AAP Professional and Scholarly Publishing Division and Int'l Association of STM Publishers
Submission in response to NIH Request for information (RFI): Including preprints and interim research products in NIH applications and reports.

Name

John Tagler for AAP-PSP and Michael Mabe for STM

Email

jtagler@publishers.org, mabe@stm-assoc.org

Name of Organization

AAP Professional and Scholarly Publishing Division and Int'l Association of STM Publishers

Types of interim research products your or your organization create/and or host.

Publishers fulfill a critical role in research and scholarly communication, transforming preliminary scholarly communications and other interim research products through peer review, editing, dissemination, and preservation among many other services into a permanent and reliable record of research and discoveries. Our primary role in the research ecosystem is providing high-quality peer-reviewed published articles about research – the "version of record" – and preserving it with integrity for the future. However, publishers are also constantly innovating and providing new services for the research community, which include the hosting, networking and dissemination of some interim research products, such as preprints and registered protocols. As defined in the RFI and used in common practice in publishing, "preprint" refers to a document prepared by a researcher that describes their research, but which has not been influenced by peer review or third-party editing. The sharing of preprints between collaborating colleagues is a long-established practice within certain disciplines in the scholarly community, and, in recent years, sharing preprints on public websites has expanded beyond early adopters like the high-energy physics community. Publishers have established policies regarding preprints, which may vary by discipline and individual organizations. Some publishers have developed preprint servers for the research communities they serve, while others encourage the use of established preprint servers like arXiv.org. As preprints are developed into published articles, additional versions of a manuscript may be prepared and disseminated, which are neither the preprint nor the version of record. This can create a problem in versioning and ascribing provenance. Although these intermediate versions are interim in the sense that they are not the final and permanent version, they are also closely related to the final publication. For example, many publishers make available to the public an "article in press," which is a version of an article upon acceptance or shortly thereafter but before its official publication date. To best support the integrity of the system and to avoid confusion in the record and for public use, these should not be considered interim research products. Some publishers are involved in the development of a relatively new form of research product, designed to increase the transparency of research. The registered research protocol, also referred to as a registered report, is a peer-reviewed description of the goals and methods to be used in a research project before the project is undertaken. Often the registration of the protocol contains with it an agreement by both the researcher and the publisher to publish an article describing the results of the research upon its conclusion. As these are peer reviewed and often published, it is not clear whether these fit NIH's definition of interim research product, but they are interim in the sense that they are produced and disseminated before the research is completed. Although registered reports are not widely adopted at this time, many organizations are exploring their use and definition, for example, https://cos.io/rr/.

Feedback on what are considered to be interim research products, and how they are used in your field.

Preprints are the primary interim research product historically considered by publishers. In general, "preprint" refers to an initial draft of a paper that is subsequently submitted to a journal (or expected to be); therefore, its status as a preprint depends on the availability and sustainability of journals to confer this status. Preprints are viewed as a rapid communication vehicle, but preprints are not the only way that articles about research can be quickly communicated. Many journals have very short turn-around from submission to publication, particularly in biomedicine. In these cases, where authoritative communications about research are rapidly and readily available, preprints are not necessary to disseminate information about cutting edge research. Communities of practice differ significantly around the use of preprints. In

some communities, the distribution of preprints works alongside peer review to improve and refine the content and presentation of research results. In others, preprints serve as a marker for the future publication. Regardless, almost all research communities depend on journals and their publishers to ensure the integrity of research and provide for long-term preservation of the scholarly record, which preprints – by their very nature as "interim" – cannot provide. Preprints can often appear in several versions, calling into question version citability and permanence. While the RFI explicitly refers to preprints as one interim research product that NIH is considering, these issues must be addressed in policy and explicitly developed to avoid confusion and the unnecessary proliferation of versions, which undermines the rigor and reproducibility of science and the public trust in its validity. We encourage NIH to work with all stakeholders to develop clear definitions and expectations regarding what is – and what is not – a preprint, and reach agreement on how to link preprints to any published article to which it is related. We strongly recommend that any preprint that does go on to be published be required to permanently link to the version of record in due course. As one example of why such clarification is needed, some research communities, including in the legal arena and some social sciences, distribute drafts of papers that are never developed into articles that are submitted to journals, and in fact it may not be clear to the author whether they ever will. Whether such drafts should be considered "preprints" or labeled as another type of interim research product is a question that NIH must consider as it pursues this policy. As noted in the response to question 1, a relatively new form of research product, designed to increase the transparency of research, is the registered research protocol, often referred to as a registered report. As these are peer- reviewed and often published, it is not clear whether these fit NIH's definition of interim research product, but they are interim in the sense that they are produced and disseminated before the research is completed. Registered reports may increase the transparency and reproducibility of research by emphasizing the importance of methods and hypothesis setting at the start of the research process.

Insight on how particular types of interim research products might impact the advancement of science.

If evaluated in context, preprints can be valuable in some areas of research. For example, they are widely used in physics, mathematics, and the social sciences, where there are established community standards about their use and status. They may accelerate communication about new research and enable researchers to receive feedback before their results are ready for full publication. However, they may be perceived as less useful and are less prevalent in fields – such as biomedicine – where publication occurs more swiftly and openly. Making preprints more widely available has the potential to accelerate the advancement of science as long as proper cautions about the use of preliminary reports are clearly visible and understandable even to non-specialists – especially where such reports may impact human health and welfare.

Preprints can also be used to provide a timestamp to establish priority and precedence in advance of full publication. In some communities, the sharing of preprints has advanced science by helping researchers get feedback on their work before submission or publication, improving the communication, analysis, or the research itself. Discussions that can occur over a preprint may help researchers form connections or collaborations where an author chooses to do so. The experience of the use of preprints in various fields has shown that, while some improvements may occur due to the sharing of preprints, the services provided by publishers – professional editing and copyediting and formal, usually anonymous, peer review, plagiarism detection, archiving in perpetuity, etc. – are critical to advancing and ensuring the integrity of science and cannot be replaced. As long as users of preprints understand their limitations, preprints can live side-by-side with authoritative publications and may create valuable interactions. As noted, in addition to their potential to benefit science, there are potential downsides to the widespread availability of preprints in biomedicine. Publishers play a critical role in ensuring the quality and integrity of research, identify and tackle plagiarism, and provide a record upon which sound science and medical practice can be built; care must be taken to ensure that preprints are not misused. If preprints are misused or their status is misunderstood, it may hinder the advancement of science by leading researchers down dead end paths. In the case of human health, there could also be significant negative effects if a reader believes preprints are as accurate as the published literature, despite the lack of peer review. Such concerns are not limited to the use of preprints by the public, but could also negatively impact clinical practice. This is of particular concern where small errors in an interim research product could lead clinicians, patients, or caregivers to provide incorrect dosages to patients. Appropriate safeguards, citation standards, and clear labelling and linking could mitigate these concerns. Such standards must be developed and required to prevent public harm. We encourage NIH to investigate the proper balance between the promotion of preprints and the reliance on the published record, particularly for journals

where preprints may not provide a significant time advantage over the published article.

Feedback on potential citation standards.

Although there are not currently widely accepted citation standards around preprints, there are many organizations (NISO, RDA, others) already working on standards for linking and citing research products, and NIH would be well served by working in partnership with these organizations to support standards that can be used broadly and across disciplines.

Collaborative projects such as SCHOLIX (http://www.scholix.org/) are the best way forward to ensure the widespread consistency and adoption of standards. Any citation standards developed should be consistent with those used for formal publications, but clearly distinguish between an interim product and the version of record. We support ASAPbio's recommendation that preprints be clearly identified as not yet peer reviewed, both in the preprint itself and in any citation thereto. Preprints that go on to be published can be cited using similar standards to formal publications, and we strongly recommend that any preprint that does go on to be published be required to permanently link to the version of record in due course. Both the preprint document and any citation should be clear as to its status: whether the article has been submitted and/or published as well as whether it has otherwise been reviewed or vetted. Such tagging and linking would help readers properly situate the document in the literature and ensure the integrity of the scholarly record. Permanent and updated links, in particular, could identify where a preprint has been published, and emphasize where the preprint was either not submitted or was rejected in the peer review process. Our member publishers participate in a variety of permanent archiving arrangements that ensure the long-term preservation, citability, and discoverability of research; it would be counterproductive to recreate such arrangements for the preprint at the interim stage. Standards should be promulgated to enable and encourage the use of existing infrastructure, where appropriate.

Insight on the possible need and potential impact of citing interim products on peer review of NIH applications.

Publishers are bringing high-quality, peer-reviewed articles to the public with increasing speed in the digital age. These versions of record, as the name implies, are the permanent record of scholarly discovery. Although the RFI is correct that journal submissions under review are not always public, versions of record are very much public and permanent. While under review, submissions are still subject to change and correction, or rejection for inaccuracy. The need for dissemination of interim products may vary by research community. As one example, disciplines with rapid publication times may have less of a need for making preprints available or citable than disciplines with slower publication times. We hope that this RFI receives significant input from the stakeholder community about the need and appropriateness of using preprints in this way in various research communities, as one size does not fit all, even in those research areas funded by NIH. We should be careful not to sacrifice the reliability of the published record to provide a citable reference merely days or weeks before the article may be available; in fact, in some research areas it is acceptable to cite an article in press, and such a citation may be preferable to linking to a preprint. Even in those disciplines where there may be community desire to use preprints in this way, care still needs to be taken. As preprints do not undergo rigorous peer review, citing them carries a greater risk of promoting flawed science and information potentially harmful to human health and welfare. Although much of the research funded by NIH is of the highest quality, errors are still known to occur and appear in early drafts of papers that are often caught in journals' editorial and peer-review process. As noted earlier, we therefore recommend that any preprint, and all citations thereto, include an appropriate indication of the status of the preprint. Such warnings or disclaimers would greatly mitigate any potential negative impact of their use. An additional impact of allowing applicants to cite interim products will occur on the review process itself. Staff and external reviewers will need to spend extra time on assessing the quality and value of interim products, as they have not gone through a peer-review process that would help with this assessment. In particular, NIH will need to consider how to appropriately provide guidance on the reviewers' consideration of interim research products in the proposal process.

Advice on how NIH reviewers might evaluate citations of interim research products in applications.

Citations of interim research products will be very difficult to evaluate if the products themselves have not been peer reviewed by outside reviewers or the NIH reviewers. Proxies – such as the quality of the journal to which an article has been submitted – are of limited use in evaluating preprints, and are likely meaningless before an article has been reviewed and accepted or rejected. We therefore recommend that evaluators look to whether

the preprint has been accepted or published subsequent to the proposal submission, and use that citation, where available, in place of the preprint. Where the preprint has not yet been accepted or published, care must be taken to evaluate the contents of the preprint. NIH should consider the additional burdens on reviewers of such requirements. As noted above, training and explicit guidelines will be necessary to ensure that preprints and other interim research products are used appropriately in the evaluation process. As research practice and technology evolve, interim research products may take on new forms within discipline communities. Ensuring that products are defined and evaluated with this flexibility in mind will be important in ensuring appropriate recognition for interim research products that are shared openly — whether these may be preprints, abstracts, posters, conference/meeting presentations, data or other artifacts.

Any other relevant information.

Publishers appreciate NIH's effort to provide more recognition to researchers for the discoveries they make; providing such recognition is part of why scholarly and professional publishers exist: our success in providing high-quality outlets for authoritative communications about scholarly research is part of why our journals are embedded in the research enterprise and relied upon by evaluators to assess research progress. We support NIH's aims in this endeavor, and have some additional thoughts for consideration as NIH seeks to strike the right balance in this principle. Academic freedom is a bedrock principle of the research enterprise, and it is important that whatever policy is promulgated that researchers continue to be free to choose where, when, and how they share communications about their research. Therefore, we recommend that there be no new requirement for listing interim products in their research, explicit or even implicit. Given the difficulty that evaluators may have in assessing interim research products that are not linked to a formal publication, we also caution that any policy be calibrated to disincentivize "padding" applications with long lists of interim products that it may be difficult to evaluate and may never be put forward for publication. Given publishers' experience with and investments in maintaining the accuracy and availability of the authoritative communications about scientific advances, we strongly emphasize that systems and incentives need to be put in place to ensure that preprints are appropriately labeled and linked. NIH should work with all stakeholders to develop clear definitions and expectations regarding what is – and what is not – a preprint, and to reach agreement on how to link preprints to any published article to which it may be related. Appropriate tagging and linking is a burden that should not be left to the author or person citing the interim product. Ensuring accurate updating is essential to maintain the reliability of the system, as the proliferation of outdated, misleading or incorrect information on the internet attests. The potential relationship of the research that NIH funds to human health and welfare makes it even more vital that such systems be put in place. Addressing these issues will have significant costs, which must be carefully considered by NIH in the face of constrained funding for research. These include the costs of making preprints available, the costs of the additional burdens on researchers and administrators, and the costs of systems needed to update and tag interim research products to ensure they are understood and used properly. These costs and burdens must be evaluated in the context of NIH's and NLM's broader mission to ensure adequate resources are available for advancing biomedical research. We strongly encourage NIH to use existing systems and infrastructure and engage in collaboration with other stakeholders, where possible, to minimize these costs and burdens.