

Developing a research data policy framework for all journals and publishers

Iain Hrynaszkiewicz¹, Natasha Simons², Azhar Hussain³ and Simon Goudie⁴

An output of the [Data policy standardisation and implementation Interest Group \(IG\)](#) of the Research Data Alliance (RDA)

1. Springer Nature, London, UK
 - a. ORCID: 0000-0002-9673-5559
2. Australian Research Data Commons, Brisbane, Australia
 - a. ORCID: 0000-0003-0635-1998
3. Jisc, London, UK
4. Wiley, Melbourne, Australia
 - a. ORCID: 0000-0002-0398-2551

Corresponding author

Iain Hrynaszkiewicz

iain.hrynaszkiewicz@nature.com

Abstract

More journals and publishers - and funding agencies and institutions - are introducing research data policies. But as the prevalence of policies increases, there is potential to confuse researchers and support staff with numerous or conflicting policy requirements. We define and describe 14 features of journal research data policies and arrange these into a set of six standard policy types or tiers, which can be adopted by journals and publishers to promote data sharing in a way that encourages good practice and is appropriate for their audience's perceived needs. Policy features include coverage of topics such as data citation, data repositories, data availability statements, data standards and formats, and peer review of research data. These policy features and types have been created by reviewing the policies of multiple scholarly publishers, which collectively publish more than 10,000 journals, and through discussions and consensus building with multiple stakeholders in research data policy via the Data Policy Standardisation and Implementation Interest Group of the Research Data

Alliance. Implementation guidelines for the standard research data policies for journals and publishers are also provided, along with template policy texts which can be implemented by journals in their Information for Authors and publishing workflows. We conclude with a call for collaboration across the scholarly publishing and wider research community to drive further implementation and adoption of consistent research data policies.

Introduction

An increasing number of publishers and journals are implementing policies that require or recommend that published articles be accompanied by the underlying research data ¹. These policies are an important part of the shift toward reproducible research and contribute to the availability of research data for reuse ².

While uptake of journal data policies is on the rise, there is wide variation between policies on aspects such as content, discoverability, ease of interpretation, infrastructure integration and support for compliance. This makes it challenging for journal editors to develop and support a data policy, difficult for researchers in understanding and complying with data policies, and complex for infrastructure providers and research support staff to assist with data policy compliance. There is clear benefit in a more standardised approach, as evidenced in the findings of the Jisc UK Journal Data Registry Project and the pioneering work of publishers, such as Springer Nature, to develop and support standard policy types for their journals ^{3,4}.

This research data policy framework is intended to help journal editors and publishers to navigate the creation or enhancement of a research data policy. It reflects international efforts by the Research Data Alliance (RDA) Data Policy Standardisation and Implementation Interest Group ⁵ to identify the key elements of a good data availability policy and to standardise data policies.

Methods

The initial list of research data policy features included in this policy framework was developed by reviewing, combining and harmonising requirements from existing scholarly publishers' research data policies - Springer Nature, Elsevier, Wiley, PLOS ⁶⁻⁹. The CODATA best practice guidelines for research data policy ¹⁰ and the TOP guidelines ¹¹ were also included in the review of existing policy frameworks. The first version of the framework also incorporated feedback on, and requirements for, research data policy gathered during RDA plenary meetings and community conference

calls/web meetings that were conducted during 2017. The first draft version of the framework (v1.2) ¹² was made available for public comment for a period of three weeks, shortly before the March 2018 11th RDA Plenary meeting in Berlin. More than 30 comments were received from nearly 20 reviewers. The draft framework and a synthesis of the comments received were presented at the Berlin meeting, with further feedback received from attendees. The present version of the framework aims to address important feedback received from the community of reviewers on issues of scope, presentation, and clarity. It also aims to serve as a tool for editors and publishers to understand and implement standardised research data policies at their journals.

Results and Discussion

Table 1 defines what each of the 14 research data policy features are and the reason for their inclusion as part of the policy framework. Figure 1 summarises which features are included in which policy type and provides a visual representation of how the feature is implemented.

The features are arranged into six types of research data policy, with increasing numbers of features and policy stringency as one progresses from the first type of policy through to the sixth.

The list of policy features and whether they are enforced through action is prescriptive. However, exactly how each policy feature and its requirements are implemented is not prescriptive, as the operations and resources available to different journals varies greatly. We, however, provide some implementation guidance and templated policy text for editors and publishers, which journals are encouraged to reuse. We acknowledge that wording and implementation methods will vary between journals, publishers and research disciplines. The scope of this document does not extend to supporting guidance and resources that are linked to from several policy features. For example, lists of recommended data repositories and criteria for assessing data repositories are not in scope. The scope of this document also does not include detailed guidance on preparing data availability statements, or detailed guidance on implementing data citation at scholarly publishers. Where appropriate, this document links to other initiatives that have or are defining more detailed guidance in these areas ^{13,14}.

Table 1: Policy feature definitions

Feature	What	Why
Definition of research data	Define which research data the policy applies to, and the types of research data covered by the policy.	This enables the policy to define its scope and, where appropriate, provide general or discipline specific information on research data and file and format types ¹⁰ . Specifying non-numeric data types (images, video, text etc) helps ensure relevance and applicability across research disciplines.
Definition of exceptions	Define what data do not need to be, or should not be made publicly available, under the policy and the alternative options for describing the availability of these data.	Ensures data policy is applicable to all research publications, but acknowledges legitimate exceptions and makes clear the policy does not create new legal or ethical precedents.
Embargoes	Define if and what embargoes on data release are permitted.	Researchers' reasonable right of first use of data generated during their research is a widely accepted principle of data sharing ¹⁵ , but reasonable lengths of embargo may vary by discipline, data type and study.
Supplementary materials	Define the journal/publisher's position on data sharing via supplementary materials, and if and when sharing data supplementary materials is permitted under the policy.	While many policies preference sharing of data in repositories ¹⁶⁻¹⁸ , sharing data as supplementary materials remains very common. Some journals have strong data sharing policies and specify supplementary materials as the mechanism for data sharing. Supplementary materials are often a solution for researchers without discipline specific repositories and the definitions of supplementary material and research data often overlap.
Data repositories	State position on the use of data repositories. Data repositories are the preferred mechanism for sharing data with community/discipline specific repositories preferred to general	Lack of an appropriate repository or lack of awareness of repositories are common reasons reported by researchers for not sharing data ¹⁹ . Journal and publisher information for authors is an important way of raising awareness of the

	repositories, where they are available.	availability of repositories for the majority of research data ²⁰ .
Data citation	Statement on the journal/publisher's support for the provision of persistent identifiers for research data that support publications, and statement of support for data persistent identifiers to be included in the reference list as formal citations. Includes whether data citation is encouraged or required.	Citing and linking to data increases visibility of research, increases academic credit and has been correlated with published articles receiving more citations ²¹⁻²⁵ . This benefits authors, journals, publishers, and society. Data citation in reference lists occurs in a fraction of published literature but is steadily increasing ²⁶ . To ensure data citation happens consistently in published articles requires additional effort from authors and editors, and therefore operational costs, but enhances reader and use experience by consistently linking important research outputs ¹³ .
Data licensing	Define position on licensing and copyright for research data.	Lack of understanding of copyright and licensing of research data is a common reason why researchers don't share data ¹⁹ . Journal/publisher policy can help increase awareness and prevalence of explicit, and ideally open, licenses for research data. However, many established repositories do not have open data-conformant licenses and this is unlikely to change in the foreseeable future ²⁷ . Whether the journal/publisher requires copyright transfer in datasets is a frequently asked question of publishers. Publishers issued a joint statement in 2006 declaring they would not take copyright in research data ²⁸ .
Researcher/author support	Information on who authors should contact at the journal or publisher for more information on complying with the policy.	Research data sharing remains a new concept for some journals and disciplines and common questions can be answered by journal and publishing staff, such as on writing data availability statements, finding repositories and on exceptions to the policy ²⁹ .

<p>Data availability statements (DASs)</p>	<p>Define position on provision of data availability statements..</p>	<p>Data availability statements are a simple, consistent, human - and increasingly, machine - readable way of expressing data availability and policy compliance.</p> <p>They are already encouraged, expected or required by many journals and publishers and some funding agencies ^{14,30}.</p>
<p>Mandatory data sharing (specific papers)</p>	<p>Statement on whether data sharing is mandatory for specific types of research data, such as where there is a community or journal-specific mandate, and the mechanism(s) by which these types of data must be shared. Examples include DNA and RNA sequence data, and macromolecular structure data.</p>	<p>Where there are established community mandates for data sharing, journals and publishers have an obligation to support editors and authors in upholding community standards as part of their service to the research communities they serve ³¹.</p>
<p>Data formats and standards</p>	<p>State position on the use of community/discipline-specific data standards - whether encouraged, required in some cases, or required in all cases. Also state whether certain file formats, such as open formats, are preferred or required.</p>	<p>Data prepared according to community standards are more interoperable and reusable, and data available in open formats are more accessible ³².</p> <p>Data standards are distinct from reporting standards, which are not within the scope of a research data policy. (e.g. MIAME as a reporting standard for papers describing for microarray experiments).</p>
<p>Mandatory data sharing (all papers)</p>	<p>Sharing of research data via an external mechanism (repositories or supplementary information) is a condition of submission or publication for all articles published.</p>	<p>Mandatory data sharing policies that are enforced during the peer-review and publishing process, and supported with data repositories, are the most effective policies ². These policies can also be more costly (time consuming) to implement and have the greatest impact on editors and authors ³³. They could however have the most benefits in terms of increasing citations and visibility of papers.</p>

Peer review of data	<p>Statement on whether peer review of data is expected or required, and if so what the expectations of peer reviewers are in their assessment of data files.</p> <p>Reviewers can also or alternatively be asked to assess compliance with research data policy.</p>	<p>Where data are made available with research articles they are accessible to peer reviewers, but for journals with a strong focus on data, such as data journals, consistent review of the data and the description of those data can be required. Peer review traditionally focuses on manuscripts rather than data, but more consistent availability of data for validation and reuse can improve the reproducibility - and quality - of published research ³⁴.</p>
Data Management Plans (DMPs)	State position on sharing of DMPs.	<p>This is currently uncommon in journal and publisher policy although encouraging their provision is analogous to how many medical journals encourage or require sharing or publication of study protocols. Furthermore, they are increasingly required by funding agencies. Some journals, such as RIO journal, publish them as articles.</p>

Features arranged by policy type and implementation method

The 14 features are arranged into six types or tiers of policy, with more features and requirements as one moves from policy one through to six (Figure 1). The six tiers allow for more nuanced, step-wise and robust implementation of policies by different journals. This tiered approach to policy guidelines and frameworks is already in place at numerous large publishers, which from 2018 also includes Taylor & Francis ³⁵, and BMJ ³⁶. This tiering also acknowledges that the later features require the most effort to implement, such as data peer review (policy 6). This six-policy approach also provides a specific policy for journals that mandate data sharing but do not carry out data peer review routinely (policy 5). An option is also available for journals that wish to mandate data availability statements, but which do not have the means to check the contents of those statements in detail or enforce any data sharing mandates (policy 3). Journals that can commit to enforcing all relevant mandates for their communities adopt policy 4. Policy 2 enables a journal to provide full information on data sharing standards and good practice, but without the need to enforce any aspect.

Figure 1: Fourteen journal research data policy features arranged as six policy types (tiers)

14 journal research data policy features arranged as 6 policy types (tiers)

	Policy 01	Policy 02	Policy 03	Policy 04	Policy 05	Policy 06
Definition of the research data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exceptions to policy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Embargoes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Supplementary materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data repositories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data citation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data licensing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Researcher/ author support	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data availability statements		<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Data formats and standards				<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Mandatory data sharing (specific data types)				<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Mandatory data sharing (all papers)				<input type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>
Peer review of data				<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Data Management Plans (DMPs)				<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Provide information

The text for the policy feature will be included in the policy template but it is clear that the feature will not be enforced and checked as part of the publishing or peer review process

Provide information and action

The text of the policy feature is included and makes clear where applicable that the feature will be checked and enforced in the publishing or peer-review process

Explanation and implementation requirements of the research data policy framework features

Policy feature: Definition of research data

Every policy must define research data as being the data that support the findings or claims made in the published article.

Policies should also specify what kinds of data are included in the policy, such as tabular data, code, images, audio, video, maps, raw and/or processed data.

Policies should further include and define their coverage of:

- Data produced by the authors for the study being reported (“primary data”)
- Data reused or analysed by the authors for the study being reported (“secondary data”)
- Additional data from experiments or observations, such as raw, unprocessed data or data from time points not reported in the article
- Any other materials that might be required to reproduce or replicate the results

Implementation notes

Journal or publisher’s editorial policy text or information for authors must include the definition of research data. See template policy text for an example.

Policy feature: Definition of exceptions

The policy must define the types of data that it does not expect to be shared publicly. It should also define, if applicable, research data that are not covered by the policy - if these are not already explicit in preceding feature, Definition of research data.

Data that a journal or publisher does not expect to be shared publicly may include personal or sensitive data, such as quantitative or qualitative data that could identify an individual without their consent, locations of endangered species, and data subject to other legitimate restrictions on public availability.

Other types of sensitive data must also be defined if they are applicable to the journal/publisher’s content. Alternative options for public sharing of these data and describing their availability should be given, such as:

- Controlled access repositories
- Anonymisation and de-identification of data
- Sharing metadata only

- Working with data access committees and implementing Data Use Agreements

Aspects of this policy feature and text may be superseded, substituted or modified by the feature “Mandatory data sharing (all papers)”

Implementation notes

For policies 1-4: Journal or publisher’s editorial policy text or information for authors must include examples of the types of data that they do not expect to be shared publicly. See template policy text for an example.

The journal or publisher, via its editorial or peer-review process, must be able to identify if authors are sharing sensitive or personal data without appropriate consent. Where this is identified, it must advise authors of appropriate action, referring to resources on alternatives to public data sharing where appropriate.

For policies 5 and 6: These journals require mandatory data sharing for every publication, evidenced by datasets cited in reference lists. In such cases, raw data, such as individual participant data from clinical studies that are not anonymised, might not be publicly available but the data must be archived in a secure repository that provides a persistent identifier and landing page for the data so that the data can be cited.

Policy feature: Embargoes

The policy must include a statement about the journal or publisher’s position on embargoes. This may need to consider community norms, funding agency policies (where applicable) and enabling data generators a reasonable right of first use. The policy should provide information on any relevant community-specific embargoes.

Implementation notes

Journal or publisher’s editorial policy text or information for authors must include information on embargoes. The journal or publisher must be prepared to respond to and resolve unreasonable embargo periods on data included in its policy’s definition of research data, if it is made aware of them. See template policy text for an example. Note that for policies 5 and 6, these tiers do not permit embargoes on data access and data must be accessible to readers at the publication date, and at minimum have been accessible to editors and peer reviewers before publication.

Policy feature: Supplementary materials

Data repositories are the preferred method for sharing data supporting publications and this must be stated in the policy. The policy must also specify if sharing research data via supplementary materials, or an equivalent method by which data objects are archived by the publisher as part of the published article, is permitted.

Implementation notes

Journal or publisher's editorial policy text or information for authors must include information on if data sharing via supplementary materials is permitted. This statement may need to reference and be consistent with the publisher's existing policy on supplementary materials. Journals and publishers that deposit supplementary materials files in third-party repositories, such as figshare, that assign persistent identifiers to each file should provide further information on these services. See template policy text for an example.

For policies 5 and 6: Data sharing via supplementary materials is not permitted and the journal must support this requirement with checks in the editorial or peer review process to ensure that datasets supporting the claims in the paper are deposited in appropriate repositories and cited in the reference list.

Policy feature: Data repositories

Data policies must be supported with data preservation, which will require a list of recommended, trusted or supported data repositories. This could be the journal/publisher's own list, a community/discipline-specific list, or a curated and trusted third-party list, such as those available from FAIRsharing.org or a repository finder tool or service such as <https://repositoryfinder.datacite.org/>. Data policies must also preference the use of community/discipline-specific data repositories over general data repositories, where community/discipline-specific repositories exist. The provided list must include general repositories, if community/discipline-specific repositories cannot support all research data included in the definition of the policy.

If the publisher's own list of repositories is provided, it must include criteria for adding repositories to the list and a position statement on its support for institutional data repositories. Different standards for assessing trusted data repositories exist - such as the CoreTrust Seal and Springer Nature/*Scientific Data*'s criteria for recommended

repositories - but it is beyond the scope of this document to define standard criteria for trusted data repositories.

Implementation notes

For policies 1-3: Journal or publisher's editorial policy text or information for authors must include information on data repositories. See template policy text for an example. Information on data repositories, and a reference to the research data policy in general, should also be communicated to authors at an appropriate point during submission of manuscripts. This could be communicated via the journal or publisher's manuscript submission system and/or in standard email correspondence sent to authors during the editorial and peer-review process. Journals and publishers must be prepared to respond to requests from authors for advice on finding appropriate data repositories.

For policies 4-6: The use of data repositories, for specific (policy 4) or all (policies 5 and 6) datasets supporting publications is mandatory. This requirement must be enforced by checks in the editorial or peer review process to ensure datasets are deposited.

Policy feature: Data citation

The policy must enable, and for policies 5 and 6 require, authors to cite datasets in the reference lists (bibliographies) of their articles. It must also include the journal/publisher's style(s) for referencing datasets. One or more examples of data citation should be included. The policy should also specify if the journal or publisher has any restrictions on which datasets can be cited in reference lists, such as those that have particular types of persistent identifier (e.g. Digital Object Identifiers [DOIs], accession codes, etc).

The policy should include links to more examples of data citation in published articles, and further information on the benefits of citing and linking data are desirable.

Implementation notes

For policies 1-4: Journal or publisher's editorial policy text or information for authors must include information on data citation. See template policy text for an example. Journals and publishers must also ensure that authors who cite data in their references do not receive conflicting information during the publishing process, nor should data citations be arbitrarily removed from reference lists.

For policies 5 and 6: Accurate and consistent provision of data citations must be enforced through editorial, peer-review and/or article production procedures. This

requirement must be included in the policy and supported by checks on manuscripts that ensure publicly available, persistently-identified datasets are cited in reference lists. Automation to identify dataset identifiers, by publishers, can aid the identification of datasets that must appear in reference lists.

Implementation of data citation by publishers also has implications for content structure and XML production workflows, which is beyond the scope of this document. Journals and publishers that are implementing data citation should consult the data citation roadmap for scholarly publishers ¹³.

Policy feature: Data licensing

The policy must specify:

- What license is applied to research data published in the journal itself. It must also specify that copyright in research data is not transferred to the publisher.
- If the journal or publisher has expectations for the licenses authors make their research data available under, and what those expectations are.

The policy must express a preference for Open Data conformant licenses (such as CC BY, CC0). Licensing can also be addressed in part with recommended repositories, as criteria for trusted repositories often include requirements for licensing.

Implementation notes

Journal or publisher's editorial or publishing policy text or information for authors must include information on data licensing. See template policy text for an example. Journals and publishers cannot enforce specific licenses for research data that are deposited in third party repositories, as the licenses applied by repositories are generally outside of the publisher's control. Journals and publishers must be prepared to respond to questions from authors about licensing and copyright of research data.

Policy feature: Researcher/author support

The policy must include information on who authors can contact with questions about compliance with the policy. This might include email addresses, phone numbers and/or web-based customer support tools. It may also include information on other services or organisations that researchers can approach for support for sharing research data.

Implementation notes

Journal or publisher's editorial policy text or information for authors must include contact information for author support. See template policy text for an example.

Policy feature: Data availability statements

The policy must include a definition of a data availability statement (DAS) and where it should be placed in the manuscript. It must also specify if such statements are mandatory and must state if authors are permitted to make research data “available on reasonable request”. Numerous examples of DASs and template DASs exist. Defining standards for DASs is beyond the scope of this document but is the topic of other initiatives ¹⁴.

Implementation notes

For policy 1: This feature does not apply to policy 1.

For policy 2: Journal or publisher’s editorial policy text or information for authors must include information on DASs. Template DASs for the most common types of DAS should be provided in the policy and further guidance and examples linked to. Contextual examples, such as DASs from published articles including DASs, should be provided. See template policy text for example text.

Where DASs are a mandatory

Where DASs are a mandatory part of published articles, this requirement must be communicated to authors at an appropriate point during submission of manuscripts as well as in the information for authors or editorial policy text. This can be communicated via the journal or publisher’s manuscript submission system and/or in standard email correspondence sent to authors during the editorial and peer-review process. Journal staff or editors who are responsible for ensuring mandatory sections of articles are included in published articles must update their standard operating procedures and documentation, such as manuscript templates provided to authors, to support this requirement. Journals and publishers implementing mandatory DASs should determine the likely impact, in time and cost, of this change on their authors, journal staff, and editors, and modify the resources available to the journal to support this requirement. An analysis by the Nature Research journals found adding mandatory DASs increases the time it takes to process a manuscript by several minutes ³³.

For policy 3: Accuracy of DASs is based on trust and there is no expectation for the journal or publisher to verify the accuracy of the statements for every publication.

For policies 4-6: Where data are not shared publicly, authors publishing in these journals must be willing to respond to reasonable requests from other researchers for copies of the data, where data are not publicly available, to verify or reproduce results reported in the paper. The journal and publisher must also facilitate readers’ access to data supporting publications, for example if no response is received to requests for data or the response to a request for data is not consistent with the policy. Journals must

take action where necessary if it transpires after publication that their data policy has not been adhered to. This might include contacting authors or their institutions directly, and in some cases publishing corrections, expressions of concern or retractions, in accordance with publication ethics guidelines.

Policy feature: Data standards and formats

The policy must express support for community-endorsed data standards and formats if and where any may be applicable to the journal or publisher's publications. The policy should also provide one or more specific examples of data standards and formats. A data standard is a common and interoperable way of representing, labeling or structuring data. A data format refers to the way data are stored or archived, commonly the digital file type or extension. Depositing data in a community/discipline specific data repository can often achieve adherence to domain-specific data standards.

The policy must also define its position on open and proprietary formats, and encourage the most interoperable file formats where this is practical to achieve. For example, encouraging or requiring open file formats (e.g. CSV for tabular data). Resources such as FAIRsharing.org should be linked to.

Implementation notes

For policies 1-3: This feature does not apply to policies 1-3.

For policies 4-5: Journal or publisher's editorial policy text or information for authors must include information on data formats and standards. See template policy text for example text.

For policy 6: Datasets must be shared in the appropriate standard and format, and this must be enforced through the editorial or peer review process. Enforcing deposition in community (discipline) specific data repositories can often achieve this requirement as community specific repositories often require data submission in specific formats and according to specific standards.

Policy feature: Mandatory data sharing (specific papers)

The policy must specify the data sharing mandate(s) that must be followed as a condition of submission and/or publication and the mechanisms for demonstrating compliance, such as deposition in specific repositories. Data sharing mandates typically relate to specific types of data, for which data sharing is an established norm and for which community/discipline specific data repositories exist for the data type(s) covered by the mandate. The policy must also specify if and how these mandates are enforced by the journal, such as by checks by editors, reviewers or journal staff. These mandates mostly apply to specific types of research data generated in life science disciplines. A

list of these established community data sharing mandates is available from Nature Research ³¹.

Implementation notes

For policies 1-3: This feature does not apply to policies 1-3.

For policies 4-6: These mandates are enforced and journals and publishers will need procedures in place to ensure they are enforced consistently. For journals and publishers that publish in multiple research disciplines, and where mandates may only apply to certain papers, implementation of its enforcement mechanism needs special attention and the impact of introducing enforcement measures on authors and editors determined. Enforcement can be enabled by editorial checklists (as used for example by the Nature Research journals ³⁷) and, potentially, supported by artificial intelligence tools, such as <https://www.penelope.ai/>

Policy feature: Mandatory data sharing (all papers)

For these journals, “available on reasonable request” DASs are not acceptable. The mechanism(s) for complying with the policy, such as integrated data repositories available to the journal/publisher, must be specified. The data must be available in a data repository or with the article as supplementary material and this must be verified as part of the publishing or peer review process. For clinical or sensitive data published under this policy, public sharing of raw data may not be required but deposition in a repository that supports controlled access and has independent governance procedures (such as data access committees and data use agreements) is required. Journals that wish to permit other types of exception to the policy, such as commercial restrictions should not adopt policy 5 or 6.

Implementation notes:

For policies 1-4: This feature does not apply to policies 1-4.

For policies 5 and 6: Journal or publisher’s editorial policy text or information for authors must include on its mandatory data sharing policy. See template policy text for an example. Journals must carry out checks on every manuscript that is sent for peer review to ensure that any datasets on which the claims are based are available in accordance with the policy. How this is achieved will depend on the systems and operations of the journal but the procedures build logically upon those that implement Data Availability Statements, mandatory data sharing for specific papers, and Data citation. The availability of data, such as links to datasets in repositories, must be visible to peer reviewers. Some manuscript submission systems can offer integration with general data repositories that enable confidential access to data during peer review,

such as figshare and Dryad, to enable authors and journals to comply efficiently with this policy.

Policy feature: Peer review of data

The policy must state:

- What the expectations are for peer reviewers to access data
- Whether peer reviewers should or must assess data policy compliance
- Whether peer reviewers should or must assess sufficiency of description of data files for understanding and reuse
- Whether peer reviewers should or must assess the data files themselves, such as for structure, completeness, reusability, etc
- Whether peer reviewers are expected to validate, replicate or reproduce claims/statistics reported in the paper

Criteria for peer reviewers' assessment of data files are included in Springer Nature's data policy framework ⁶, and a guide to data peer review has been produced by PLOS (<https://plos-marketing.s3.amazonaws.com/Marketing/Peer+Reviewing+Datasets.pdf>).

Implementation notes

For policies 1-3: This feature is not applicable to Policies 1-3.

For policy 4: Peer reviewers are not expected to routinely access and assess supporting datasets, although are not discouraged from doing so. For policy 4, peer reviewers are expected to include in their assessment of papers recommendations on whether the authors have complied with the journal's data sharing policy, rather than on the data files themselves. This requirement must be communicated, such as in the journal or publisher's guide to peer reviewers or peer reviewer forms. See template policy text for example text.

For policy 5: All reviewers, under this policy, will have the opportunity to see supporting data files and guidelines for reviewers must provide information on what reviewers should consider when accessing and assessing datasets. See template policy text for example text.

For policy 6: Peer reviewers are required to access supporting datasets, enabled by the journal's editorial process and data policy. Reviewers must be aware of the journal's expectations for data peer review and utilise these in their assessment of manuscripts and supporting datasets. Including these requirements in peer reviewer forms and checklists is highly desirable.

Where applicable, implementation will need to address issues with double blind peer review, as datasets may contain information that can identify the authors.

Policy feature: Data Management Plans

The policy must define if and how it incorporates the preparation and sharing of Data Management Plans (DMPs). Options include:

- Always encouraging or requiring preparation of DMPs
- Requiring DMPs when funding agencies or institutions also require them
- Encouraging publication or sharing of DMPs

Implementation notes

For policies 1-3: Feature does not apply to policies 1-3.

For policies 4-6: Journal or publisher's editorial policy text or information for authors must include information on the preparation and/or sharing of DMPs. See template policy text for an example. Under no policy type is the use, or sharing with the journal, of DMPs mandatory or enforced, as this reflects current practice in scholarly publishing and in funding agencies' policies.

Conclusion and next steps

This paper provides a comprehensive journal research data policy framework that can be adopted by and aligns with the policy requirements of all scholarly journals and publishers. It is an output of the Data Policy Standardisation and Implementation Interest Group of the RDA and has been produced with open, research community and publishing industry consultation over a period of two years. The framework is practical and pragmatic, enabling any journal to implement a research data policy that is compatible with the editorial model and procedures of the journal, and the level of support for data sharing in the journal's author and reader community. While some of the policy types in this framework might be viewed as unambitious, "overreach" has been identified as a factor associated with policy failure³⁸ and our goal is to provide a research data policy framework that is usable by the widest possible audience. Implementation, adoption and endorsement of this framework by journals and publishers is critical to its success and this process has begun in 2019. Success of this initiative can be measured, in the short term, by the number of journals and publishers who adopt this policy framework or align their existing data policy options with this framework. Longer term, success should be measured by increased levels of data sharing and reuse, which means enabling journals, editors and researchers to implement the policy types 3 and above. Policy types 3 and above require data availability statements in published articles and these are a recognised compliance monitoring and data-discovery tool. Policy implementation should be combined with ongoing evaluation of the impact (costs, as well as benefits) of data policies. It is also assumed that data reuse and reproducible research are enabled by data sharing, and further research is needed to test these assumptions in large cohorts and in multiple research disciplines.

Definitions

This section aims to provide definitions of common terms used in the document:

Primary data: Data that are collected directly from first-hand sources, using methods such as surveys, interviews, or experiments.

Secondary data: Data gathered from studies, surveys, or experiments that have been conducted by other people or for other research.

Paper, article, publication: In this document paper, article and publication are used to refer to outputs that are published in journals. Beginning as unpublished manuscripts of research not previously published, these typically undergo a peer-review process by one or more academic referees before being accepted or rejected for publication within a journal.

Data availability / accessibility statements (DAS): A data availability statement (also referred to as a 'data accessibility statement') indicates where the data associated with a paper is available, under what conditions the data can be accessed, including links (where applicable) to the data set.

Community / discipline-specific repository: A public data repository designed for housing data for a given domain of research.

Data standard: A common and interoperable way of representing, labelling or structuring data.

Data format: The way data are stored or archived, commonly the digital file type or extension.

Data citation: A reference to a published or unpublished data source, for the purpose of acknowledging the relevance of the works of others concerning the topic being discussed.

Supplementary materials: Any material that adds detail, background, or context to an article by providing, for example, multimedia objects such as audio clips and applets;

additional XML-tagged sections, tables, or figures; raw data in a spreadsheet, or a software application in a repository.

Acknowledgements

The authors thank all members of the Research Data Policy Standardisation and Implementation IG and any other individuals who provided comments on previous versions of this paper and/or contributions to our community calls, and plenary meetings. We also acknowledge previous co-chairs of the group, David Kernohan and Simone Taylor for their contributions in establishing this initiative.

Competing interests

At the time of writing Iain Hrynaszkiewicz was an employee of Springer Nature and Simon Goudie an employee of Wiley.

Appendix: Policy feature standard texts for policy template construction

Standard text for policy 1

Policy summary for authors	
<p>Authors are encouraged to make the research data that support their publications available but are not required to do so. The decision to publish will not be affected by whether or not authors share their research data.</p> <p>Required</p> <ul style="list-style-type: none"> • No action <p>Optional</p> <ul style="list-style-type: none"> • Data citation • Data sharing via repositories for all research data 	
Policy summary for Editors	
<p>The journal will encourage authors to share data but the editor's decision to publish will not be affected by whether or not authors share their research data.</p> <p>Required</p> <ul style="list-style-type: none"> • Respond to questions about this policy. <p>Optional</p> <ul style="list-style-type: none"> • Data citation • Data sharing via repositories for all research data 	
Feature	Text
Definition of research data	<p>This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors ("primary data") and data from other sources that are analysed by authors in their study ("secondary data"). Research data includes any recorded factual material that are used to produce the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.</p>

Definition of exceptions	<p>Research data that are not required to verify the results reported in articles are not covered by this policy.</p> <p>This policy does not require public sharing of quantitative or qualitative data that could identify a research participant unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Alternatives to public sharing of sensitive or personal data include:</p> <ul style="list-style-type: none"> ● Deposition of research data in controlled access repositories ● Anonymisation or deidentification of data before public sharing ● Only sharing metadata about the research data ● Stating the procedures for accessing your research data in your article and managing data access requests from other researchers
Embargoes	Embargoes on data sharing are permitted.
Supplementary materials	Sharing research data as supplementary information files is discouraged.
Data repositories	<p>The preferred mechanism for sharing research data is via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories.</p>
Data citation	<p>The journal encourages authors to cite any publicly available research data in their reference list. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.</p>
Data licensing	<p>The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.</p>
Researcher/ author support	Questions about complying with this policy should be sent to [email] .

Standard text for policy 2

Policy summary for authors

Authors are encouraged to share research data and encouraged to provide data availability statements but are not required to. The decision to publish will not be affected by whether or not authors share their research data.

Required

- No actions

Optional

- Data availability statements
- Data citation
- Data sharing via repositories for all research data

Policy summary for Editors

The journal will encourage data sharing and will publish data availability statements in articles, when they are provided by authors. The decision to publish will not be affected by whether or not authors share their research data.

Required

- Respond to questions about this policy.
- Publish data availability statements when they are provided

Optional

- Data citation
- Data sharing via repositories for all research data

Definition of research data

This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors (“primary data”) and data from other sources that are analysed by authors in their study (“secondary data”). Research data includes any recorded factual material that are used to produce the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.

Definition of exceptions

Research data that are not required to verify the results reported in articles are not covered by this policy.
 This policy does not require public sharing of quantitative or qualitative data that could identify a research participant unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Alternatives to public sharing of sensitive or personal data include:

- Deposition of research data in controlled access repositories
- Anonymisation or deidentification of data before public sharing
- Only sharing metadata about the research data

	<ul style="list-style-type: none"> • Stating the procedures for accessing your research data in your article and managing data access requests from other researchers
Embargoes	Embargoes on data sharing are permitted.
Supplementary materials	Sharing research data as supplementary information files is discouraged.
Data repositories	The preferred mechanism for sharing research data is via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories.
Data citation	The journal encourages authors to cite any publicly available research data in their reference list. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.
Data licensing	The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.
Researcher/ author support	Questions about complying with this policy should be sent to [email] .

<p>Data availability statements</p>	<p>The journal encourages authors to include in any articles that report results derived from research data to include a Data availability statement. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. Where research data are not publicly available, this must be stated in the manuscript along with any conditions for accessing the data. Data Availability statements must take one of the following forms (or a combination of more than one if required for multiple types of research data):</p> <ul style="list-style-type: none"> · The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS] · The datasets generated during and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request. · The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. · Data sharing not applicable to this article as no datasets were generated or analysed during the current study. · All data generated or analysed during this study are included in this published article [and its supplementary information files]. · The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name]. <p>More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available at <link></p>
--	---

Standard text for policy 3

<p>Policy summary for authors</p>
--

By publishing in the journal authors are required to provide a data availability statement in their articles. Authors are encouraged to share their data but not required to. The decision to publish will not be affected by whether or not authors share their research data.

Required

- Data availability statements

Optional

- Data citation
- Data sharing via repositories for all research data

Policy summary for Editors

The journal will encourage data sharing and ensure that every published article includes a data availability statement. The decision to publish will not be affected by whether or not authors share their research data.

Required

- Respond to questions about this policy.
- Data availability statements

Optional

- Data citation
- Data sharing via repositories for all research data

Feature	Text
Definition of research data	This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors (“primary data”) and data from other sources that are analysed by authors in their study (“secondary data”). Research data includes any recorded factual material that are used to produce the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.
Definition of exceptions	Research data that are not required to verify the results reported in articles are not covered by this policy. This policy does not require public sharing of quantitative or qualitative data that could identify a research participant unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Alternatives to public sharing of sensitive or personal data include: <ul style="list-style-type: none"> • Deposition of research data in controlled access repositories • Anonymisation or deidentification of data before public sharing • Only sharing metadata about the research data

	<ul style="list-style-type: none"> • Stating the procedures for accessing your research data in your article and managing data access requests from other researchers
Embargoes	Embargoes on data sharing are permitted.
Supplementary materials	Sharing research data as supplementary information files is discouraged.
Data repositories	The preferred mechanism for sharing research data is via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories.
Data citation	The journal encourages authors to cite any publicly available research data in their reference list. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.
Data licensing	The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.
Researcher/ author support	Questions about complying with this policy should be sent to [email] .

<p>Data availability statements</p>	<p>The journal requires authors to include in any articles that report results derived from research data to include a Data availability statement. The provision of a Data availability statement will be verified as a condition of publication. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. Where research data are not publicly available, this must be stated in the manuscript along with any conditions for accessing the data. Data Availability statements must take one of the following forms (or a combination of more than one if required for multiple types of research data):</p> <ul style="list-style-type: none"> · The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS] · The datasets generated during and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request. · The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. · Data sharing not applicable to this article as no datasets were generated or analysed during the current study. · All data generated or analysed during this study are included in this published article [and its supplementary information files]. · The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name]. <p>More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available at <link></p>
--	--

Standard text for policy 4

<p>Policy summary for authors</p>
<p>By publishing in the journal authors are required to make research data available to editors and reviewers, and to readers upon request. For some research data, deposition in repositories is required and this is encouraged for all research data. For some papers, the decision to publish will be affected</p>

by whether or not authors share their research data.

Required

- Data sharing via repositories for some research data
- Data availability statements
- Data sharing on request

Optional

- Data citation
- Data sharing via repositories for all research data
- Prepare and share Data Management Plans

Policy summary for Editors

The journal will require data sharing for specific types of research data and promote public data sharing by all its authors. The journal will ensure that its policy on data sharing is adhered to, including helping reviewers and readers access research data where needed. For some papers, the decision to publish will be affected by whether or not authors share their research data.

Required

- Data sharing via repositories for some research data
- Data availability statements
- Support readers and reviewers in making reasonable requests for data
- Respond to questions about this policy.

Optional

- Data citation
- Data sharing via repositories for all research data
- Peer reviewers assess compliance with research data policy

Feature	Text
Definition of research data	This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors (“primary data”) and data from other sources that are analysed by authors in their study (“secondary data”). Research data includes any recorded factual material that are used to produce the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.

Definition of exceptions	<p>Research data that are not required to verify the results reported in articles are not covered by this policy.</p> <p>This policy does not require public sharing of quantitative or qualitative data that could identify a research participant unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Alternatives to public sharing of sensitive or personal data include:</p> <ul style="list-style-type: none"> ● Deposition of research data in controlled access repositories ● Anonymisation or deidentification of data before public sharing ● Only sharing metadata about the research data ● Stating the procedures for accessing your research data in your article and managing data access requests from other researchers
Embargoes	<p>Embargoes on data sharing are only permitted with the agreement of the Editors.</p>
Supplementary materials	<p>Sharing research data as supplementary information files is discouraged. Research data of the types listed in “Mandatory data sharing (specific papers)” must not be uploaded as supplementary information files. The journal will require authors to deposit these in an approved repository as a condition of publication.</p>
Data repositories	<p>The preferred mechanism for sharing research data is via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories. Research data of the types listed in “Mandatory data sharing (specific papers)” must be uploaded to an appropriate repository. The journal will require authors to deposit these in an approved repository as a condition of publication.</p>
Data citation	<p>The journal encourages authors to cite any publicly available research data in their reference list. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.</p>
Data licensing	<p>The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.</p>
Researcher/ author support	<p>Questions about complying with this policy should be sent to [email].</p>

<p>Data availability statements</p>	<p>The journal requires authors to include in any articles that report results derived from research data to include a Data availability statement. The provision of a Data availability statement that is compatible with the journal’s research data policy will be verified as a condition of publication. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. Where research data are not publicly available, this must be stated in the manuscript along with any conditions for accessing the data. Data Availability statements must take one of the following forms (or a combination of more than one if required for multiple types of research data):</p> <ul style="list-style-type: none"> · The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS] · The datasets generated during and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request. · The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. · Data sharing not applicable to this article as no datasets were generated or analysed during the current study. · All data generated or analysed during this study are included in this published article [and its supplementary information files]. · The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name]. <p>More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available at <link></p>
<p>Data formats and standards</p>	<p>The journal encourages authors to share research data using data formats and standards recognised by their research community. Please see FAIRsharing.org for more information on established data sharing formats and standards.</p> <p>The journal prefers research data to be shared in open file formats – those that do not require proprietary software to access - where possible. For example, tabular data should be shared as CSV files rather than XLS files.</p>

<p>Mandatory data sharing (specific papers)</p>	<p>For the following types of research data, submission to a community-endorsed, public repository is mandatory: <link to list of data types where deposition is mandatory> The journal will require authors to deposit data of these types in an approved repository as a condition of publication.</p>
<p>Research data and peer review</p>	<p>Peer reviewers are encouraged to check the manuscript's Data Availability statement. Where applicable, they should consider if the authors have complied with the journal's policy on the availability of research data, and whether reasonable effort has been made to make the data that support the findings of the study available for replication or reuse by other researchers. Peer reviewers are entitled to request access to underlying data (and code) when needed for them to perform their evaluation of a manuscript.</p>
<p>Data Management Plans</p>	<p>The journal encourages authors to prepare Data Management Plans before conducting their research and encourages authors to make those plans available to editors, reviewers and readers who wish to assess them.</p>

Standard text for policy 5

<p>Policy summary for authors</p>
<p>By publishing in the journal authors are required to make research data available to editors and reviewers, and to readers. For all research data deposition in repositories is required. For all papers, the decision to publish will be affected by whether or not authors share their research data.</p> <p>Required</p> <ul style="list-style-type: none"> ● Data sharing via repositories for all research data ● Data availability statements ● Data sharing on request ● Data citation <p>Optional</p> <ul style="list-style-type: none"> ● Prepare and share Data Management Plans
<p>Policy summary for Editors</p>
<p>The journal will require data sharing in repositories for all research data. The journal will ensure that its policy on data sharing is adhered to including ensuring reviewers and readers can access research data for all papers, if they wish to. The decision to publish will be affected by whether or not authors share their research data.</p>

Required

- Data sharing via repositories for all research data
- Data availability statements
- Support readers and reviewers in making reasonable requests for data
- Respond to questions about this policy
- Data citation

Optional

- Peer reviewers assess compliance with research data policy
- Peer reviewers assess research data files

Feature	Text
Definition of research data	<p>This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors (“primary data”) and data from other sources that are analysed by authors in their study (“secondary data”). Research data includes any recorded factual material that are used to produce the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.</p>
Definition of exceptions	<p>Research data that are not required to verify the results reported in articles are not covered by this policy.</p> <p>This policy does not require public sharing of quantitative or qualitative data that could identify a research participant (“personal data”) unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Personal or sensitive data must be shared in a secure or controlled access way, in agreement with the Editors. Methods for sharing sensitive or personal data include:</p> <ul style="list-style-type: none"> • Deposition of research data in an approved controlled access repositories and sharing metadata publicly about the research data • Anonymisation or deidentification of data before public sharing <p>In these cases the procedures and conditions for accessing your research data must be included in your manuscript.</p>
Embargoes	Embargoes on data sharing are not permitted.
Supplementary materials	Sharing research data as supplementary information files is not permitted. The journal will require authors to deposit these in an approved repository as a condition of publication.

<p>Data repositories</p>	<p>Research data must be shared via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories. The journal will require authors to deposit these in an approved repository as a condition of publication.</p>
<p>Data citation</p>	<p>The journal requires authors to cite any publicly available research data in their reference list, and will verify this as a condition of publication. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.</p>
<p>Data licensing</p>	<p>The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.</p>
<p>Researcher/ author support</p>	<p>Questions about complying with this policy should be sent to [email].</p>
<p>Data availability statements</p>	<p>The journal requires authors to include in any articles that report results derived from research data to include a Data availability statement. The provision of a Data availability statement that is compatible with the journal’s research data policy will be verified as a condition of publication. Data availability statements must include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. Where research data are not publicly available, a persistent link to a metadata record or landing page for the data should be provided. Any conditions for accessing the data must be stated in the manuscript. Data Availability statements must take one of the following forms (or a combination of more than one if required for multiple types of research data):</p> <ul style="list-style-type: none"> · The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS] · The datasets generated during and/or analysed during the current study are not publicly available because [REASON WHY DATA ARE NOT PUBLIC] but are available [CONDITIONS FOR OBTAINING ACCESS] · Data sharing not applicable to this article as no datasets were generated or analysed during the current study. <p>More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available at <link></p>

<p>Data formats and standards</p>	<p>The journal encourages authors to share research data using data formats and standards recognised by their research community. Please see FAIRsharing.org for more information on established data sharing formats and standards.</p> <p>The journal prefers research data to be shared in open file formats – those that do not require proprietary software to access - where possible. For example, tabular data should be shared as CSV files rather than XLS files.</p>
<p>Mandatory data sharing (all papers)</p>	<p>The journal requires that all research data that support articles published in the journal, except those covered by the “Definition of exceptions”, must be available in public repositories.</p>
<p>Research data and peer review</p>	<p>Peer reviewers are encouraged to consider a manuscript’s Data Availability Statement (DAS), where applicable. They should consider if the authors have complied with the journal’s policy on the availability of research data, and whether reasonable effort has been made to make the data that support the findings of the study available for replication or reuse by other researchers.</p> <p>For the Data availability statement, reviewers should consider:</p> <ul style="list-style-type: none"> ● Has an appropriate DAS been provided? ● Is it clear how a reader can access the data? ● Where links are provided in the DAS, are they working/valid? ● Where data access is restricted, are the access controls warranted and appropriate? ● Where data are described as being included with the manuscript and/or supplementary information files, is this accurate? <p>For the data files, where available, reviewers should consider:</p> <ul style="list-style-type: none"> ● Are the data in the most appropriate repository? ● Were the data produced in a rigorous and methodologically sound manner? ● Are data and any metadata consistent with file format and reporting standards of the research community? ● Are the data files deposited by the authors complete and do they match the descriptions in the manuscript? ● Do they contain personally identifiable, sensitive or inappropriate information?
<p>Data Management Plans</p>	<p>The journal encourages authors to prepare Data Management Plans before conducting their research and encourages authors to make those plans available to editors, reviewers and readers who wish to assess them.</p>

Standard text for policy 6

<p>Policy summary for authors</p> <p>The journal will require data sharing for all types of research data and peer review of research data. The journal will ensure that its policy on data sharing is adhered to and, for all papers, the decision to publish will be affected by whether or not authors share their research data.</p> <p>Required</p> <ul style="list-style-type: none"> • Data sharing via repositories for all research data • Data availability statements • Data sharing on request • Data citation • Share data with editors and peer reviewers <p>Optional</p> <ul style="list-style-type: none"> • Prepare and share Data Management Plans 	
<p>Policy summary for Editors</p> <p>The journal will require data sharing for all types of research data and peer review of research data. The journal will ensure that its policy on data sharing is adhered to and, for all papers, the decision to publish will be affected by whether or not authors share their research data.</p> <p>Required</p> <ul style="list-style-type: none"> • Data sharing via repositories for all research data • Data availability statements • Support readers and reviewers in making reasonable requests for data • Respond to questions about this policy • Data citation • Peer reviewers assess compliance with research data policy • Peer reviewers assess research data files <p>Optional</p> <ul style="list-style-type: none"> • None 	
Feature	Text
Definition of research data	<p>This policy applies to the research data that would be required to verify the results of research reported in articles published in the journal. Research data include data produced by the authors (“primary data”) and data from other sources that are analysed by authors in their study (“secondary data”).</p> <p>Research data includes any recorded factual material that are used to produce</p>

	<p>the results in digital and non-digital form. This includes tabular data, code, images, audio, documents, video, maps, raw and/or processed data.</p>
<p>Definition of exceptions</p>	<p>Research data that are not required to verify the results reported in articles are not covered by this policy.</p> <p>This policy does not require public sharing of quantitative or qualitative data that could identify a research participant (“personal data”) unless participants have consented to data release. The policy also does not require public sharing of other sensitive data, such as the locations of endangered species. Personal or sensitive data must be shared in a secure or controlled access way, in agreement with the Editors, ensuring Editors and peer reviewers may access the data. Methods for sharing sensitive or personal data include:</p> <ul style="list-style-type: none"> ● Deposition of research data in an approved controlled access repositories and sharing metadata publicly about the research data ● Anonymisation or deidentification of data before public sharing <p>In these cases the procedures and conditions for accessing your research data must be included in your manuscript.</p>
<p>Embargoes</p>	<p>Embargoes on data sharing are not permitted.</p>
<p>Supplementary materials</p>	<p>Sharing research data as supplementary information files is not permitted. The journal will require authors to deposit these in an approved repository as a condition of publication.</p>
<p>Data repositories</p>	<p>Research data must be shared via data repositories. Please see <recommended repository list> or https://repositoryfinder.datacite.org/ for help finding research data repositories. The journal will require authors to deposit these in an approved repository before they consider sending their paper for peer review.</p>
<p>Data citation</p>	<p>The journal requires authors to cite any publicly available research data in their reference list, and will verify this as a condition of publication. References to datasets (data citations) must include a persistent identifier (such as a DOI). Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style.</p>
<p>Data licensing</p>	<p>The journal encourages research data to be made available under open licences that permit reuse freely. The journal does not enforce particular licenses for research data, where research data are deposited in third party repositories. The publisher of the journal does not claim copyright in research data.</p>
<p>Researcher/ author support</p>	<p>Questions about complying with this policy should be sent to [email].</p>

<p>Data availability statements</p>	<p>The journal requires authors to include in any articles that report results derived from research data to include a Data availability statement. The provision of a Data availability statement that is compatible with the journal’s research data policy will be verified as a condition of publication. Data availability statements must include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. Where research data are not publicly available, a persistent link to a metadata record or landing page for the data must be provided and any conditions for accessing the data stated in the manuscript. Data Availability statements must take one of the following forms (or a combination of more than one if required for multiple types of research data):</p> <ul style="list-style-type: none"> · The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS] · The datasets generated during and/or analysed during the current study are not publicly available because [REASON WHY DATA ARE NOT PUBLIC]. Metadata are available at [PERSISTENT WEB LINK] and data are available under the following conditions: [CONDITIONS FOR OBTAINING ACCESS] · Data sharing not applicable to this article as no datasets were generated or analysed during the current study. <p>More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available at <link></p>
<p>Data formats and standards</p>	<p>The journal requires authors to share research data using data formats and standards recognised by their research community. The sharing of data using appropriate formats and standards will be verified as a condition of publication. Please see FAIRsharing.org for more information on established data sharing formats and standards.</p> <p>The journal prefers research data to be shared in open file formats – those that do not require proprietary software to access - where possible. For example, tabular data should be shared as CSV files rather than XLS files.</p>
<p>Mandatory data sharing (all papers)</p>	<p>The journal requires that all research data that support articles published in the journal, except those covered by the “Definition of exceptions”, must be available in public repositories.</p>

<p>Research data and peer review</p>	<p>Peer reviewers are required to consider a manuscript's Data availability statement (DAS), and the supporting research data</p> <p>Peer reviewers must consider if the authors have complied with the journal's policy on the availability of research data, and whether reasonable effort has been made to make the data that support the findings of the study available for replication or reuse by other researchers.</p> <p>For the Data availability statement, reviewers should consider:</p> <ul style="list-style-type: none"> ● Has an appropriate DAS been provided? ● Is it clear how a reader can access the data? ● Where links are provided in the DAS, are they working/valid? ● Where data access is restricted, are the access controls warranted and appropriate? ● Where data are described as being included with the manuscript and/or supplementary information files, is this accurate? <p>For the data files, where available, reviewers should consider:</p> <ul style="list-style-type: none"> ● Are the data in the most appropriate repository? ● Were the data produced in a rigorous and methodologically sound manner? ● Are data and any metadata consistent with file format and reporting standards of the research community? ● Are the data files deposited by the authors complete and do they match the descriptions in the manuscript? ● Do they contain personally identifiable, sensitive or inappropriate information?
<p>Data Management Plans</p>	<p>The journal encourages authors to prepare Data Management Plans before conducting their research and encourages authors to make those plans available to editors, reviewers and readers who wish to assess them.</p>

References

1. Jones, L., Grant, R. & Hrynaszkiewicz, I. Implementing publisher policies that inform,

- support and encourage authors to share data: two case studies. *UKSGI* **32**, (2019).
2. Vines, T. H. *et al.* Mandated data archiving greatly improves access to research data. *FASEB J.* **27**, 1304–1308 (2013).
 3. Naughton, L. & Kernohan, D. Making sense of journal research data policies. *UKSGI* **29**, 84–89 (2016).
 4. Hrynaszkiewicz, I. *et al.* Standardising and harmonising research data policy in scholarly publishing. *IJDC* **12**, 65 (2017).
 5. Hrynaszkiewicz, I., Simons, N., Goudie, S. & Hussain, A. Research Data Alliance Interest Group: Data policy standardisation and implementation. at <https://www.rd-alliance.org/groups/data-policy-standardisation-and-implementation>
 6. Data Policy Types | Authors | Springer Nature. at <https://www.springernature.com/gp/authors/research-data-policy/data-policy-types/12327096>
 7. Research Data Guidelines. at <https://www.elsevier.com/authors/author-resources/research-data/data-guidelines>
 8. Data Sharing & Citation | Wiley. at <https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/data-sharing-citation/index.html>
 9. PLOS Data availability policy. at <https://journals.plos.org/plosone/s/data-availability>
 10. Hodson, S. & Molloy, L. Current Best Practice for Research Data Management Policies. (2015). at https://zenodo.org/record/27872#.WlczsIVI_IV
 11. Transparency and Openness Promotion (TOP) Guidelines. (2014). at https://osf.io/xd6gr/?_ga=2.251468229.297610246.1542300800-587952028.1539080384
 12. Hrynaszkiewicz, I., Simons, N., Goudie, S. & Hussain, A. Journal and publisher research

data policy master framework (RDA IG draft output) DRAFT v1.2 Feb 2018. at

<https://docs.google.com/document/d/1DTAfOKkE1a2n2f_1hGcrXlrw-5Tq_AL5tk-ju8B82_E/edit?usp=sharing>

13. Cousijn, H. *et al.* A data citation roadmap for scientific publishers. *Sci. Data* **5**, 180259 (2018).
14. Murphy, F. & Samors, R. J. Belmont Forum Data Accessibility Statement Policy and Template - Endorsed 18 October 2018. (2018).
15. *Concordat on Open Research Data*. (2016). at <<http://www.rcuk.ac.uk/documents/documents/concordatonopenresearchdata-pdf/>>
16. McCarthy, J. Supplementary online material: potential and precautions. *Augment. Altern. Commun.* **25**, 4–6 (2009).
17. Santos, C., Blake, J. & States, D. Supplementary data need to be kept in public repositories. *Nature* **438**, 738 (2005).
18. Evangelou, E., Trikalinos, T. A. & Ioannidis, J. P. A. Unavailability of online supplementary scientific information from articles published in major journals. *FASEB J.* **19**, 1943–1944 (2005).
19. Stuart, D. *et al.* *Whitepaper: Practical challenges for researchers in data sharing*. (2018). at <https://figshare.com/articles/Whitepaper_Practical_challenges_for_researchers_in_data_sharing/5975011>
20. Schmidt, B., Gemeinholzer, B. & Treloar, A. Open data in global environmental research: the belmont forum's open data survey. *PLoS ONE* **11**, e0146695 (2016).
21. Piwowar, H. A. & Vision, T. J. Data reuse and the open data citation advantage. *PeerJ* **1**, e175 (2013).
22. Piwowar, H. A., Day, R. S. & Fridsma, D. B. Sharing detailed research data is associated

- with increased citation rate. *PLoS ONE* **2**, e308 (2007).
23. Data Sharing Effect on Article Citation Rate in Paleoceanography. at https://figshare.com/articles/Data_Sharing_Effect_on_Article_Citation_Rate_in_Paleoceanography/1222998/1
 24. Henneken, E. A. & Accomazzi, A. Linking to Data - Effect on Citation Rates in Astronomy. 4 (2011). at <http://arxiv.org/abs/1111.3618>
 25. Dorch, B. F., Drachen, T. M. & Ellegaard, O. The data sharing advantage in astrophysics. *IAU* **11**, 172–175 (2015).
 26. Glad You Asked: A Snapshot of the Current State of Data Citation. at <https://blog.datacite.org/citation-analysis-scholix-rda/>
 27. Open for business. *Sci Data* **4**, (2017).
 28. STM & ALPSP. Databases, data sets, and data accessibility – views and practices of scholarly publishers. (2006).
 29. Astell, M., Hrynaszkiewicz, I., Grant, R., Smith, G. & Salter, J. Have questions about research data? Ask the Springer Nature Helpdesk. in at https://figshare.com/articles/Providing_advice_and_guidance_on_research_data_a_look_at_the_Springer_Nature_Helpdesk/5890432
 30. Announcement: Where are the data? *Nature* **537**, 138 (2016).
 31. Availability of data, materials, code and protocols : authors & referees @ npg. at <https://www.nature.com/authors/policies/availability.html>
 32. Sansone, S.-A. *et al.* FAIRsharing as a community approach to standards, repositories and policies. *Nat. Biotechnol.* **37**, 358–367 (2019).
 33. Grant, R. & Hrynaszkiewicz, I. The impact on authors and editors of introducing Data Availability Statements at Nature journals. *IJDC* **13**, 195–203 (2018).

34. Let referees see the data. *Sci. Data* **3**, 160033 (2016).
35. Understanding our data sharing policies - Author Services. at
<<https://authorservices.taylorandfrancis.com/understanding-our-data-sharing-policies/>>
36. Data sharing - BMJ Author Hub. at <<https://authors.bmj.com/policies/data-sharing/>>
37. Checklists work to improve science. *Nature* **556**, 273–274 (2018).
38. Neylon, C. Building a culture of data sharing: policy design and implementation for research data management in development research. *RIO* **3**, e21773 (2017).