F1000Research
Open for Science

Guide to open science publishing

For the online version of this guide please visit **f1000.com/open**

Introduction

What is the history of open access? Where should we share our data? What are the benefits of peer review? Who is using post-publication peer review? How can I get involved with open science?

To answer these and other key questions, *F1000Research* has put together this guide to open science publishing. The first four chapters each focus on one of the underlying principles of *F1000Research*, and the last chapter addresses other areas of open science.

We hope this will be a useful resource for researchers, students, and others with an interest in the increasingly open nature of scientific research and publishing.

Finally, we'd like to express our thanks to Peter Kraker, Fabiana Kubke, Ross Mounce, Liz Neeley, Anna Sharman, Lenny Teytelman, Kaitlin Thaney and Liz Wager for useful feedback on individual chapters of this document.

About the Author

Eva Amsen
Outreach Director at *F1000Research*

Eva holds a PhD in Biochemistry from the University of Toronto, and is interested in all forms of communication between researchers, from hallway conversations to academic papers.

Before joining *F1000Research*, she worked at the journal *Development*, where she launched and ran the Node, a community website for developmental biologists.

Open access explained

Although the basic idea of open access is indeed to give everyone access to the contents of a research paper, the term "open access" implies more than just making the content free.

A short history of open access

Open access as we know it was defined in a meeting in Budapest in 2001. This meeting produced the Budapest Open Access Initiative: a statement of principles about open access publishing.

The resulting document, published in 2002, described the opportunities that the internet could provide in opening up scientific literature:

> *"...lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge."*

Despite this occasional lofty wording, the document is very clear on the definition of open access:

> *"By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."*

At this point there were not yet many open access publishers, although researchers could make their manuscripts available by self-archiving. Within the life sciences, BioMed Central was the first open access publisher to launch in 2000. (It was founded by Vitek Tracz, who went on to found *F1000Research*.) The Public Library of Science announced their intent to publish shortly after, in 2001.

Open access today

Since these early open access publishers, many others have followed suit, and open access is more popular than ever. In August 2013, the European Commission announced that

> *"open access is reaching the tipping point, with around 50% of scientific papers published in 2011 now available for free."*

The European Union's research framework Horizon 2020 requires all research they fund to be made available via open access. Similar requirements are made by large funding agencies around the world, such as the Wellcome Trust in the UK and the National Institutes of Health in the USA.

Green and Gold open access

How can researchers comply with these funder regulations? There are currently two possible open access routes: Green and Gold.

- ***Green open access*** is archiving of accepted manuscripts in accessible repositories, for example in their institutional repository or in PubMed Central. While this allows researchers to publish in any journal they want and deposit later, the system has some limitations: Some journals only allow the archiving of a final accepted manuscript, not of the published and formatted paper. Some journals open up access to all their archived articles after a certain time period, but in other cases authors will have to remember to deposit their own paper, which can be a time consuming process.
- ***Gold open access*** is publisher-mediated open access. The benefit of this is that the article is immediately made open access, and authors don't have to take any extra steps, but there can be a cost associated with it. Usually the entire journal will be available as an open access journal, but some journals operate a hybrid model, where researchers can pay to publish an open access article in an otherwise non-open-access journal.

Three misconceptions about open access

As the above summary illustrates, open access has a distinct definition, but can be applied in different ways. That can lead to some confusion, and there are a few misconceptions about open access floating around, such as the following.

1 Open access just means "free to read"

Not exactly. Although the basic idea of open access is indeed to give everyone access to the contents of a research paper, the term "open access" implies more than just making the content free. It also requires making the material available for others to re-use, and allowing content mining (e.g. for meta-analysis).

Usually, open access articles are accompanied by a Creative Commons licence that describes the details of what can be done with the content of the paper. Most of these licences require giving credit to the authors, who generally retain copyright. (See more about Creative Commons licences in the "further reading" section below).

2 Gold open access means "author pays"

No, it just means that it's publisher-mediated open access. Because the publisher does not rely on subscription fees from readers and libraries for those articles, they cover their costs in other ways. Some smaller journals can be funded entirely by institutes or societies. Larger open access publishers often use article processing charges (APCs), which allow them to scale their business model once they start receiving more submissions. The Directory of Open Access Journals (DOAJ) lists which model each journal uses.

Although APCs are applied on a per-article basis, in reality many authors do not pay these fees out of pocket. Funders and institutes who require open access publication often support authors and cover the cost, and many publishers will waive APCs for low income countries via the HINARI programme or on a case-by-case basis for specific situations of economic difficulty.

3 Open access implies bad quality

No. Whether an article is free to access or hiding behind a paywall says nothing about the quality of the research itself or about the peer review carried out on the paper.

This misconception comes from the fact that there is indeed a small group of so-called "predatory" publishers who are charging researchers to publish articles in their journals for the sole purpose of making money, without considering the scientific quality and often without even inviting peer reviewers to look at the papers. On the surface, this charge for publication may resemble the APC model used by many open access journals, but it is different. Reliable journals use APCs to cover the cost of managing the editorial and peer review process, and to develop new features to support their authors and readers, whereas "predatory" journals don't invest in these processes.

But how can you distinguish a good journal from a bad one? First of all, you can check if you know people who have published in this journal, or members of the editorial board. Has the journal attended or sponsored conferences or supported other initiatives? Is the journal a member of the Committee on Publication Ethics (COPE)? Is the journal included in scholarly databases such as PubMed? So far, that all applies equally to both open access and subscription journals. For open access journals, there is even an extra level of scrutiny: The Open Access Scholarly Publishers Association (OASPA) regularly re-reviews its members, and publishers must comply with very strict guidelines to remain a member, so you can check their member list. The DOAJ has also recently tightened their criteria for inclusion, and has removed several journals that do not fit these criteria. Their member list is easy to search, and provides detailed information about each publication.

Finally, if a journal has a transparent peer review model, where names of reviewers and/or content of referee reports or editorial decision letters are made public, you can see for yourself what the peer review process looks like and make an informed decision about the journal.

REFERENCES & RESOURCES

European Commission: Open Access to Research Publications Reaching 'Tipping Point' (21 August 2013) http://europa.eu/rapid/press-release_IP-13-786_en.htm

Budapest Open Access Initiative declaration (February 14, 2002) http://www.budapestopenaccessinitiative.org/read

Wellcome Trust open access policy http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD002766.htm

NIH public access policy https://publicaccess.nih.gov/

Directory of Open Access Journals (DOAJ) http://doaj.org/

Richard van Noorden: Open Access Website gets tough , 6 August 2014, *Nature*; 512(7512) 17-17

Open Access Scholarly Publishers Association (OASPA) http://oaspa.org/

OASPA's second statement following the article in Science entitled "Who's Afraid of Peer Review?" (11 November 2013) http://oaspa.org/oaspas-second-statement-following-the-article-in-science-entitled-whos-afraid-of-peer-review/

Committee on Publication Ethics (COPE) http://publicationethics.org/

FURTHER READING

Claire Redhead/OASPA: Why CC-BY? (23 October 2012) http://oaspa.org/why-cc-by/

Paul Klimpel (2013) Free Knowledge based on Creative Commons Licenses. Consequences, Risks, and side-effects of the license module Non-Commercial – NC, Wikimedia Germany

The seer of science publishing. Interview of Vitek Tracz by Tania Rabesandratana, 4 October 2013, *Science* 342 (6154), 66-67 http://www.sciencemag.org/content/342/6154/66.full

Richard Poynder: The state of open access. http://richardpoynder.co.uk/the-state-of-open-access.html

Stephen Curry: Open Access - Yes you can. (20 April 2014) http://occamstypewriter.org/scurry/2014/04/20/open-access-yes-you-can/ .

The merits of open peer review

Anonymous reviewers can be biased against the authors of the paper, and lean toward rejection or acceptance for unscientific reasons. Often, the closest "peers" in someone's area of research are also that researcher's direct competitors!

History of traditional (closed) peer review

Even though scientific publishing has been around since the 17th century, formal peer review of submitted articles by external academics is relatively new. The journals *Science* and *JAMA*, for example, introduced formal peer review in the 1940s, and *Nature* didn't introduce it until 1967.

The peer review system adopted in the 20th century has now become the norm for many journals. It involves an editor (usually a practising researcher, but sometimes a journal staff member in the case of journals like *Nature*) sending out a paper to a few experts in the field, who then provide comments for the paper's authors. Although the reviewers can generally see who the authors are, they themselves remain anonymous to the author, and only the editor knows everyone's identity.

Problems with traditional, semi-blind, peer review

This "single-blind" system is not without problems. Anonymous reviewers can be biased against the authors of the paper, and lean toward rejection or acceptance for unscientific reasons. Often, the closest "peers" in someone's area of research are also that researcher's direct competitors! One solution is to remove the authors' names from the manuscript, but this double-blind system is not fool-proof, and a reviewer will still often recognize which lab a paper comes from. In addition, any bias towards competitors of the reviewer still remain, even if that competitor is anonymised.

Another drawback of traditional peer review is that the referee reports are visible only to the authors and the editor. Nobody else can see what the reviewers thought of the paper. Especially in situations where reviewers disagree, and a single editor makes the final decision, it can be very informative to see what the reviewers thought of an article, and whether the editor's decision was in line with their opinion. Reviewers are usually in a position to put the work in a broader context of the field, and often mention this context in their reports. They can also point out where the work could be expanded into new areas, and may still have some lingering questions. All of this is useful for everyone to read – not just the authors. It's also important to remember that not all journals use the same criteria for publication. Some journals may turn a great paper down just because it doesn't fit the scope of the journal. Other journals publish all sound science, including some papers that get extremely high praise in the referee reports.

A timeline of open and transparent review

Within the life sciences in particular, several journals have opened their peer review process to address some of the issues discussed above. Sometimes this involves publicly naming reviewers and/or editors. Other journals publish some or all reviewer comments.

1999	After studying various peer review models, *BMJ* starts revealing reviewer names to authors.
2000	BioMed Central launches, and soon after that starts including reviewer names and pre-publication history for published articles in all medical journals in their BMC series of publications.
2001	*Atmospheric Chemistry and Physics* introduces a system where manuscripts are placed online as a "discussion paper", which is archived with all comments and reviews, even before approved and peer-reviewed articles appear in the journal.
2006	Launch of *Biology Direct*, which includes reviewer comments and names with published articles.
2007	Frontiers launches, and includes reviewer names with articles.
2010	*EMBO Journal* starts publishing review process file with articles. Editors are named, but referees remain anonymous.
2011	*BMJ Open* launches, and includes all reviewer names and review reports with published articles.
2012	Several journals launch with an open peer review model: » ***GigaScience*** - publishes pre-publication history with articles and names reviewers (opt-out system) » ***PeerJ*** - Peer review reports published with author approval, reviewer names published with reviewer permission. » ***eLife*** - Decision letter published with author approval. Reviewers anonymous. » ***F1000Research*** - All peer review reports and reviewer names are public, and appear after article is published online.
2014	More journals open their peer review process: » ***Science Open*** launches journal with an open post-publication review model. » ***BMJ*** - moves to a fully open peer review model, where reports and reviewer names are published with each article.

Benefits of open review

Benefits for authors and readers

- Author can see who reviewed their work
- Reviewer comments put paper in context, which is useful additional information for readers
- Reduces bias among reviewers
- More constructive reviews
- Published reports can serve as peer review examples for young researchers

Benefits for reviewers

- Shows the reviewer's informed opinion of the work
- Demonstrates experience as a reviewer
- Can take credit for the work involved in conducting the review

To make it easier for referees to take credit for their work, some journals, including *F1000Research*, now provide unique identifiers (DOIs) for referee reports. In addition, *F1000Research* is co-chairing a working group investigating how to include peer review output in ORCiD profiles.

Challenges

Although open peer review is becoming more common, and addresses several of the issues of anonymous review, a few challenges still remain. A study in the early days of open review suggested that naming referees slightly reduced the likelihood of finding reviewers but did not affect the quality of review. Other studies suggest that open review provides more constructive reports.

REFERENCES & RESOURCES

Michael Nielsen: Three myths about scientific peer review (8 January 2009) http://michaelnielsen.org/blog/three-myths-about-scientific-peer-review/

Timeline: History of the journal *Nature* http://www.nature.com/nature/history/timeline_1960s.html

Kathleen Fitzpatrick (2011) The history of peer review, in *Planned Obsolence: publishing, technology, and the future of the academy*, NYU Press

Working double-blind. *Nature* 451, 605-606 (7 February 2008)

Van Rooyen S, Godlee F, Evans S, Black N, Smith R (1999) Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial, *BMJ* 318, 23-7

Maria K Kowalczuk, Frank Dudbridge, Shreeya Nanda, Stephanie L Harriman, Elizabeth C Moylan (2013) A comparison of the quality of reviewer reports from author-suggested reviewers and editor-suggested reviewers in journals operating on open or closed peer review models. F1000Posters, 4: 1252

ORCID & CASRAI Kick-off New Standards Project on 'Peer Review Services'. (7 April 2014) http://casrai.org/435

Van Rooyen S, Delamothe T, Evans SJW (2010) Effect on peer review of telling reviewers that their signed reviews might be posted on the web: randomised controlled trial. *BMJ* 341

Smith R (1999) Opening up BMJ peer review: A beginning that should lead to complete transparency. *BMJ*; 318:4

Shanahan DR, Olsen BR (2014) Opening peer-review: the democracy of science, *Journal of Negative Results in BioMedicine* 13:2

Godlee F (2002) Making Reviewers Visible: Openness, Accountability, and Credit, *JAMA* 287(21):2762-2765

What is 'open peer review', as operated by the medical journals in the *BMC* series? http://www.biomedcentral.com/authors/authorfaq/medical

Atmospheric Chemistry and Physics – publication policy http://www.atmospheric-chemistry-and-physics.net/general_information/publication_policy.html

Koonin EV, Landweber L, Lipman DJ (2006) A community experiment with fully open and published peer review. *Biology Direct* 1:1

Frontiers review system http://www.frontiersin.org/about/reviewsystem

About the *EMBO journal* – review process files http://emboj.embopress.org/about#Review_Process

About *BMJ Open* http://bmjopen.bmj.com/site/about/

About *GigaScience* – Publication and peer review process http://www.gigasciencejournal.com/about#publication

PeerJ – How it works https://peerj.com/about/how-it-works/

eLife review process http://elifesciences.org/about#process

F1000Research – How it works http://*F1000Research*.com/about

Science Open – How does it work? http://about.scienceopen.com/how-does-it-work/

BMJ – Peer review process http://www.bmj.com/about-bmj/resources-authors/peer-review-process

FURTHER READING

A guide to peer review in ecology and evolution, published by the British Ecological Society www.britishecologicalsociety.org/wp-content/uploads/Publ_Peer-Review-Booklet.pdf

Peer review: the nuts and bolts, published by Sense About Science http://www.senseaboutscience.org/resources.php/99/peer-review-the-nuts-and-bolts

How to write a good peer review report, published by *F1000Research* http://*F1000Research*.com/peer-reviewing-tips

Arjun Raj: How to review a paper (19 April 2014) http://rajlaboratory.blogspot.co.uk/2014/04/how-to-review-paper.html

Benos D *et al.* (2007) The ups and downs of peer review. *Advances in Physiology Education,* Vol. 31 no. 145-152

Walbot V (2009) Are we training pit bulls to review our manuscripts? *Journal of Biology* 8:24

Group work image by Eva Amsen. Used with permission.

Understanding post-publication peer review

When you come across a mention of post-publication peer review, always check which "flavour" it is [...] All types of post-publication peer review serve a purpose, but they don't all serve the same purpose.

In the past few years, the phrase "post-publication peer review" has popped up in various discussions about scientific publishing, as either an add-on to, or a replacement of, pre-publication peer review. As is becoming increasingly apparent, pre-publication peer review doesn't pick up all problems that may exist with a manuscript. But what is post-publication peer review, and how can it address these issues? Confusingly, the term can refer to a number of different models, which each work in a different way – some by introducing a new peer review system within a journal, others by providing a platform to discuss any published articles.

Types of post-publication peer review:

1 *Review by formally invited reviewers, after publication of the un-reviewed article*

This type of post-publication peer review is used by publications such as *F1000Research* and the Copernicus journals. Here, peer review is carried out by invited reviewers, like it's done at most journals, but the article is already published online (after an editorial check) before the peer review process starts. Articles that pass peer review are clearly marked as such and are indexed in scholarly databases.

2 *Review by volunteer reviewers, after publication of the un-reviewed article*

This is also a publisher-driven method of post-publication peer review, and also involves articles being published online before peer review, but in this case the reviewers are not invited by the journal. Each publisher may use different criteria to determine who can review, and whether the reviews change the status of the published article. At *Science Open*, a reviewer must have at least five of their own published articles in their ORCID profile. At *The Winnower*, any registered user can leave a review on any of their published articles.

This system closely resembles the commenting system that several journals have implemented in addition to a formal (invited) peer review system, but journals may ask their volunteer reviewers to address specific aspects of the article, as with invited review.

3 *Comments on blogs or third party sites, independent of any formal peer review that may have already occurred on the article*

In recent years, most discussions surrounding post-publication peer review have been about new platforms that allow researchers to comment on published

research articles. PubPeer allows anonymous researchers to comment on any article with a DOI, or those published as preprints in arXiv. PubMed Commons gives authors with at least one of their own publications indexed in PubMed the ability to comment on any other articles in the database but here they cannot be anonymous. Both services were launched to encourage online discussion about scientific articles – a practice already taking place on blogs and on social media, and to a much smaller extent in the comments sections of journal articles itself.

This sort of discussion can be very valuable and highlights some of the problems of traditional anonymous pre-publication peer review. Discussions on social media and on PubPeer successfully identified issues with the STAP paper that was published in Nature in early 2014, and will probably continue to bring to light other issues with high-profile papers in the future.

Sometimes F1000Prime is also mentioned in the context of post-publication peer review. This service uses a network of 10,000 international Faculty Members to recommend articles from the life sciences. However, F1000Prime focuses on recommendations only, as a service to highlight important articles, whereas the other methods focus more on criticism and debate.

Why does the distinction matter?

Post-publication peer review is still new, and is facing several challenges. However, different types of post-publication peer review are not all affected in the same manner. This is highlighted in the list of challenges included below.

Challenges for post-publication peer review

- **Participation** Not all published articles receive comments via systems such as PubPeer and PubMed Commons, although they have been shown to serve as a useful platform for discussion of controversial articles. But for journals that use voluntary post-publication peer review (such as *Science Open* or *The Winnower*), low levels of participation might mean that articles remain unreviewed. Journals that invite reviewers for post-publication peer review (such as *F1000Research* and Copernicus journals) use a system quite similar to "traditional" peer review, and can make sure that all articles are seen by reviewers.
- **Reviewer expertise** Inviting reviewers also allows journals to ensure their reviewers have the adequate expertise to review each particular paper. Some voluntary review systems have also built in a checkpoint to control expertise level: for example, *Science Open* requires reviewers to have five articles in ORCID, and PubMed Commons requires one article in PubMed. However, neither system is able to check that the previously published work of the reviewer is in the field of the article they're commenting on.
- **Fragmentation of discussion** One critique of the variety of post-publication peer review systems is that discussion happens in multiple places. The same article can have comments on the article itself (if that feature is available), in PubPeer, on PubMed Commons, on ResearchGate, on blogs, on Twitter, on F1000Prime and elsewhere.

As these examples show, the different versions of post-publication peer review all deal with different types of challenges, so it is important to clearly distinguish between them.

How to tell which type of post-publication review you're dealing with

As described in the previous chapter, open peer review can mean named reviewers, or public referee reports, or both. In all cases, though, "open peer review" refers to review by invited reviewers. "Post-publication peer review", on the other hand, can be named or anonymous, and reviews can in some cases be written by uninvited reviewers who may not necessarily be literal "peers" in the field.

The many different uses of the phrase are confusing, and currently the only way to know which is being used is to look into each specific case to find out what is meant. For example, at *F1000Research* we noticed that many people assume that "post-publication peer review" means that anyone can provide the formal peer reviews on our articles. In fact, all *F1000Research*'s peer review is carried out by invited reviewers.

When you come across a mention of post-publication peer review, always check which "flavour" it is: are reviewers invited or is review voluntary? Is there a check for reviewer expertise? Are reviews published on the article itself or on a third-party site? Are reviewers anonymous? Does the post-publication review replace traditional peer review or is it an add-on service?

All types of post-publication peer review serve a purpose, but they don't all serve the same purpose.

REFERENCES & RESOURCES

The Guardian view on the end of peer review. *The Guardian* editorial, July 6, 2014 http://www.theguardian.com/commentisfree/2014/jul/06/guardian-view-end-peer-review-scientific-journals

STAP retracted: Two retractions highlight long-standing issues of trust and sloppiness that must be addressed. *Nature* 511, 5–6 (03 July 2014)

Richard Smith: What is post-publication peer review? (6 April 2011) http://blogs.bmj.com/bmj/2011/04/06/richard-smith-what-is-post-publication-peer-review/

Aimee Swartz: Post-publication peer review mainstreamed (22 October, 2013) http://www.the-scientist.com/?articles.view/articleNo/37969/title/Post-Publication-Peer-Review-Mainstreamed/

Post publication peer-review: Everything changes, and everything stays the same, by Bonnie Swoger on the Scientific American blog *Information Culture*, 26 March 2014

F1000Research – How it works http://*F1000Research*.com/about

Copernicus peer review model http://publications.copernicus.org/services/public_peer_review.html

Science Open – How does it work? http://about.scienceopen.com/how-does-it-work/

The Winnower https://thewinnower.com/

What is F1000Prime? http://f1000.com/prime/about/whatis

FURTHER READING

STAP paper discussion on PubPeer https://pubpeer.com/publications/8B755710BADFE6FB0A848A44B70F7D

Summary of the "post-publication peer review" session at the Science Online 2014 conference. https://storify.com/jalees_rehman/discussion-about-post-publication-peer-review-scio

Kelly Servick: Researcher files lawsuit over anonymous PubPeer comments (26 October 2014) http://news.sciencemag.org/scientific-community/2014/10/researcher-files-lawsuit-over-anonymous-pubpeer-comments

What is open data?

The open data movement wants to apply the principles of open data to not just big publicly funded projects like the human genome project, but to all kinds of data however large or small.

Open data in science

"Open data" is a broad concept that doesn't just apply to research data, but also to, for example, the opening up of government data. Many of the underlying ideas are similar: the goal of open data, whether it involves research data or census data, is to make data available to anyone and reusable by anyone for further analysis.

In the sciences, data have not always been easy to come by. Before the internet, journal articles could not feasibly include all the relevant data. If you wanted to use another group's data, you had to ask them for it.

One of the first, and one of the best-known, data sharing projects in biology is the human genome project. The sequencing of the human genome was a massive undertaking, by many researchers across the world. The results of their efforts have greatly advanced many areas of research and healthcare over the past decade and a half, but none of that would have been possible if the genomic sequences had not been widely available. Imagine if every time you wanted to align a DNA sequence or generate PCR primers you had to ask for permission, or worse, pay for use of the information.

Instead, anyone can freely download human genomic data, use it without asking for explicit permission, re-analyse and interpret it, and use it for anything from art projects to teaching to data mining to including versions of it in their own work. That is what open data is.

The open data movement wants to apply the principles of open data to not just big publicly funded projects like the human genome project, but to all kinds of data however large or small.

Why use open data in science?

As illustrated by the human genome example above, opening up research data makes it much easier for other scientists to build upon that work and advance the field. Another advantage of open data is that availability of the underlying data used to generate the figures in a paper makes it easier for others to reproduce the work. This complete transparency of data also discourages researchers from falsifying figures in their publication. Too often people get away with Photoshopped images or duplicating images from different studies, and that is much easier to catch if the underlying data is available. Another important advantage of open data is that it allows datasets to be easily aggregated for meta-studies.

Regulations and principles for data sharing in biomedical research

There are a number of organisations that recommend, regulate, or advise the use of data sharing in research. A few of them are listed here, and each of their websites includes much more information:

- NIH data sharing policies http://www.nlm.nih.gov/NIHbmic/nih_data_sharing_policies.html
- Biosharing – a resource of various policies, standards and databases for the sharing of research data. http://biosharing.org/
- Wellcome Trust Guidance for researchers: Developing a data management and sharing plan http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/index.htm
- Panton Principles for open data in science http://pantonprinciples.org/

Incentives for data sharing

Guidelines are a good first step, but there also needs to be an incentive for researchers to comply with the guidelines. Funders may ask you to share your data, but often lack the resources to ensure that you really do. To overcome a similar lack of (mandated) open access publication, NIH no longer renews grants if the grantholder did not make their work available by open access standards. A similar enforcement for open data is not (yet) in place.

At the moment, if you want to publish work based on certain formats of data, such as microarray screens or protein structures, journal editors will ask you to deposit your data in a suitable database within a certain period of publishing your article, but they often aren't able to follow up and make sure that an author has really deposited their data within the required period after publication.

To encourage data sharing of all types of data, *F1000Research* and (since early 2014) PLOS require their authors to make all data underlying their articles openly available from the moment of publication of the article.

Credit for data publication

Another incentive for data sharing is to provide credit for data. Researchers now generally get professional credit only for published articles. A few journals now allow researchers to publish data sets in the form of a research article, such as *F1000Research* (data notes), *GigaScience*, *Scientific Data* and *Data*. The requirements for such articles (often called "data notes" or "data descriptors") are that they include only a brief introduction, methods, and results – but no interpretation. *F1000Research* has had confirmation from many major publishers that this sort of publication will still allow researchers to later use these same data sets in another, more in-depth, publication.

Over time, a better way to receive credit for data would be for funders and institutes to formally recognize data deposition and open data sharing as a valuable contribution to research, but until that happens, this is one way to formally turn unanalysed data into a tangible credit.

REFERENCES & RESOURCES

All about the human genome project http://www.genome.gov/10001772

Panton Principles for open data in science http://pantonprinciples.org/

Biosharing – a resource of various policies, standards and databases for the sharing of research data http://biosharing.org/

NIH data sharing policies http://www.nlm.nih.gov/NIHbmic/nih_data_sharing_policies.html

F1000Research data sharing information http://*F1000Research*.com/data-preparation

FAQs about PLOS data sharing policy http://www.plos.org/plos-data-policy-faq/

Upcoming Changes to Public Access Policy Reporting Requirements and Related NIH Efforts to Enhance Compliance (November 16, 2012) http://grants.nih.gov/grants/guide/notice-files/NOT-OD-12-160.html

Varsha Khodiyar: Moving to a new research project? Give your old data a new lease of life... (6 May 2014) http://blog.*F1000Research*.com/2014/05/06/give-your-data-a-new-lease-of-life/

Can I publish an analysis of my published dataset in other journals? http://*F1000Research*.com/data-policies

Format of *Scientific Data* data descriptors http://www.nature.com/sdata/for-authors#format

Format of *GigaScience* Data Notes http://www.gigasciencejournal.com/authors/instructions/datanote

FURTHER READING

Wikipedia page on open science data http://en.wikipedia.org/wiki/Open_science_data

Open knowledge foundation's Open definition https://okfn.org/opendata/

Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 (11 December 2013) http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

What is open science?

Major funders are increasingly insisting that underlying data be shared openly, in addition to publications.

Crowdfunded researchers are also often incentivised to share their work with their funders.

Open science is the concept of opening up all aspects of scientific research, to allow others to follow the process and collaborate. There is no formal definition of open science, but it usually incorporates some of the aspects covered in previous chapters, such as open access, open peer review, post-publication peer review, and open data. Additionally, it includes other ways to make science more transparent and accessible during the research process, and we will discuss them here: open notebook science, citizen science, and aspects of open source software and crowdfunded research projects.

Open notebook science

While some groups use online, password protected, lab notebooks to share notes with collaborators, open notebook science takes this a step further by making day-to-day lab notes available in real time. By keeping notes online, rather than in an offline lab notebook, open notebook scientists are giving everyone direct insight into their work, and enabling easier collaboration. For example, you can find open notebooks on OpenWetWare (biology and biological engineering), Open Notebook Science Network (chemistry and other disciplines), or The IPython Notebook (interactive computational science).

This is quite a radical form of openness, and few bench scientists use a fully open notebook system at the moment. The general reluctance of many researchers to share ongoing research data is the fear of being scooped by competing groups in academia or industry, as well as being unsure whether they can still publish the work in their journal of choice afterwards. With the increased use of preprints in biology, more journals are developing guidelines about whether they will consider publishing previously shared research, which may alleviate some of the concerns about putting lab notes online.

Open notebook science shares some similarities with open data: both make the underlying research data public. However, where data sharing can occur at the point of publication of the resulting journal article, or after a conference (e.g. by uploading a conference poster or slides), open notebook science happens “live”: data and methods are made public at the moment of collection.

Citizen science

One of the most traditional uses of collaborative open research data predates the open science movement: citizen science. Here, members of the public, who are often not scientists themselves, participate in the collection (and sometimes analysis) of scientific data.

The oldest running citizen science project is the Audubon Christmas Bird Count, which started in the year 1900. In this bird count, and in similar surveys run subsequently by other organisations (see "further reading" below for a few examples), people are asked to take note of wildlife in their area, and report which animals they encounter. This sort of work helps ecologists survey populations in large areas, and informs long-term conservation studies. Since data are collected by citizens, they are usually made available to the public after analysis. The Audubon counts, for example, are on their website.

In recent years, citizen science projects have moved to the web. Over a million people are registered to participate in Zooniverse's citizen science projects, which involve anything from hunting for planets to counting penguins. These web-based citizen science projects have a very low threshold for participation, and only require an internet connection and a few minutes of time during coffee breaks. In all cases, they are research projects where human eyes work better than computers, and where researchers need help from a large group of individuals to analyse a large dataset.

Another type of citizen science uses computer games to perform scientific calculations: Foldit is a game that lets players find the best protein-folding conformation, after which human-driven intuitive protein folding solutions are used to optimize computational protein folding calculations.

Such projects fall under open science because the researchers are allowing anyone to interact with their data, but they do restrict and control that analysis to their own platform. In most cases, they will make the resulting publications available to everyone (both Foldit and Zooniverse provide scientific publications via their sites), and in exceptional situations, citizens who participated in these projects may even be listed as co-authors.

Crowdfunding

Another feature that is sometimes included under "open science" is the crowdfunding of research projects. It should be noted, though, that crowdfunding does not require open science: researchers can crowdfund closed research projects as well.

However, like citizen science, involvement of a large group of people in the work (now at the funding level rather than at the data collection stage) often encourages researchers to make the output openly available so that donors can have access to the results of the work.

Usually, successfully funded projects have a clear goal, a small budget, and intriguing perks. There are exceptions: one of the most highly crowd-funded science projects is the ARKYD space telescope by Planetary Resources, which raised 1.5 million US dollars. Crowdfunding often only works for distinct projects and is difficult to scale to running a lab long-term. That being said, the California-based Perlstein Lab did start out as a crowd-funded project, but is currently run as a seed-funded startup company.

Open source software

Open science also overlaps with the open source software movement, which advocates the use and development of software that has its source code made available to others to re-use and build upon.

Open source software for science includes projects like the BioJS library of graphical components, which anyone can use or build on to represent and visualise biological information.

At the moment, a lot of scientific research and communication still relies on software that is not open source, and many open science proponents will, where an equivalent open source alternative is not available, use software that is not open, but will make sure the output of their own work (including software they produced) is openly available.

Open drug discovery

Open drug discovery combines open notebook science, citizen science and open source science to find new drugs. Different groups use slightly different approaches, but all are based on open science principles. The Open Source Drug Discovery platform, based in India, uses community participation for initial candidate target discovery, and works with researchers in academia and hospitals to then synthesize and test the targets. Other groups are focused more specifically on one disease, such as the Open Source Malaria project, which uses an open notebook approach to share all ongoing work. Various other collaborations are in place to find and test drug targets for a wide range of diseases.

Future of open science

In this guide we've looked at several aspects of open science: open access, open peer review, post-publication peer review, open data, and the topics listed above. That leaves a final thought: What is the future of open science?

Open access has been growing steadily over the past decade, and open peer review is becoming more popular. Many major funding organisations are asking for not just publications, but also the underlying data to be shared openly, and crowdfunded researchers are also often incentivised to share their work with their funders. Meanwhile, citizen science has been around for over a century, and is only growing with novel web applications that enable everyone to participate in scientific research and drug discovery. Online post-publication peer review is still quite new, and open notebook science has not spread very far yet, but both of these are steadily growing as well.

So despite researchers' fear of competition, and a reward system that still favours publication in exclusive journals (where openness is not a main concern), scientific research is gradually moving towards an open science system.

REFERENCES & RESOURCES

Open Notebook Science Network
http://onsnetwork.org/

Open WetWare:
http://openwetware.org

The IPython Notebook
http://ipython.org/notebook.html

Audubon Christmas Bird Count
http://birds.audubon.org/christmas-bird-count

Audubon Christmas Bird Count data
http://netapp.audubon.org/cbcobservation/

Zooniverse
https://www.zooniverse.org/

Foldit publications
http://fold.it/portal/info/about#folditpub

ARKYD: A space telescope for everyone. (Kickstarter funded 1 July 2013) https://www.kickstarter.com/projects/arkydforeveryone/arkyd-a-space-telescope-for-everyone-0

How a biotech startup juiced investors with yeast, flies and zebra fish (1 August 2014)
http://www.bizjournals.com/sanfrancisco/blog/biotech/2014/08/perlstein-lab-ethan-martin-shkreli-ngly1-biotech.html?page=all

BioJS http://biojs.net/

Corpas M, Jimenez R, Carbon SJ et al. BioJS: an open source standard for biological visualisation – its status in 2014 [v1; ref status: indexed, http://f1000r.es/2yy] *F1000Research* 2014, 3:55 (doi: 10.12688/*F1000Research*.3-55.v1)

Open Source Drug Discovery
http://www.osdd.net/

Open Source Malaria
http://opensourcemalaria.org/

FURTHER READING

Center for Open Science
http://centerforopenscience.org/

Open Science Federation
http://opensciencefederation.com/

Mozilla Science
http://mozillascience.org/

Open Science at the Open Knowledge Foundation http://science.okfn.org/

Open Science Framework
https://osf.io/4znzp/wiki/home

Science Commons principles for open science http://sciencecommons.org/resources/readingroom/principles-for-open-science/

Woelfle M, Olliaro P, Todd M (2011) Open science is a research accelerator *Nature Chemistry* 3, 745–748

Helen Shen: Interactive notebooks: Sharing the code (5 November 2014) http://www.nature.com/news/interactive-notebooks-sharing-the-code-1.16261

The Monarch Program – monitoring monarch butterflies
http://www.monarchprogram.org/

Scistarter – database of active citizen science projects http://scistarter.com/

SciFund Challenge - helps researchers communicate more clearly about their work, so that they can be more successful in crowdfunding.
http://scifundchallenge.org/

Experiment.com
https://experiment.com/

Petridish http://www.petridish.org/

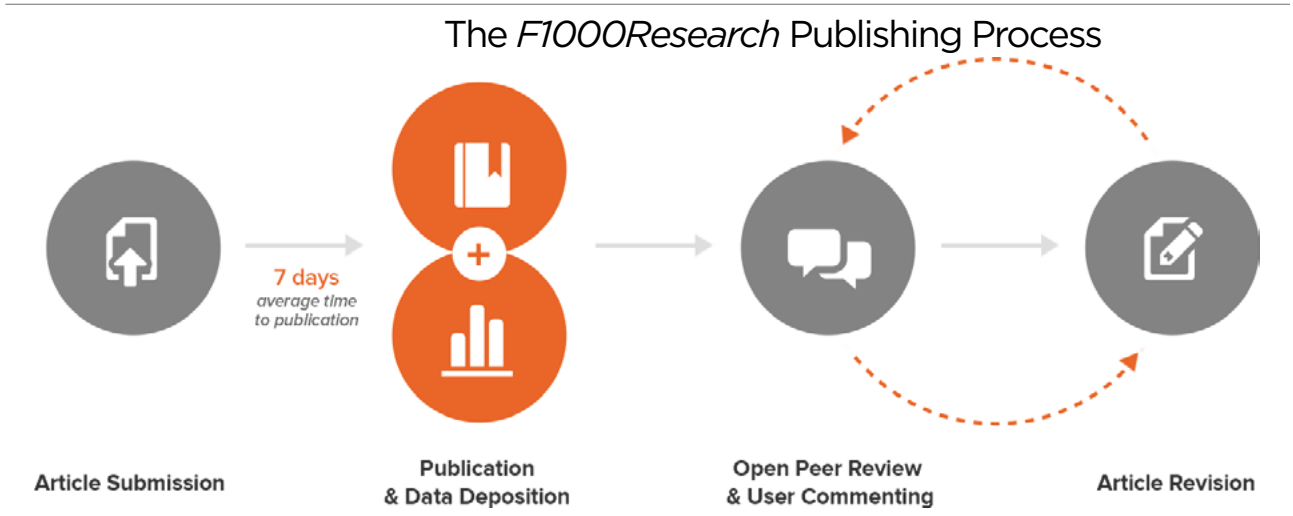
Kraker P., Leony D., Reinhardt W., & Beham G. (2011). The Case for an Open Science in Technology Enhanced Learning. *International Journal of Technology Enhanced Learning*, 6(3), 643–654.

Sabou M, Bontcheva K, Scharl A (2012) Crowdsourcing Research Opportunities: Lessons from Natural Language Processing. In: 12th International Conference on Knowledge Management and Knowledge Technologies, Special Track Research 2.0 (#STR20) , 5-7 September, Graz, Austria.

MacLean D (2013) Cutting edge: Changing the rules of the game *eLife* 2013;2:e01294

Li FW, Pryer KM (2014) Crowdfunding the Azolla fern genome project: a grassroots approach, *GigaScience* 3:16

Masum H, Rao A, Good BM, Todd MH, Edwards AM, et al. (2013) Ten Simple Rules for Cultivating Open Science and Collaborative R&D. *PLoS Comput Biol* 9(9): e1003244. doi:10.1371/journal.pcbi.1003244



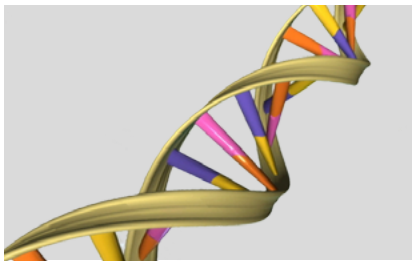
About *F1000Research*

F1000Research is an original open science publishing platform for life scientists that offers immediate open access publication, transparent post-publication peer review by invited referees, and full data deposition and sharing. All scientifically sound articles are accepted, including single findings, case reports, protocols, replications, null/negative results, as well as more traditional research articles and reviews.

F1000Research operates a fully transparent, author-driven publishing model: the authors are solely responsible for the content of their article. Invited peer review takes place openly after publication, and the authors play a crucial role in ensuring that the article is peer reviewed by independent experts in a timely manner.

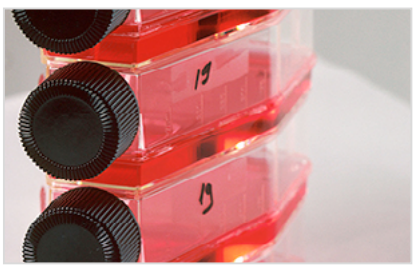
For more information, please visit f1000research.com or contact us at research@f1000.com

Some popular articles we have published



Thousands of exon skipping events differentiate among splicing patterns in sixteen human tissues [v2; ref status: indexed, http://f1000r.es/2dl]

Liliana Florea, Li Song, Steven L Salzberg

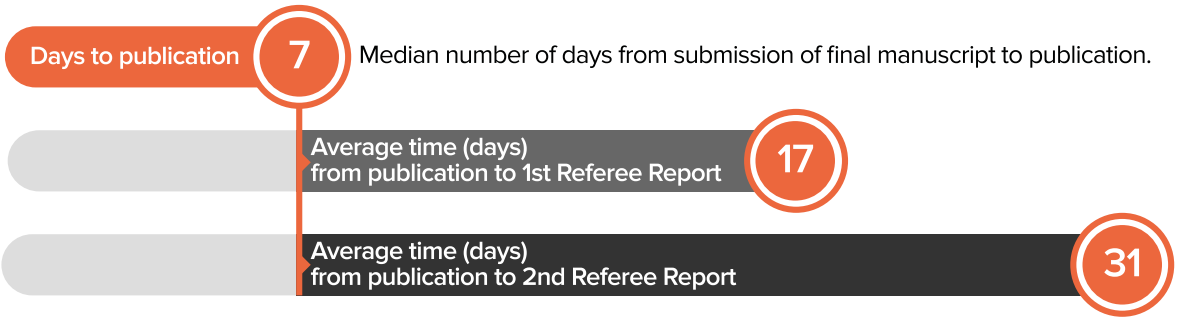


Transient acid treatment cannot induce neonatal somatic cells to become pluripotent stem cells [v1; ref status: indexed, http://f1000r.es/3dq]

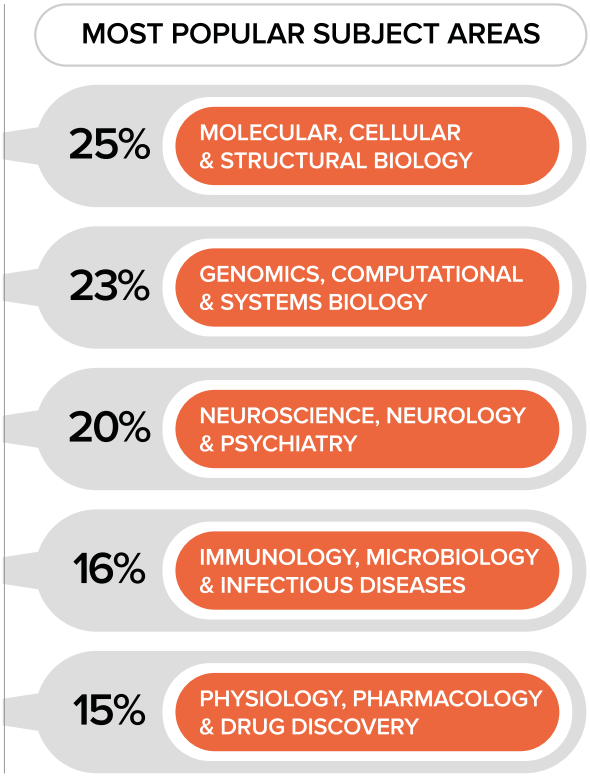
Mei Kuen Tang, Lok Man Lo, Wen Ting Shi, Yao Yao, Henry Siu Sum Lee, Kenneth Ka Ho Lee

Shaping the Future of Research: a perspective from junior scientists [v2; ref status: indexed, http://f1000r.es/4yc]

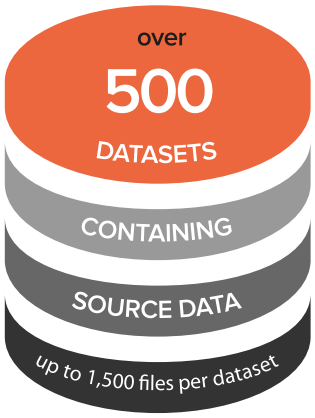
Gary S. McDowell, Kearney T. W. Gunsalus, Drew C. MacKellar, Sarah A. Mazzilli ... Ayesha Islam, Matthew D. Mattozzi, Kristin A. Krukenberg, Jessica K. Polka



Calculated over articles published in 2014



Research universities included the Russell Group in the UK, the U15 Group of Canadian Research Universities, and the Group of Eight coalition of Australian research universities.



60
SOFTWARE CODE ENTRIES STORED ON ZENODO

TOP FUNDERS

OF WORK PUBLISHED IN *F1000RESEARCH*

1. NATIONAL INSTITUTES OF HEALTH
2. NATIONAL SCIENCE FOUNDATION
3. NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL OF CANADA (NSERC)
4. EUROPEAN COMMISSION
5. WELLCOME TRUST

ADVISORY BOARD includes

3
Nobel Laureates

19
Fellows of the Royal Society

28
Members of the National Academy of Science

COUNTRIES OUR AUTHORS COME FROM

NORTH AMERICA
Aruba
Canada
Mexico
Panama
Puerto Rico
USA

EUROPE
Austria
Belgium
Bulgaria
Croatia
Cyprus
Czech Republic
Denmark
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Luxembourg
Netherlands
Norway
Poland
Portugal
Romania
Russian Federation
Serbia
Slovenia
Spain
Sweden
Switzerland
Turkey
UK
Ukraine

SOUTH AMERICA
Argentina
Brazil
Chile
Ecuador
Peru

AFRICA
Burkina Faso
Cameroon
Egypt
Ghana
Guinea-Bissau
Kenya
Malawi
Morocco
Nigeria
South Africa
Tunisia

ASIA
Armenia
Cambodia
China
Hong Kong
India
Indonesia
Iran
Iraq
Israel
Japan
Korea, South
Lebanon
Malaysia
Nepal
Pakistan
Philippines
Saudi Arabia
Singapore
Sri Lanka
Syria
Taiwan
Thailand
United Arab Emirates
Vietnam

OCEANIA
Australia
New Zealand