. However, our editorial power is limited to delaying review or publication until disputes can be resolved, making corrections to papers, adding an 'editorial expression of concern' or, in very rare cases, retracting a paper. We rely on authors to behave responsibly and we cannot investigate or adjudicate authorship disputes. We advise those embroiled in disputes to seek help from their department head, university or other employer. We also recommend speaking to an experienced neutral party familiar with the study for advice—it's human nature to often overestimate our own contributions, but it's right to speak up about unfair treatment.

Unfortunately, we do not have the space to cover all possible authorship scenarios in this short piece. We look forward to answering your questions and perhaps sparking some lively discussion on Twitter, where you can follow us at @naturemethods. ❐

Published online: 3 September 2021
https://doi.org/10.1038/s41592-021-01271-8

Editorial

https://doi.org/10.1038/s41592-023-01865-4

Data sharing is the future

Check for updates

We examine data sharing practices and explore possible future directions.

In late 2022, the US government mandated open-access publication of scholarly research and free and immediate sharing of data underlying those publications for federally funded research beginning no later than 2025. For some fields the necessary standards and infrastructure are largely in place to support these policies. For others, however, many questions remain as to how these mandates can best be met.

In this issue, we feature a Correspondence from Richard Sever that was inspired by the government mandate and the increasing demand for open science. In it, he raises important topics, including deciding which data must be shared, standardizing file formats and developing community guidelines. He also calls for a "federated system of repositories with functionality tailored to the information that they archive," to meet the needs of many distinct fields.

Nature Portfolio journals have several data deposition requirements. These largely cover fields where data sharing has been standard practice for years. We strongly support data sharing and expect our authors to make data available immediately upon publication as well as over the long term after publication. We also actively ask authors to avoid 'data available upon request' statements except for exceptionally large datasets. Our papers also have stand-alone data availability statements to help to guide readers to source data. In addition, we are beginning a new collaboration with FigShare to host larger source data associated with our papers beginning at the peer review stage.

The fields of genomics and transcriptomics probably come closest to representing a model for data sharing, as consensus guidelines exist regarding types of data that must be shared and the form in which those data should be stored. Appropriate repositories are available for DNA and RNA sequences, genetic variation data, functional genomics data and gene expression data. The history of data sharing in genomics makes data storage and sharing the expectation from the onset of experimental design. A caveat is that genomics is expanding rapidly, especially with the rapid rise of spatial omics technologies, which have their own unique requirements for sharing and for which relatively few databases exist. As these methods emerge and grow, and as they become increasingly multimodal, new standards and databases may be needed.

Proteomics and structural biology are comparably mature. There are established repositories for protein sequences and proteomics data, and structural databases associated with crystallography, nuclear magnetic resonance structure determination and cryo-electron microscopy. And again, in these fields, data sharing formats are largely agreed upon, and sharing is often mandated by publishers and is certainly expected among researchers.

Other fields are still in the process of developing best practices for data sharing. For example, immunology research involves diverse methodological approaches and data types. As such, there is no one 'catch-all' repository for immunological data, nor are there many mandatory data sharing requirements apart from those for omics data. That being said, repositories are available that cover many widely used data types, such as flow cytometry data, immunogenomics and immune receptor repertoires. Efforts in this field are underway to further develop and implement best-practice guidelines for data sharing and also to improve the diversity and quality of data in databases.

Neuroscience is another field with diverse data storage and sharing needs, where a single common repository for all neuroscience data may be difficult to envisage given the differing needs of, for example, magnetic resonance imaging, microscopy, behavioral and electrophysiology data. Nevertheless, there has been substantial progress in the development of high quality, reliable databases and a strong community effort to promote data sharing. For example, the International Neuroinformatics Coordinating Facility has developed an infrastructure portfolio to help researchers to find solutions for their data sharing needs, including structural and functional neuroscience data from multiple modalities, large-scale projects and neurogenomics data.

Microscopy does not have a long history of data sharing, and most journals have no microscopy data deposition mandates. The challenges this field faces are many and include, huge dataset sizes, diverse data output from different modalities, questions surrounding what counts as 'raw data', the need to store and save multiple versions of files due to data processing, optimal file formats, best practices for metadata recording, and cost. However, groups like Quarep-LIMI, REMBI, Global BioImaging, BioimagingNorth America and more are developing guidelines for data reporting and sharing that, should enable meaningful sharing and reuse of bioimaging data. And although not yet meeting the needs of all microscopists, image data resources and repositories such as the Image Data Resource and BioImage Archive are growing and setting standards for the field.

A few themes emerge when examining data storage and sharing solutions across different fields. For one, data size matters a great deal to the feasibility of long-term data storage, let alone data sharing. Resources such as FigShare and Zenodo are becoming widely used, but they may have associated costs, especially for very large datasets. Moreover, issues of data privacy are paramount for many types of data involving human subjects and must be considered a top priority. Issues of data provenance and metadata standards are also crucial when it comes to reuse of data. Expectations within a field for data sharing are important for experimental planning, to create data that are not only shared but are also actually reusable. Perhaps the clearest theme of all, however, is that fields that share data as a matter of routine are richer for it, especially in the age of big data and artificial intelligence. Data sharing and reuse are more important now than ever.

We think the best path forward for all researchers involves smart guidelines and community consensus best practices to avoid ad hoc data storage and sharing solutions and promote reproducibility and reuse of data of all types. We hope that grant-funding agencies take note of the great needs of these diverse communities and continue to fund and develop stable databases to help to take storage and sharing burdens off the shoulders of individual laboratories.

Published online: 12 April 2023

Check for updates

editorial

Registered Reports at *Nature Methods*

Nature Methods is introducing a new article format: Registered Reports. We encourage all authors interested in submitting comparative analyses of the performance of established, related tools or methods to familiarize themselves with this alternative approach to peer review.

Much has been written about the reproducibility crisis in scientific research. The pressures on researchers to publish novel and exciting results can lead some to poor scientific practices ranging from cherry picking, HARKing (hypothesizing after the results are known), *P*-hacking and, at worst, outright fraud.

One approach some journals are taking to avoid such poor practices is by offering an article format known as a Registered Report. On the surface, a published Registered Report appears much like a traditional research paper, but this format differs radically in the approach to peer review: the review of the experimental design plan and acceptance 'in principle' by the journal occurs at an early stage, often before any experiments have been carried out. This helps shift reviewers' and editors' focus to the soundness of the research question and experimental design, rather than on the perceived importance of the results. By its very nature, this format encourages greater transparency in reporting and the publication of 'negative' results.

Registered Reports were initially introduced in 2013 by the journals *Cortex* and *Perspectives in Psychological Science*. This format has since become relatively commonplace in the social sciences. It is not well known in the life sciences community, though interest is growing. Several journals publishing basic biology research now offer the Registered Reports format, including *eLife*, *PLoS Biology*, *PLoS One*, *Scientific Reports* and several BioMed Central journals.

We have been particularly inspired by our colleagues at *Nature Human Behavior*, who have offered this format since they launched in 2017. By adopting their workflow and adapting their guidelines to suit the unique needs of life sciences research, we are now introducing the Registered Reports format at *Nature Methods* for papers describing comparative analyses of tools or methods.

Although Registered Reports were initially designed for hypothesis-driven research and replication studies, we realized that they are also ideal for comparative analyses. The key contributions of such papers are whether a comparison is valuable for a research community and whether it is scientifically robust — not whether a particular method performs best. (Note that the Registered Reports format is not suitable for method development papers themselves.)

In traditional peer review, if reviewers identify a fundamental flaw in a large-scale comparative analysis study, there is usually little the authors can do to address this: coordinating a completely revised analysis, often performed by researchers at different institutes, would be a logistical nightmare. However, when peer review takes place before any experiments have been carried out, this enables authors to rework their design plan to ensure that it is robust and meets standards in the field.

The general process for peer review of Registered Reports at *Nature Methods* is as follows. First, authors submit a 'Stage 1' manuscript, which should include an Introduction that justifies the value of the comparative study and a detailed experimental plan, including a data analysis plan. If the Stage 1 manuscript meets our editorial criteria for scope, novelty, potential interest and comprehensiveness, it will be peer reviewed. If, over one or more rounds, reviewers find the experimental plan to be valuable and scientifically sound, the editors will offer an 'accepted in principle' decision. At this point, authors must register their Stage 1 paper in an appropriate repository, such as Figshare. Next, authors carry out their experiments and then resubmit the full 'Stage 2' manuscript, now including Results and Discussion. Reviewers perform a final technical evaluation of the Stage 2 manuscript, but the editors will not reject papers at this stage for reasons such as scooping or the perceived importance of the results.

The 'accepted in principle' decision is conditional on the assumption that authors will not substantially deviate from their Stage 1 manuscript. Further exploratory analysis of the results is allowed, but must be clearly stated as such in the Stage 2 submission. Should authors with an 'accepted in principle' agreement realize that they need to make significant changes to their experimental plan, they should contact the editors as soon as possible. To avoid this, we encourage authors (especially for wet lab studies) to include pilot data in the Stage 1 manuscript that demonstrate the feasibility of the proposed broader analysis. Authors must also agree in writing at Stage 1 to make their data, code and unique materials available upon publication.

Over the past two years we have closely engaged with several research groups and sent three Registered Reports out for peer review. We are very pleased to report that one of these papers, reporting an experimental comparison of near-infrared fluorescent proteins, is now accepted in principle, and the Stage 1 manuscript is available via Figshare.

We've learned several valuable lessons during this trial period: for example, that our guidelines need to be sufficiently flexible to allow for minor changes to experimental design. We've also realized that Registered Reports really do need to be submitted for review before a large-scale study gets underway — otherwise, as with a regular Analysis paper, it becomes too logistically difficult to address reviewer concerns about experimental design. We've also found that reviewers may have a kid-in-a-candy-shop tendency to request experiments that go beyond the reasonable scope of a study; thus, the role of the editor in giving clear advice to authors about what experiments we expect to see in a revision is essential.

We encourage all authors interested in submitting comparative analyses to *Nature Methods* to familiarize themselves with our guidelines for Registered Reports and reach out to us at the early stages of a project with presubmission inquiries and questions.

We hope the expert guidance authors will receive from Stage 1 manuscript peer review and the ability for the editors to provide an 'accepted in principle' decision at Stage 1 will make it more attractive for researchers to pursue scientifically valuable and sound comparisons of tools or methods. We also hope with this announcement that we may raise awareness in the life sciences community of the many benefits of this alternative approach to peer review. ❐

Published online: 10 February 2022
https://doi.org/10.1038/s41592-022-01407-4

In summary

Transparency in published manuscripts is key to research integrity

- Methods, materials and code reporting
- Data reporting

Newer initiatives by Nature Research journals help support research integrity

- New peer review initiatives
- New format: Registered Reports

Dominique Morneau-Brosnan
Chief Editor
Nature Reviews Methods Primers

How Reviews journals can promote research integrity

Nature Reviews

Clinical Sciences: Cardiology, Clinical Oncology, Endocrinology, Gastroenterology & Hepatology, Nephrology, Neurology, Rheumatology, Urology

Life Sciences: Cancer, Molecular Cell Biology, Genetics, Immunology, Neuroscience, Microbiology, Drug Discovery, Psychology

Physical sciences: Materials, Chemistry, Physics, Earth & Environment, Bioengineering

Primers: Disease Primers, Methods Primers

Reviews journals and research integrity

- Agenda-setting.
- A platform for critiquing the body of literature on a given subject.
- Nature Reviews journals publish consensus statements, evidence-based guidelines and expert recommendations.
- Technical reviews, Tools of the trade, Primers.

Research integrity

- Methodological details
  - Hardware
  - Software
- Analytical details
  - 'Black box' programs
  - Statistics
- Data storage, management and sharing

Nature Reviews Methods Primers

A Methods Primer is intended to outline best practices at every stage of an experiment – from design to data analysis to data sharing.

The structure is rigid and set by the journal – we make sure that each Primer contains the key elements readers need to adopt a method at each stage of the experiment.

We encourage authors to explain analytical steps in detail, even for 'black box' methods where possible.

Authors in fields without open science/data standards are encouraged to discuss minimum reporting and basic repositories.

Experimental design checklists

Box 1 | Checklist for getting started with CRISPR screening

1. What are the most suitable models and the most relevant phenotypic read-outs for the investigated biological process? Simple read-outs such as cell proliferation or a selectable marker tend to be easier, cheaper and more scalable. By contrast, complex read-outs including single-cell RNA sequencing (RNA-seq) and spatial imaging can provide much more detailed information already as part of the screen.
2. What are realistic cell numbers and what are the proliferation characteristics of the chosen model? What is the largest cell population that can be maintained at acceptable cost and effort throughout the screen?
3. Is the model amenable to lentiviral transduction? If not, which alternative delivery methods might be applicable? For example, what about transient transfection of plasmids, ribonucleoprotein (RNP) complexes or mRNA, or transposon-mediated integration of the Cas protein into the genome? What are suitable delivery vectors? The list of popular plasmids on the website of the non-profit Addgene repository provides an up-to-date starting point.
4. Has the delivery and perturbation been optimized for best efficiency? Has the model been tested with suitable positive and negative controls and optimized to provide a high signal-to-noise ratio?
5. Considering points 1–4, is the model compatible with genome-wide screening? If not, what is the maximum library size that can be screened with adequate coverage? As a guideline, we recommend a coverage of 100–200 cells per target gene for positive selection screens and 500–1,000 cells per target gene for negative selection screens.
6. What type of guide RNA (gRNA) library is most suitable? Does the selected library include all relevant controls or do they need to be added separately? Genome-wide libraries are readily available from Addgene and other sources. Nevertheless, many applications benefit from the creation of custom libraries.
7. Before performing the first screen, which gRNAs and genes are expected to be enriched or depleted based on the study design (positive and negative controls) and relevant biological knowledge? After the first screen, did the expected results materialize with good signal-to-noise ratio? If not, what are potential problems and how can the screen be improved?
8. After running the screen in three biological replicates, how consistent are the results? Does the screen need to be optimized and repeated? Which hits are

Box 2 | Checklist for planning SFX experiments, including time-resolved studies

Experiment cost–benefit analysis

- Consider the scientific impact or technical advances of the proposed research in view of the resources and effort involved in serial femtosecond crystallography (SFX) data collection.

Use of SFX data collection

- Decide whether SFX is needed or whether another method — macromolecular crystallography or serial synchrotron crystallography — can be used instead. This depends on the sample to be studied, the required spatial resolution and in case of time-resolved experiments, the time scale and structural changes expected.
- If SFX is the best approach, choose an appropriate facility and instrument for data collection.

Sample properties and availability

- Sample-related considerations include the material availability; stability; aggregation; crystallization conditions, as this affects the choice of sample delivery; and crystal properties, including space group, unit cell, diffraction resolution and physical characteristics.

Sample delivery method

- The choice of method is dependent on the available crystal quantity, crystal size and size homogeneity, crystal growth medium, and scattering strength.
- For time-resolved experiments, the choice of delivery method is additionally influenced by the triggering method and time delay.
- Available options include jets (gas dynamic virtual nozzle, high-viscosity extrusion and microfluidic electrokinetic sample holder); drop-on-demand or drop-on-tape methods; and chips or fixed targets.

Risk assessment

- The risk assessment should evaluate what assumptions are involved in the experiment, what can go wrong, how to respond, and steps that can be taken to rescue the beam time if the planned experiment cannot be performed (plan B).

Additional points to consider for time-resolved experiments

- The reaction scheme, including the kinetic rates and the time scales and [illegible] in the

- Provide a guideline for different decision points in experimental design.
- Decisions at each point need to be verified ahead of time and optimized.
- Expert advice on sample and library sizes and where to find resources.

Comparison of analytical tools

- Provide a guide for selecting the best analytical tool.
- Can be both wet-lab and computational.
- Easy to compare between methods based on research question and data type.

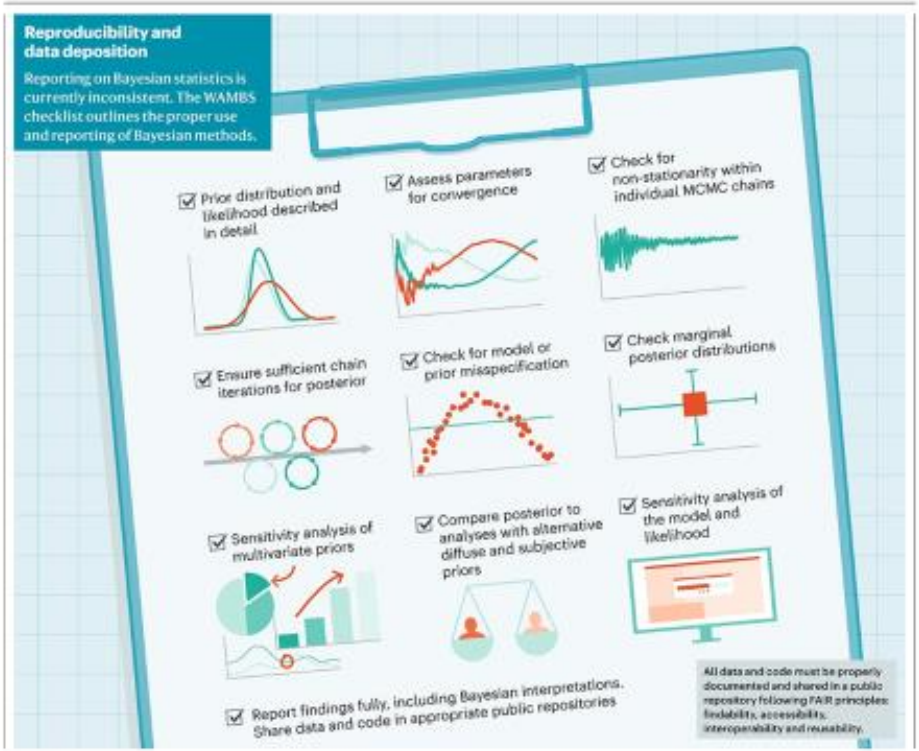
Table 2 | **Online separation options for glycopeptide analysis**

Technique	Description	Separation modality	Resolves glycopeptide isomers?	Advantages	Disadvantages	Degree of use
Liquid chromatography-based techniques						
Reverse phase	Separation based on interactions with hydrophobic stationary phase	Hydrophobicity; mostly peptide backbone, some glycan influence	Poor resolution	Simple, robust, MS amenability[a], high peak capacity[b]	Poor separation of glycan isomers	Widespread
Hydrophilic liquid interaction chromatography (HILIC)	Separation based on interactions with hydrophilic stationary phase	Hydrophilicity; mostly glycan and charge, some peptide influence	Yes	MS amenability[a], multiple resin types, improved separation of glycopeptide isomers relative to RP	Lower peak capacity than RP, glycan class biases	Moderate
Porous graphitized carbon (PGC)	Separation based on complex interactions with chemical surfaces of a crystalline graphitic stationary phase	Hydrophobicity and charge; complex retention, peptide and glycan	Yes	MS amenability[a]	Difficult to elute glycopeptides, complicated retention mechanism	Limited
Other techniques						
Capillary electrophoresis	Separation based on electrophoretic mobility induced by an applied voltage	Charge-to-size ratios; significant glycan influence	Yes	High reproducibility, improvements in sensitivity and peak capacity over LC separations, glycopeptide isomer separation	Not entirely orthogonal to m/z measurements, limited mobile and stationary phase combinations	Limited but growing
Ion mobility	Separation based on mobility of gas-phase ions through a carrier gas	Gas-phase conformation; multiple characteristics determine conformation, multiple approaches available	Yes	Rapid, potentially compatible with other online separations	Specialized instrumentation, limited peak capacity	Limited but growing

LC, liquid chromatography; MS, mass spectrometry; m/z, mass to charge ratio; RP, reverse phase. [a]MS amenability refers to ease with which a separation technique can be coupled to MS measurements in an online fashion. For example, separations like RP are highly MS amenable because they do not require buffers with salts. [b]Peak capacity is defined as the maximum number of peaks that can be theoretically separated on a column given certain conditions and serves as a general measure of elution peak width per unit of gradient time.

Reporting standards checklists

- Provide a standard for reporting practices.
- Help in interpretation of data and results.
- Allow researchers to assess the validity of data and parameters used in analysis, enabling errors to be identified and fixed.



Data repositories

Table 2 | **Commonly used databases for archiving and distributing chromatin accessibility data**

Database type	Database	Description
General epigenomic databases	Gene Expression Omnibus (GEO)	Repository that archives and distributes microarray and high-throughput sequencing data submitted by the research community
	ArrayExpress	Repository that stores and allows sharing of data from high-throughput functional genomics experiments
Databases to deposit raw sequencing data	Sequence Read Archive (SRA)	Largest publicly available repository of high-throughput sequencing data
	European Nucleotide Archive (ENA)	Database for archiving and sharing all types of nucleotide sequencing data
	DNA Data Bank of Japan (DDBJ)	Database for archiving and sharing all types of nucleotide sequencing data
	European Genome-phenome Archive (EGA)	Database for archiving and sharing all types of personally identifiable genetic and phenotypic data resulting from biomedical research projects
	Database of Genotypes and Phenotypes (dbGaP)	Repository for archiving and distributing individual-level human data and results from studies that have investigated the interaction of genotype and phenotype
Databases to deposit code	GitHub	Platform on which researchers can host software development and perform version control using Git
	Zenodo	Repository for the deposition of both code and data
	Kipoi	Repository of ready to use trained machine learning models for genomics
Databases that make processed data easily accessible: portals of large consortia	Encyclopedia of DNA Elements (ENCODE)	Consortium with the goal of building a comprehensive list of functional genomic elements in the human genome using various omics assays
	Roadmap Epigenomics	Consortium aiming to deliver a collection of normal epigenomes (via histone ChIP-seq, DNase-seq, etc.) across a broad range of cell types that can serve as a reference for future studies
	BLUEPRINT	Consortium effort to map epigenomes of the haemopoietic system for healthy and diseased individuals
Databases that make processed data easily accessible: portals based on meta-analyses	ChIP-Atlas	Integrative database for visualizing and making use of public ChIP–seq data
	ReMap	Platform of integrative analysis of Homo sapiens and Arabidopsis thaliana transcriptional regulators from DNA-binding experiments, including ChIP–seq
Databases that make processed data easily accessible: study-specific portals	Many, e.g. mouse sci-ATAC-seq Atlas	Laboratory-specific, often include several tabs covering, e.g., data visualization and data download

Table 2 | Guidelines for standardized and reproducible extracellular vesicle (EV) research and data deposition

	Repository	Description	Curated	Updated
Reproducibility and standardization	MISEV	Minimal information for Studies of EVs (MISEV): a position statement of the International Society for Extracellular Vesicles	NA	Yes
	EV-TRACK	Transparent Reporting And Centralizing Knowledge in EV research	Yes	Yes
EV data deposition	Vesiclopedia	A compendium of RNA, proteins, lipids and metabolites in EVs	Yes	Yes
	Exocarta	A web-based compendium of exosomal cargo	Yes	No
	EVpedia	A community web-resource for prokaryotic and eukaryotic EV research	Yes	No
	ExoRBase	An atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids	Yes	Yes
	EVmiRNA	A database of miRNA profiling in EVs	Yes	Yes
	EV-ADD	Database for EV-associated DNA in human liquid biopsy samples	Yes	Yes
Cargo-specific data deposition	PRIDE	The proteomics identification database	Yes	Yes
	The Global Proteome Machine Database (GPMDB)	Open-source system for analysing, validating, and storing protein identification data	No	Yes
	MassIVE	Mass spectrometry interactive virtual environment	No	Yes
	PeptideAtlas	A multi-organism, publicly accessible compendium of peptides identified in a large set of tandem mass spectrometry proteomics experiments	No	Yes
	exRNA	The data repository of the Extracellular RNA Communication Consortium (ERCC)	No	Yes
	Gene Expression Omnibus (GEO)	A database repository of high-throughput gene expression data and hybridization arrays, chips and microarrays	No	Yes
	Exceropt	A comprehensive analytic platform for extracellular RNA profiling	No	Yes
Method-specific data deposition	MIFlowCyt MIFlowCyt-EV	A framework for standardized reporting of (EV) flow cytometry experiments	NA	Yes
	MIQE guidelines	Minimum information for publication of quantitative real-time PCR experiments	NA	Yes

Documentation of code

- Primers outline appropriate repositories for code, metadata and data, including entry requirements where appropriate.
- Those Primers that include code for illustrative models and data analysis are shared for readers to access.

All outputs – such as data, software, hardware and others – should be published in a dedicated data repository. For environmental and ecological sciences, this could be the GBIF or Dryad. Multidisciplinary repositories such as OSF or Zenodo are also appropriate. These repositories provide a Digital Object Identifier (DOI), which allows a dataset to have a permanent citable reference. The Registry of Research Data Repositories provides a list of additional data repositories[182]. While software and hardware design often occur on version control platforms such as GitLab or GitHub, copies should be deposited in these repositories. The Cos4Cloud project hosts various online services to aid citizen science data interoperability and reproducibility for uptake into the European Open Science Cloud.

Code availability

Example code for a nutrient–phytoplankton–zooplankton–detritus (NPZD) model, parameter optimization and state estimation can be found in refs.[57,81,201], respectively.

Code availability

The two example data analysis workflows for neuron and mitochondria segmentation can be found at: https://github.com/kreshuklab/vem-primer-examples.

Versioning

Repositories like Zenodo provide a DOI for code and data, to make it all trackable and citable.

GitHub provides versions and documents changes and updates made to code.

Containers encapsulate and isolate all tools for data analysis, including all versions, libraries and dependencies needed.

Key points

- Reviews are an important part of promoting and ensuring research integrity.
- Checklists help researchers determine the best course of action at each decision point in their experiment, ensuring that experiments and analyses are robust and meet current best practices.
- Data and code sharing are key to reproducibility and replicability. Information on the best repositories for sharing different data types and code will help researchers make the best choice.
- Code must be versioned and made accessible in a public repository, preferably with a DOI or other identifier.