Preprinting and Publishing in the Life and Biomedical Sciences

II: Understanding and Engaging with Preprints

ASAPbio

Section 1

What are preprints? What is the impact of preprinting?

What is a preprint?
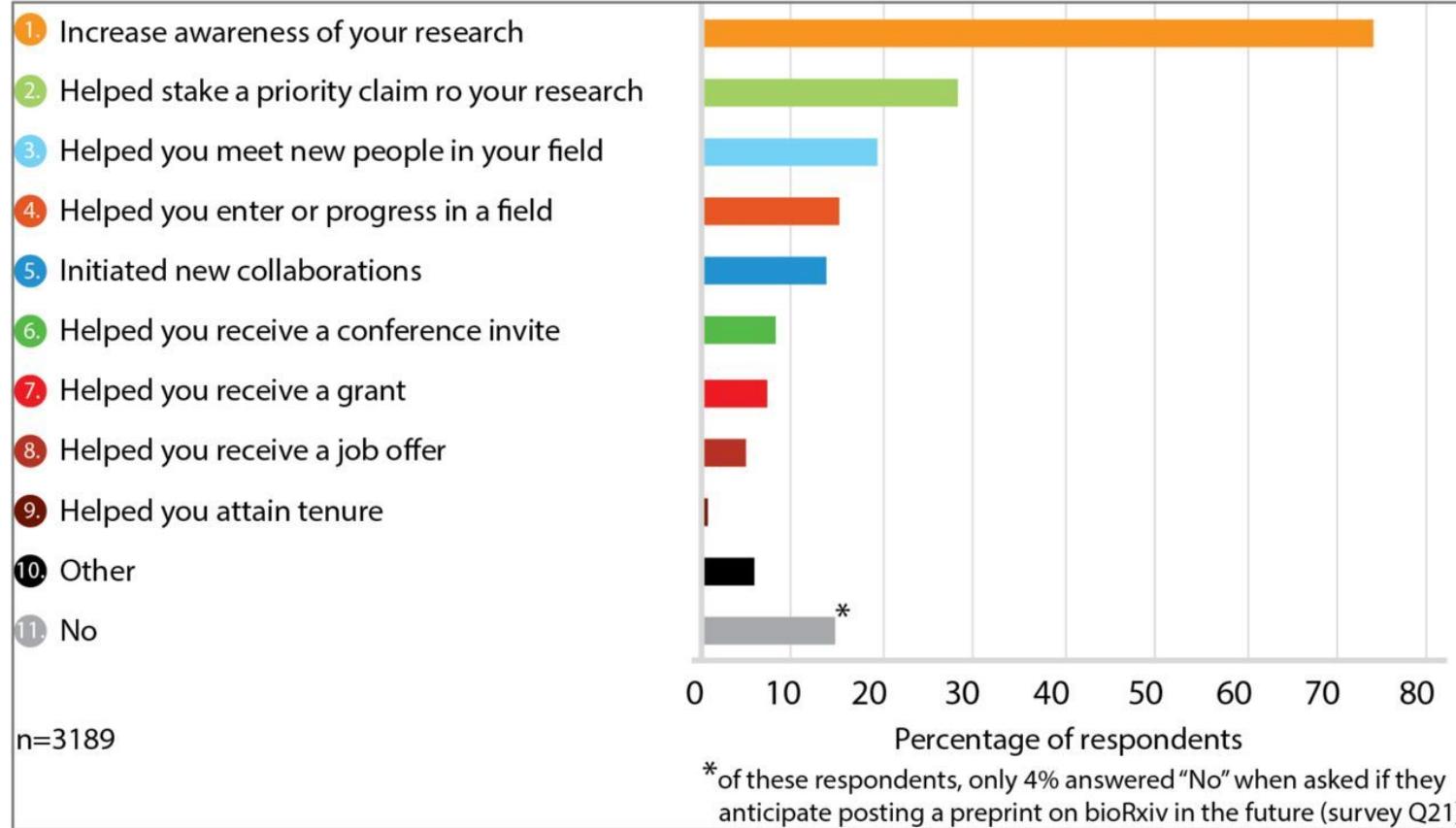
"A preprint is a full draft research paper that is shared publicly before it has been peer reviewed."

- Complete scientific manuscript posted to a preprint server, which is a publicly accessible platform to everyone around the world
- Once document is uploaded, quality inspection occurs
- Once accepted, preprints receive a DOI or persistent ID that can be cited
- Preprints can be updated at any time by the authors

Preprints make work available almost immediately

Preprint server

<48 hrs screening process

Public

Peer reviewed paper

Community feedback, ideas, discussion

Private

Manuscript

Submit

Journal 1

Peer Review

Journal 2

Revise

Journal 3

Revise

Months to years

ASAPbio

Benefits of preprints for scientists

1. Increase awareness of your research
2. Helped stake a priority claim ro your research
3. Helped you meet new people in your field
4. Helped you enter or progress in a field
5. Initiated new collaborations
6. Helped you receive a conference invite
7. Helped you receive a grant
8. Helped you receive a job offer
9. Helped you attain tenure
10. Other
11. No

n=3189

Percentage of respondents

*of these respondents, only 4% answered "No" when asked if they anticipate posting a preprint on bioRxiv in the future (survey Q21)

ASAPbio   @ASAPbio_ | #ASAPbio | @{yourhandle}

Preprints do not replace the peer review process, rather they can enhance it

- Disentangles scientific disclosure from peer-review validation
- Peer Review comments on preprint can strengthen the manuscript prior to submission to a peer-reviewed journal
- Community efforts, such as PREreview, provide platforms in which scientists can submit suggestions and comments on preprints
- Crowdsourcing peer review

Section 2

What are the top concerns about preprints?

Preprints and Quality Control

- Misinformation?
- Risk of public health or society?
- Deluge of Poor Papers?
- Peer Review is undermined?

'Sharing preprints can cause premature media coverage and subsequent misinformation'

- Misinformation concern is also shared with traditional peer-reviewed manuscripts
- Can be addressed by inclusion of a research summary dedicated to lay persons (non-scientists)
- If findings can directly be used in patient treatment or prevention, authors must make the study's limitations clear
- Preprints can be more easily retracted than "published" articles in journals

'Without peer review, there is a risk to public health'

Risk mitigation framework - medRxiv

Is it nonsense?

Is it non-science?

Is it a paper?

Is it research?

Is it plagiarized?

Is it a health threat?

Is there a benefit to sharing now vs. after peer review?

1	Author undertakings
2	Automated check
3	CSHL Check
4	medRxiv Affiliate check
5	Escalation 1- experienced clinician-editor(s)
6	Escalation 2 - medRxiv leadership
7	Posting and public discussion

Theo Bloom, presentation at FORCE2019
http://bit.ly/preprints-FORCE2019

'Without peer review, there is a risk to public health'

medRxiv requires declarations in line with those required for reporting of clinical work in peer-reviewed literature:

- Competing interests
- Funding statement
- Ethical approval/consent
- Clinical trial registration

As well as data statements (beyond what some journals operate)

Increased plasma heparanase activity in COVID-19 patients

Baranca Buijsers, Cansu Yanginlar, Inge Grondman, Aline de Nooijer, Marissa L Maciej-Hulme, Inge Jonkman, Nico Janssen, Nils Rother, Mark de Graaf, Peter Pickkers, Matthijs Kox, Leo Joosten, Tom Nijenhuis, Mihai G Netea, Luuk Hillbrands, Frank van de Veerdonk, Raphael Duivenvoorden, Quirijn de Mast, ⓘⅅ Johan van der Vlag

doi: https://doi.org/10.

This article is a pr [what does this mean?]. It reports l and so should *not* be used

Competing Interest Statement

The authors have declared no competing interest.

Clinical Trial

This study was performed according to the latest version of the declaration of Helsinki and guidelines for good clinical practice. The local independent ethical committee approved the study protocol (CMO 2020-6344, CMO 2020-6359, CMO 2016-2923).

Funding Statement

This study was financially supported by the Radboud university medical center PhD fellow program and consortium grant LSHM16058-SGF (GLYCOTREAT; a collaborative project financed by the PPP allowance made available by Top Sector Life Sciences & Health to the Dutch Kidney Foundation to stimulate public-private partnerships) coordinated by JvdV. MGN was supported by an ERC Advanced grant (#833247) and a Spinoza Grant of the Netherlands Organization for Scientific Research.

Author Declarations

I confirm all relevant ethical guidelines have been followed, and any necessary IRB and/or ethics committee approvals have been obtained.

Yes
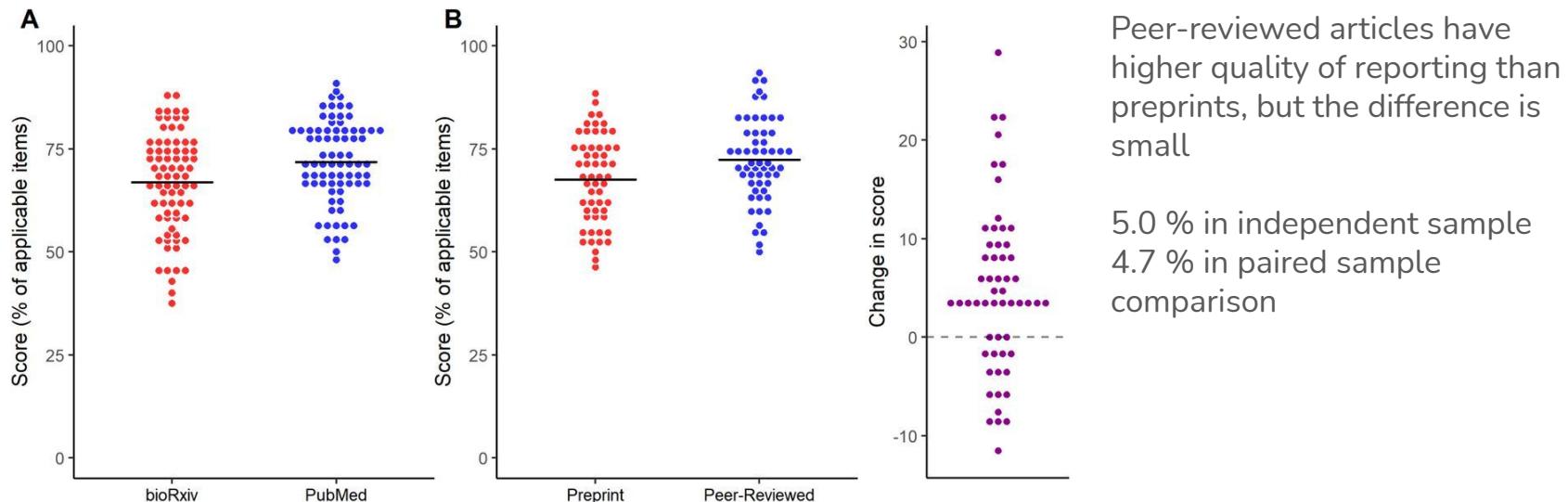
The details of the IRB/oversight body that provided approval or exemption for the research described are given below:

This study was performed according to the latest version of the declaration of Helsinki and guidelines for good clinical practice. The local independent ethical committee approved the study protocol (CMO 2020-6344, CMO 2020-6359, CMO 2016-2923).

All necessary patient/participant consent has been obtained and the appropriate institutional forms have been archived.

'Preprints will lead to a deluge of poor papers'

Peer-reviewed articles have higher quality of reporting than preprints, but the difference is small

5.0 % in independent sample
4.7 % in paired sample comparison

Comparison of random sample (76) of bioRxiv preprints to peer-reviewed articles from PubMed, and a paired comparison of a sample (43) of bioRxiv preprints to their own peer-reviewed article versions

Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. Carneiro *et al.* bioRxiv 581892; doi: https://doi.org/10.1101/581892

▶ASAPbio   @ASAPbio_ | #ASAPbio | @{yourhandle}

'Preprints can undermine the value of peer review'

- That is certainly not the goal of preprints; in fact, preprints are meant to encourage more peer review!
- PREreview allows scientists to submit review of preprints
- These reviews can be potentially integrated in the publishing workflow

Preprints and Scientific Careers

- Scooping?
- Journal won't publish my work?
- What's in it for me?

'My work will be scooped'

Has posting a preprint negatively affected you in any of the following ways	% of respondents
No	89.6
Limited your choice of journal for publication	6.43
Prevented you from publishing in your journal of choice because another lab published before you	0.70
Affected your priority claim to the research	1.25
Other	4.41
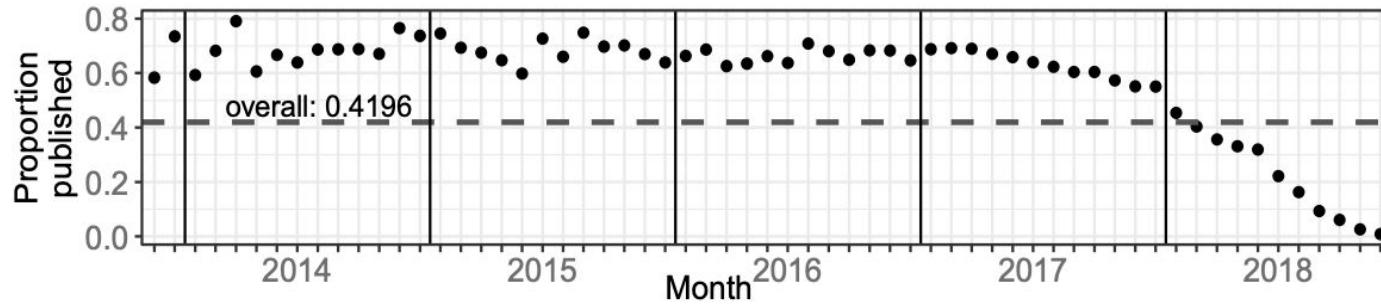
bioRxiv survey N=3127 'bioRxiv: the preprint server for biology' https://doi.org/10.1101/833400

There is no evidence that preprints increase risk for scooping

Paul Ginsparg, founder of arXiv on scooping: *"It can't happen, since arXiv postings are accepted as date-stamped priority claims."*

Resources on scooping available on ASAPbio website: https://asapbio.org/preprint-info/preprint-faq A number of journals operate scooping protection policies: EMBO, *eLife*, PLOS journals

'The journal will not publish my work'

⅔ of preprints are published within two years

The study by Addill & Blekhman focused on preprints in bioRxiv, the same statistic has been reported for preprints in arXiv (Larivière *et al.* Journal of the American Society for Information Science and Technology, 65(6): 1157−1169)

'The journal will not publish my work'

- SHERPA/RoMEO lists over 1,200 publishers with policies that accept preprints
- **TRANSPOSE database** (https://transpose-publishing.github.io/#/) provides information on preprint policies at journals
- Some journals give the authors the option to post the paper at a **partner preprint platform in parallel** to consideration at the journal
- Some journals have dedicated editors who check preprints to invite submission to the journal (see https://asapbio.org/journal-policies for more info on innovative journal practices)

'What's in it for me?'
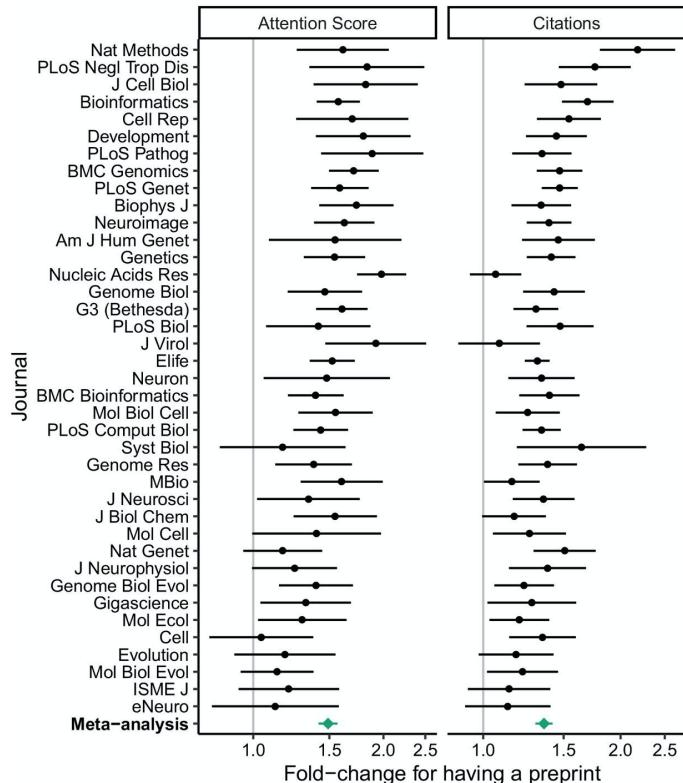
A number of funders encourage preprints as evidence of productivity in grant applications & reports
List and links to policies at asapbio.org/funder-policies

'What's in it for me?'
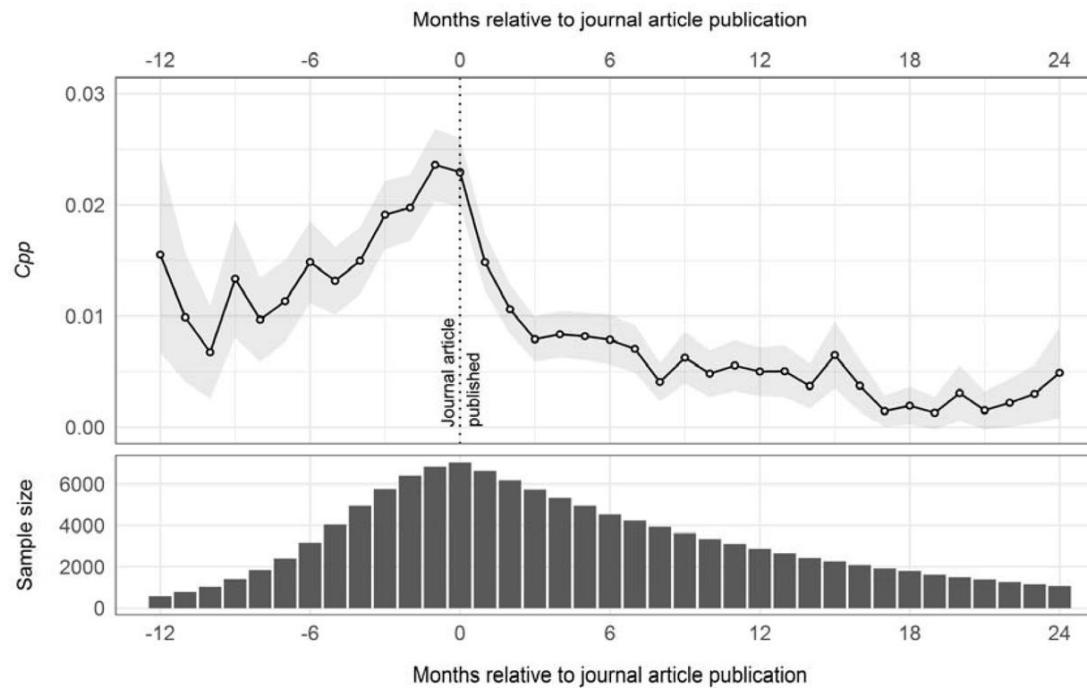
Having a preprint on bioRxiv is associated with a higher Altmetric Attention Score and more citations of the peer-reviewed article

'What's in it for me?'
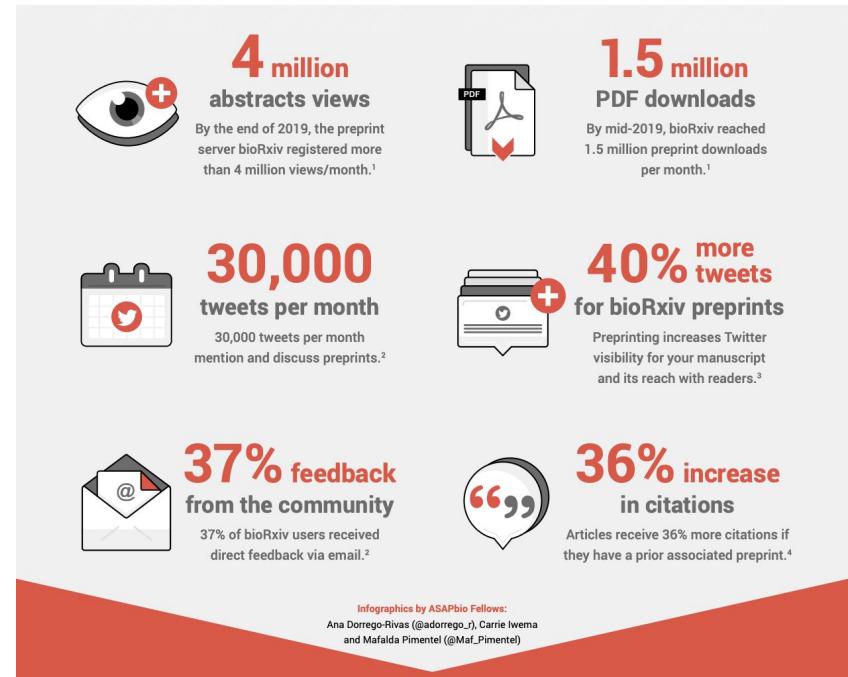
Papers posted to bioRxiv receive citations prior to journal publication

Preprints can **extend the reach** of the work

The effect of bioRxiv preprints on citations and altmetrics. Nicholas Fraser, Fakhri Momeni, Philipp Mayr, Isabella Peters. bioRxiv 673665; doi: https://doi.org/10.1101/673665

What's in it for me?

Preprints increase the visibility of your work!

4 million abstracts views
By the end of 2019, the preprint server bioRxiv registered more than 4 million views/month.[1]

1.5 million PDF downloads
By mid-2019, bioRxiv reached 1.5 million preprint downloads per month.[1]

30,000 tweets per month
30,000 tweets per month mention and discuss preprints.[2]

40% more tweets for bioRxiv preprints
Preprinting increases Twitter visibility for your manuscript and its reach with readers.[3]

37% feedback from the community
37% of bioRxiv users received direct feedback via email.[2]

36% increase in citations
Articles receive 36% more citations if they have a prior associated preprint.[4]

Infographics by ASAPbio Fellows:
Ana Dorrego-Rivas (@adorrego_r), Carrie Iwema and Mafalda Pimentel (@Maf_Pimentel)

Section 3

What are the components of a preprint?

A. Manuscript:

Complete scientific work

Structure and content:
Should contain all the sections relevant to a scientific article
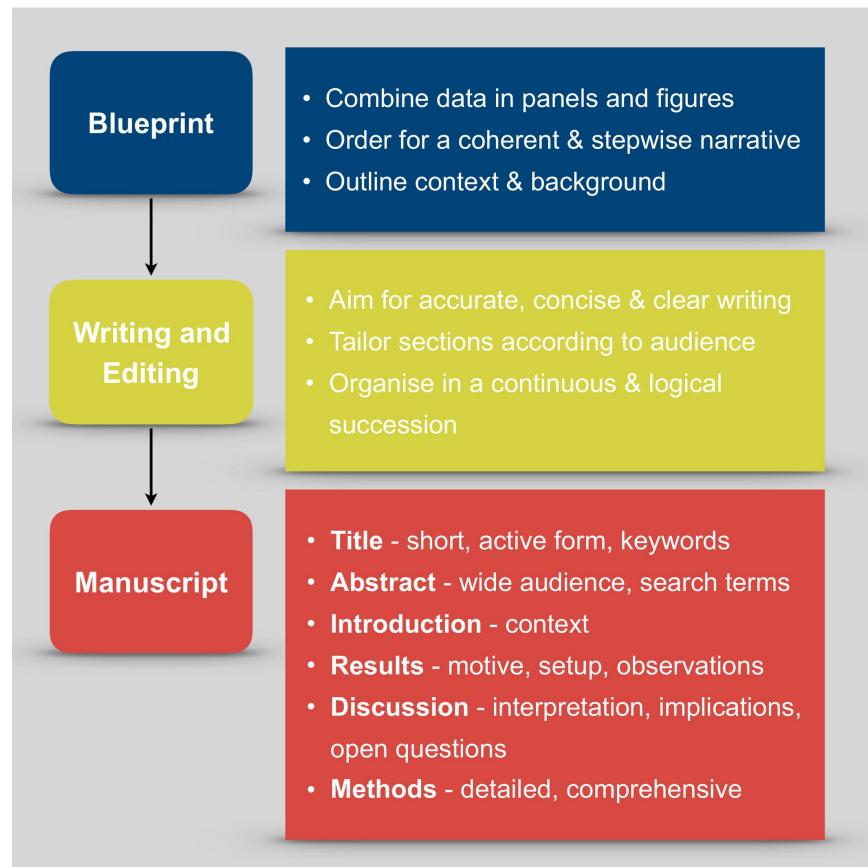
Manuscript length and format:
A preprint has more freedom compared to a journal submission

On quality:
Review the manuscript to avoid scientific and grammatical errors
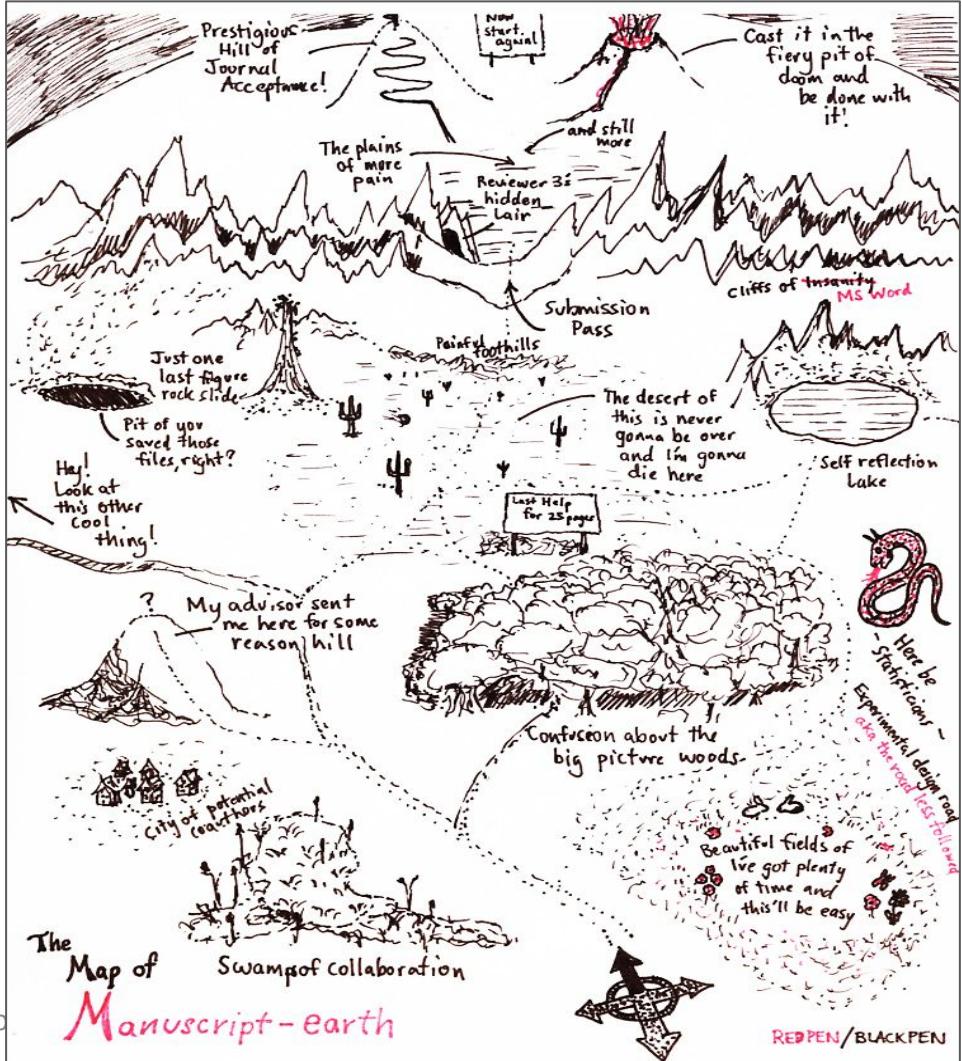
Blueprint	• Combine data in panels and figures • Order for a coherent & stepwise narrative • Outline context & background
Writing and Editing	• Aim for accurate, concise & clear writing • Tailor sections according to audience • Organise in a continuous & logical succession
Manuscript	• **Title** - short, active form, keywords • **Abstract** - wide audience, search terms • **Introduction** - context • **Results** - motive, setup, observations • **Discussion** - interpretation, implications, open questions • **Methods** - detailed, comprehensive

Manuscript:

The process of research and writing a manuscript is lengthy and with a lot of hurdles to overcome.

ASAPbio   @ASAPbio

Manuscript: when to post

1. Before Journal submission
2. At journal submission AND/OR before peer-review
3. After peer-review but before acceptance

Subject to journal policies which vary.
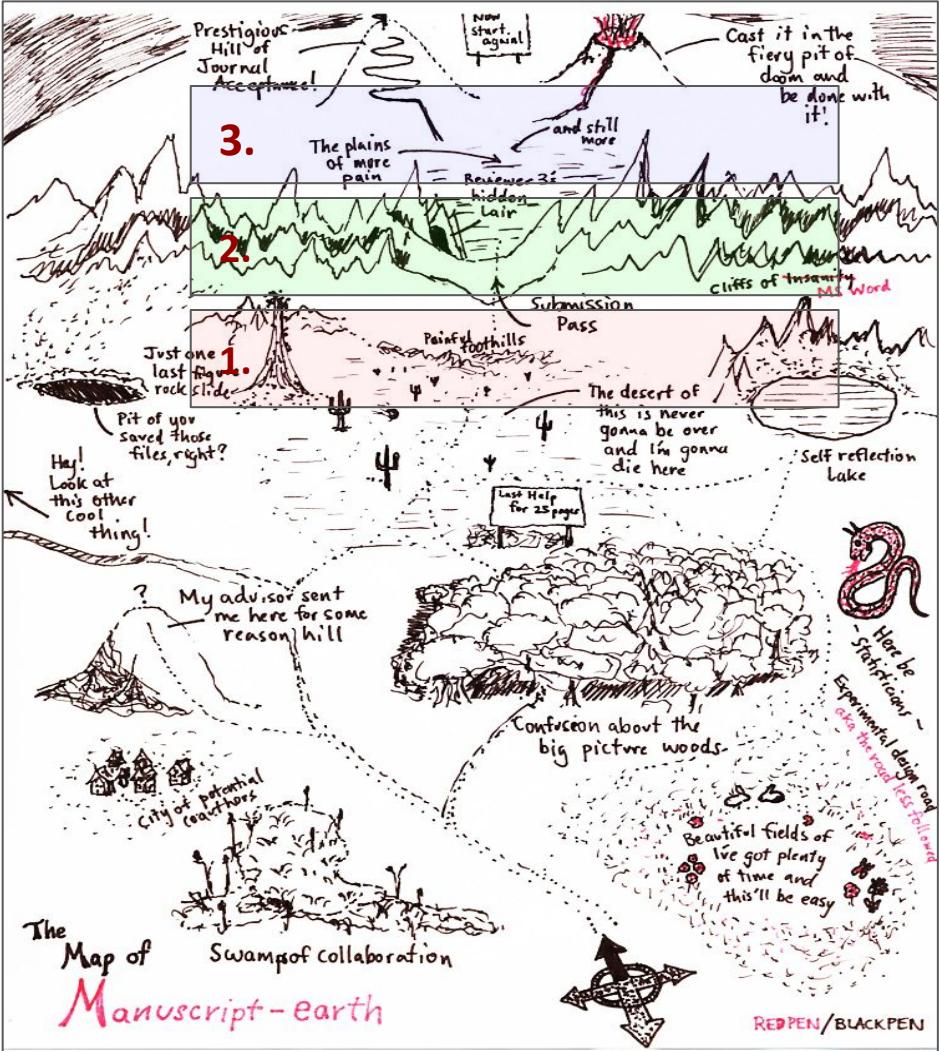
https://www.redbubble.com/people/redpenblackpen/shop?ref=artist_title_name

When can preprints be posted?

"Nature Portfolio journals encourage posting of preprints of primary research manuscripts on preprint servers… …preprints may be posted at any time during the peer review process"

\- Nature Research

"… we do not support posting to preprint servers revisions that respond to editorial input and peer review or final accepted manuscripts. Once your paper is published, we encourage you to update the preprint record with a link to the final published article. Please see our prepublication publicity policies …"

\- Cell Press

Posting on Multiple servers

"We recommend that a preprint is posted on only one server. bioRxiv provides metrics for article views, PDF downloads, and attention scores." bioRxiv

Multiple versions/revisions:

New pre-print versions may be displayed preferentially

Versions are considered permanent citable scientific communications - can only be withdrawn due to "significant ethical or legal concerns"

"An article posted on bioRxiv can be revised at any time, until it is accepted for publication....To correct errors in your article or Supplementary file, please submit a revised version of your article."

-bioRxiv

Multiple versions/revisions:

Exploring correlations in cultural and genetic variation across language families in Northeast Asia

Hiromi Matsumae, Peter Ranacher, Patrick E Savage, Damian E Blasi, Thomas E Currie, Kae Koganebuchi, Hideyuki Tanabe, Takehiro Sato, Nao Nishida, Atsushi Tajima, Steven Brown, Mark Stoneking, Kentaro K K Shimizu, Hiroki Oota, Balthasar Bickel

bioRxiv 513929; doi: https://doi.org/10.1101/513929   Revision   New Results
+ Add to Selected Citations

Beating your neighbor to the berry patch
Alan R. Rogers

bioRxiv 2020.11.12.380311; doi: https://doi.org/10.1101/2020.11.12.380311   Revision New Results
+ Add to Selected Citations

Large-scale study validates that regional fungicide applications are major determinants of resistance evolution in the wheat pathogen Zymoseptoria tritici in France.

Maxime Garnault, Clementine Duplaix, Pierre Leroux, Gilles Couleaud, Olivier David, Anne-Sophie Walker, Florence Carpentier

bioRxiv 2020.07.17.208728; doi: https://doi.org/10.1101/2020.07.17.208728   Revision New Results
+ Add to Selected Citations

A demogenetic agent based model for the evolution of traits and genome architecture under sexual selection

Louise Chevalier, François de Coligny, Jacques Labonne

bioRxiv 2020.04.01.014514; doi: https://doi.org/10.1101/2020.04.01.014514   Revision New Results
+ Add to Selected Citations

Testing methods of linguistic homeland detection using synthetic data
Soeren Wichmann, Taraka Rama

bioRxiv 2020.09.03.280826; doi: https://doi.org/10.1101/2020.09.03.280826   Revision New Results
+ Add to Selected Citations

Wolbachia and host intrinsic reproductive barriers contribute additively to post-mating isolation in spider mites

Miguel Alfredo Cruz, Sara Magalhaes, Elio Sucena, Flore Zele

bioRxiv 2020.06.29.178699; doi: https://doi.org/10.1101/2020.06.29.178699   Revision New Results
+ Add to Selected Citations

178699v4

Wolbachia and host intrinsic reproductive barriers contribute additively to post-mating isolation in spider mites

Miguel Alfredo Cruz, Sara Magalhaes, Elio Sucena, Flore Zele

doi: https://doi.org/10.1101/2020.06.29.178699

This article is a preprint and has not been certified by peer review [what does this mean?].

Abstract | Info/History | Metrics | Preview PDF

ARTICLE INFORMATION

doi   https://doi.org/10.1101/2020.06.29.178699

History   November 14, 2020.

ARTICLE VERSIONS

Version 1 (July 5, 2020 - 21:57).
Version 2 (July 9, 2020 - 20:45).
Version 3 (October 30, 2020 - 01:56).

You are viewing Version 4, the most recent version of this article.

Copyright   The copyright holder for this preprint is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

REVISION SUMMARY

- Version 4 of this preprint has been peer-reviewed and recommended by Peer Community In Evolutionary Biology (https://doi.org/10.24072/pci.evolbiol.100116).

Open Science & Preprints
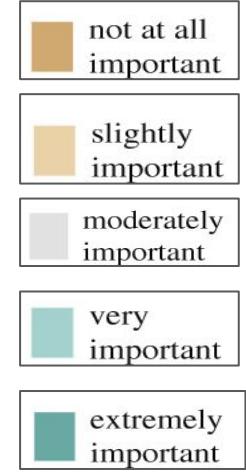
Credibility of preprints Vs Shared Information

Credibility of preprints: an interdisciplinary survey of researchers, Volume: 7, Issue: 10, DOI: (10.1098/rsos.201520)

Submit preprints and view/download

- Spearman correlations for the view/download and submit correlations

- Pearson correlation for favourability

Credibility of preprints: an interdisciplinary survey of researchers, Volume: 7, Issue: 10, DOI: (10.1098/rsos.201520)

	view/download preprints	submit preprints	favor use
author's previous work	−0.10	−0.06	−0.11
author's institution	−0.10	−0.08	−0.10
professional identity links	−0.07	−0.05	−0.02
COI disclosures	−0.06	−0.11	0.01
author's level of open scholarship	−0.06	−0.07	0.04
funders of research	−0.10	−0.10	−0.00
preprint submitted to a journal	−0.20	−0.22	−0.26
usage metrics	−0.02	0.02	0.07
citations of preprints	0.01	0.01	0.10
anonymous comments	−0.03	−0.03	0.06
identified comments	0.03	−0.02	0.12
simplified endorsements	−0.05	−0.02	0.04
link to study data	0.13	0.03	0.15
link to study analysis scripts	0.17	0.05	0.17
link to materials	0.11	0.01	0.13
link to pre-reg	0.06	−0.03	0.11
info about indep groups accessing linked info	0.11	0.04	0.18
info about indep group reproductions	0.08	−0.02	0.10
info about indep robustness checks	0.04	−0.02	0.08

Section 4

What are the steps in submitting a preprint?

So you decided to preprint - now what?

1. Prepare your preprint.
2. Get all co-authors on board with preprinting. Refer to the resources in the Preprint Info Center (including these FAQ).
3. Double check journal policies on when and where preprints may be posted.
4. Choose a preprint server. Consider visibility, funder recommendations, and features like preservation and indexing, which are cataloged in the Preprint Server Directory.
5. Choose a license.
6. Upload any code/data/reagents you want to share to appropriate repositories.
7. Post the preprint!
8. Invite feedback via social media or email

https://asapbio.org/preprint-info/preprint-faq#qaef-4524
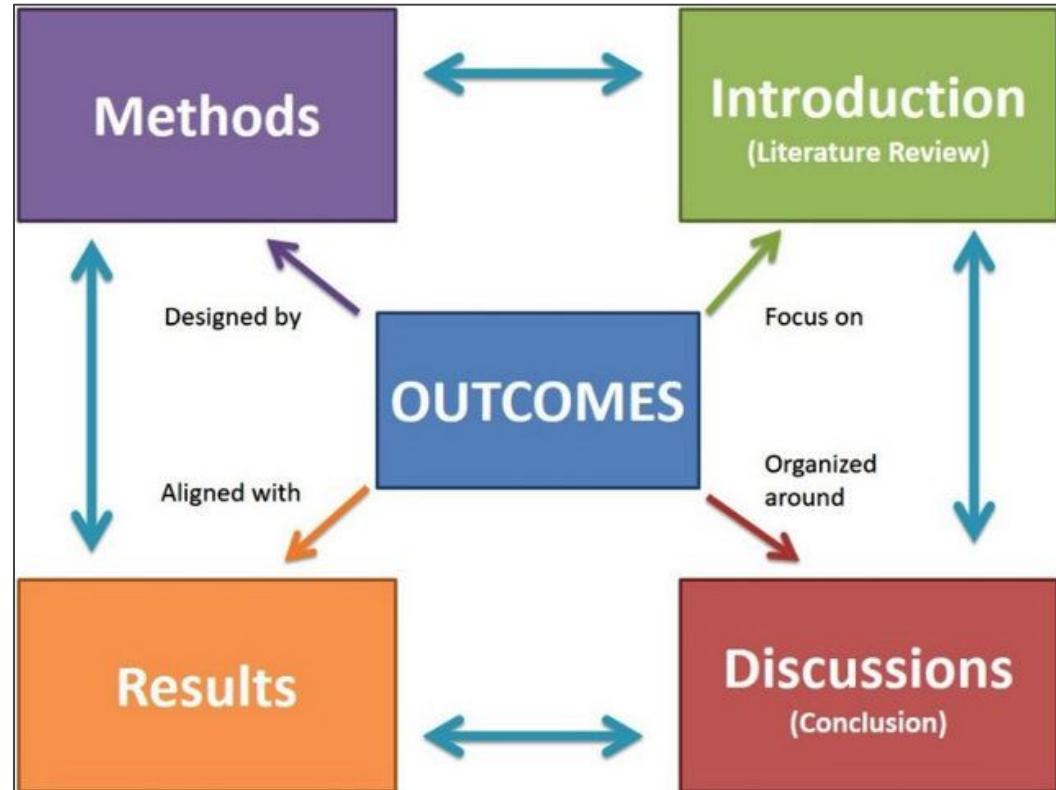
1 Prepare your Preprint

Prepare you preprint

- Complete the Scientific work
- Prepare a complete manuscript
- Review the manuscript to avoid scientific and grammatical errors

Yusoff MSB. ABC of manuscript writing. Education in Medicine Journal. 2018;10(2):61–67. https://doi.org/10.21315/eimj2018.10.2.8

2

Seek Approval from Coauthors

- Find preprints by your colleagues in the field or at your institution
- Discuss preprints in other settings: in a journal club, etc

Resources for preprints

asapbio.org/preprint-info

Learn more about preprints

Visit the preprint FAQ to learn more about submitting preprints, what they mean for scooping, and preprints in general.

We also have resources on choosing a license for your preprint.

Browse our collection of further readings about preprints.

Keep up to date

Learn about the use of preprints in COVID-19 research.

View statistics on the growth of preprints over time.

See a listing of preprint server products and services.

Policies about preprints

See journal practices and policies.

See university policies about preprints.

See funder policies recognizing and encouraging preprints and other interim research products.

Take action

What happens when you preprint? Hear first-hand stories from biologists.

Ready to spread the word about preprints? There are many ways to help awareness grow.

Take action

3 Double check journal policies

SHERPA/RoMEO is a fantastic start

ASAPbi

...but journal policies are much more nuanced

eg, what types of servers are allowed?

- The Royal Society of Chemistry journals allow deposition with "==non-commercial== repositories" such as ArXiv and ChemRxiv - http://www.rsc.org/journals-books-databases/journal-authors-reviewers/processes-

- Development "supports authors who wish to post primary research manuscripts on ==community preprint servers== such as bioRxiv." - http://dev.biologists.org/content/journal-policies#preprint

- Biophysical Journal "will consider for publication manuscripts that have been posted informally on a private website or on ==arXiv or bioRxiv==, but will not consider manuscripts that have been posted on other preprint servers or "virtual journal" websites." - https://www.cell.com/pb-assets/journals/society/biophysj/PDFs/author-guidelines.pdf?code=cell-site

4 Choose your server

Preprints for all disciplines, languages, & communities

Landscape of platforms

- Access to money, staff, time, publishing know-how
- Philosophy on amount of gatekeeping versus speed & transparency
- Motivations: from publisher-driven preprints to publishing-disruptive preprints

Multi-disciplinary platforms owned by or affiliated with for-profit publishers

Elsevier

SSRN & First Look platforms:
Cell Press Sneak Peek, NeuroImage: Clinical *First Look*, Preprints with The Lancet, Surgery Open Science *First Look*

Authorea

Wiley/Atypon

Research Square

Springer Nature: BMC & Nature Research

F1000 Research

Taylor & Francis

preprints.org

MDPI

JMIR Preprints

(PeerJ Preprints)

MitoFit Preprint Archives

Therapoid

Subject-specific platforms run by for-profit (med-tech & other) companies

Run by individual(s)

ViXra   *Some OSF communities*

MarXiv

for-profit

Academic groups, societies or funding organisations

non-profit

OSF communities:
AfricArXiv
AgriXiv
Arabixiv
EcoEvoRxiv
FocUS Archive
Frenxiv
INA-Rxiv
MetaArXiv
MindRxiv
NutriXiv
PaleoRxiv
PsyArxiv
SocArXiv
SportRxiv
Thesis Commons

bioRxiv & medRxiv

ESSOAr

OSF Preprints   ChemRxiv

arXiv

ChinaXiv

SciELO Preprints

Open Research platforms: AAS, AMRC, Gates, HRB, MNI, Wellcome

multi-disciplinary ⟷ *sub-disciplinary*

Slide by Naomi Penfold - https://zenodo.org/record/3626770#.XraqE2hKh3h

›ASAPbio @ASAPbio_ | #ASAPbio | @{yourhandle}

Preprint servers differ in screening, withdrawal, commenting policies

5 Choose your license

The majority of "free" literature is not open

How open is your preprint?

The license you choose has a big impact on how your work will be shared & reused.

The Creative Commons (CC) licenses described here break down the barriers to sharing by communicating rights and permissions up front with everyone.

⧉ NCBI Resources ⌄ How To ⌄

PMC
US National Library of Medicine
National Institutes of Health

PMC ▼

About PMC ▼   For Publishers ▼   Rel

Open Access Subset

2,080,010

out of

5,384,701

* any CC license as of 2019-02-14

http://asapbio.org/licensing-faq

CC0 waiver

CC0 places work in the public domain, waiving all copyright and related rights.

Allows anyone to repost or reuse your preprint in any medium for any purpose, even without attributing it to you.

Often used for works created by U.S. government employees, as these are already in the public domain in the U.S.

Ideal for datasets.[3]

CC BY

Attribution (BY)

Allows anyone to repost or modify your preprint in any medium for any purpose, but requires that users provide attribution to you and include a link back to the original whenever the material is used and shared.

Encouraged by NIH.[1]

Fits the original definition of open access.[2]

-NC,-ND,-SA

These terms can be added to the CC BY license to produce 5 other licenses
https://creativecommons.org/licenses/

Noncommercial (NC)

Prohibits commercial use of the material.

If you select it, you *don't* grant permission to:

○ Republish a figure in a paywalled journal
○ Use the preprint to advertise products
○ Reprint the work in a textbook sold commercially

No derivatives (ND)

Prohibits the sharing of adaptations of the material.

If you select it, you *don't* grant permission to:

○ Translate the preprint to another language
○ Create a copy of the preprint with extensive annotations
○ Adapt a diagram or drawing for use in another paper

ShareAlike (SA)

Requires adaptations of the material to be released under the same license.

For example, a figure that is modified from your preprint would have to *also* be published under a CC BY-SA license. (However, a book containing that modified figure could have its own, more restrictive license).

This license is used by Wikipedia and Wikimedia Commons.

No license

All rights reserved

If you do not select a license, you do not give default permission to reuse the work (beyond what is required to post to the preprint server).

As a result, you *don't* grant permission to:

○ Repost your paper, unchanged, on a class website

Using a figure in academic talks or text & data mining may also be prohibited in countries without a fair use or equivalent doctrine. Note that some servers (bioRxiv, arXiv, etc) allow TDM for all manuscripts.

Remember…

- All CC licenses require reusers to indicate if changes have been made, which alerts others that the work as modified is not the same as the original.
- As long as you retain the copyright in your work, you can always grant additional permissions on an individual basis. This includes giving permission for someone to reproduce or modify your work, commercialize your work, or transferring copyright to a journal or signing a license to publish agreement.
- Licenses are permanent, but don't stop authors from releasing other versions under other licenses.
- Professional norms for citation and plagiarism apply regardless of how content is licensed, and even for works dedicated to the public domain under CC0. Often, those norms are more restrictive than the attribution requirements of CC licenses.
- Fair use and other limitations and exceptions apply regardless of which license is selected.

References
1. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html
2. http://www.soros.org/openaccess/boai-10-recommendations
3. https://creativecommons.org/share-your-work/public-domain/cc0/

http://asapbio.org/licensing-faq

ASAPbio   @ASAPbio_ | #ASAPbio | @{yourhandle}

5

Upload code, data, or deposit reagents

Preprints servers offering clear links to data/code

OSF

bioRxiv

○ Previous

Posted October 15, 2020.

📄 **Download PDF**   ✉ Email
📄 Supplementary Material   ↪ Share
📄 Data/Code   🌐 Citation Tools

🐦 Tweet   👍 Like 0

Public Data: ~~(required)~~
⦿ Available   ○ No   ○ Not applicable

Links to data: (required)
https://osf.io/43z28/   ➖
https://osf.io/43z28/   ➖
                                                   ➕

Data refers to raw and/or processed information (quantitative or qualitative) used for the analyses, case studies, and/or descriptive interpretation in the preprint. Public data could include data posted to open-access repositories, public archival library collection, or government archive. For data that is available under limited circumstances (e.g., after signing a data sharing agreement), choose the 'No' option and use the comment box to explain how others could access the data.

A Catalogue of Databases for data: FAIRsharing.org

Recommendations for preprint data/materials statements:
Beck, Jeffrey, et al. "Building Trust in Preprints: Recommendations for Servers and Other Stakeholders." OSF Preprints, 21 July 2020. Web. https://osf.io/8dn4w/ | #ASAPbio | @{yourhandle}

6 Post it!
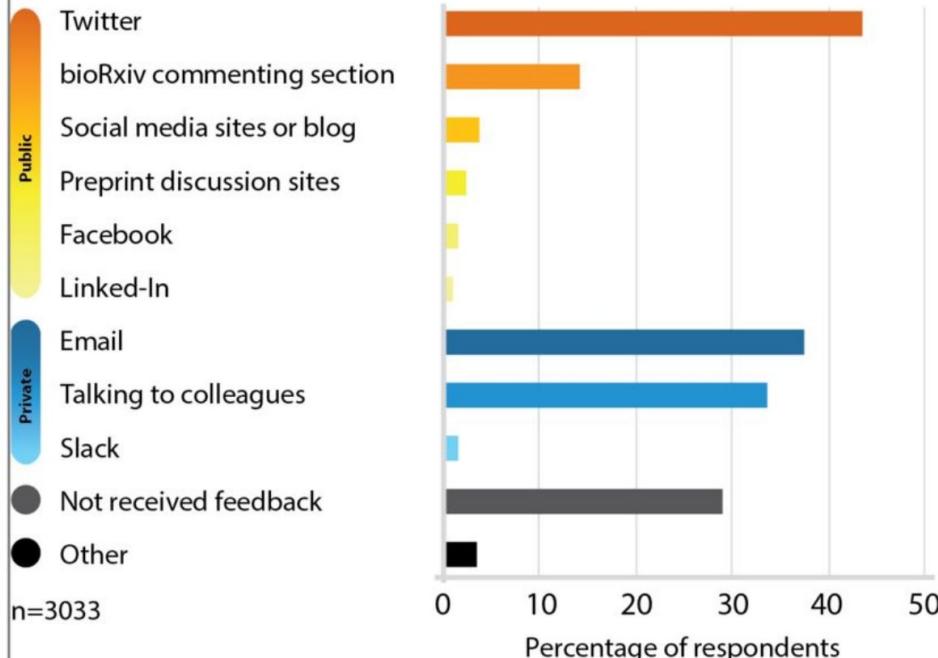
7 Request feedback

Feedback is not just in the comment section

In a survey of bioRxiv users, scientists were asked the mechanisms by which they have received feedback on papers posted on bioRxiv.

Feedback can be found all around the web

Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP

Daniel S. Quintana[1*] and Donald R. Williams[2]

https://osf.io/sgpe9/

Psychological Methods Discussion Group

Public Group

Discussion

| Daniel's post

Members

Daniel Quintana shared a link.
April 10

Just posted a preprint on Bayesian alternatives for common null-hypothesis significance tests that may be of interest to the group. Our goal was to put together a non-technical walkthrough using JASP for those unfamiliar with Bayesian alternatives. Would appreciate any feedback

osf.io
OSF.IO

https://www.facebook.com/groups/853552931365745/permalink/1349684805085886/

Daniel Quintana shared a link.
April 10

Just posted a preprint on Bayesian alternatives for common null-hypothesis significance tests that may be of interest to the group. Our goal was to put together a non-technical walkthrough using JASP for those unfamiliar with Bayesian alternatives. Would appreciate any feedback

Like   Share
13

Uli Schimmack I thought this would be a tutorial about picking alternative hypothesis to carry out a Bayesian statistical analysis because this is an important additional and new step that researchers are not familiar with. Unlike NHST where you only need to specify H0, default effect size = 0, Bayesian hypothesis testing requires also to specify H1 because BF provide information that the relative support for H0 and H1 given the data.

Alah, this is just another "tutorial" with all the wrong claims about p-values, a focus on hypothesis testing, when we really want to know how effective drugs are (effect sizes) and a total neglect of Bayesian and frequentist ways to assess the probability that a drug is not effective. Daniel Lakens

http://daniellakens.blogspot.ca/.../tost-equivalence...

Excuse me, if this is a bit harsh, but we have been discussing these issues for over a year now and I think it is fair to request a balanced and informative review of options to draw inferences from data.

Stop bashing p-values and provide some guidelines for researchers how they can pick a sensible alternative and how they BF have to be interpreted in the light of prior odds of H0 and H1.

TOST equivalence testing R package (TOSTER) and spreadsheet
I'm happy to announce my first R package (TOSTER)
DANIELLAKENS.BLOGSPOT.COM

Like 1 · April 10 at 6:42pm

Daniel Quintana Thanks for the feedback, glad to hear this while it's still a preprint. We actually cited Daniel Lakens' excellent TOSTER paper but I guess we can make this clearer.
Like 1 · April 10 at 11:46pm

Daniel Lakens Hi Daniel Quintana, I read the first few pages, and I have good news and bad news. The good news is the reviewers are all Bayesians, it will be accepted. The bad news is there are quite important misunderstandings of p-values and Bayes factors in the paper.

The hypothesis you describe in the intro (is the null true, or is there an effect larger than 0) can only be tested with p-values. It is underspecified for Bayesian stats. In Bayes, the alternative is 'is there a true effect between x and y with the distribution like z'. So the intro is an argument against Bayes factors. They don't allow you to test the hypothesis you seem interested in.

Then I stopped reading where you said Bayes factors could quantify the size of an effect. It is not true. You need to provide an effect size estimate with a Bayes factor. You can't only report a Bayes factor - it tells you nothing about the size of an effect. This is such a basic misunderstanding, I stopped reading, but you might want to reconsider getting an expert on board?

Finally, you misunderstand p-values. You are re-hashing arguments by p-value bashers. But not by experts on p-values (e.g., Benjamini, Nickerson, Frick). P-values are ONLY used for error control. Not mentioning that in the intro is the last reason this paper should not be read by novices.

Now it will be read, like crazy, because everyone thinks they need to report Bayes Factors. As I have blogged, equivalence tests outperform Bayes factors for testing the absence of any effect you care about. But to quote your excellent podcast: there are academic hipsters. They want to twist their mustaches, drink machiato's, and report Bayes factors.

There are thousands of 'intro to Bayes' factors resources. And there are 2 intermediate Bayes factorsresources. Everybody wants to know what it is, but no one really goes on to use it. Think about that.

We'll have a fun podcast recording wednesday, talking about this stuff!

Daniel Lakens Here is the critical misunderstanding error (you'll need to remove the criticisms on cohen's d from the paper, or admit you need effect sizes in addition to bayes factors) - also, the Bayes factor can not provide evidence for the presence of an effect.... See More
Like · April 11 at 1:54am · Edited

Daniel Lakens This is very good feedback, great to have extra pairs of eyes looking over this before submission. Looking forward to discussing this topic on our podcast!
Like · April 11 at 2:05am

Kyle Morrissey There are thousands of intro to Bayes factors resources? That was not my experience :S

Though I finally did have someone run me through the conceptual basics in person the other day, and it made sense.
Like · April 11 at 8:28am

Daniel Lakens Kyle , -1 for not saying that the intro in my MOOC was all you needed. You can lead a horse to the water, but you can't make them drink.
Like · April 11 at 8:58am

Stephen Martin P-values really aren't used for error control. That's conflating NP and Fisherian approaches, no?

Piggy backing off this comment thread.... See More
Like · April 12 at 12:33am · Edited

Stephen Martin After reading Donald Williams' response, I thought I should just clarify: I'm all for papers giving 'new' (or at least, newly applied) perspectives on old topics, along with critiques of old perspectives on old topics. I intended my reply to be a critique moreso of BFs and some of the specific arguments, not as a critique of you or your intentions. I realized I never actually made that explicit in my reply above.
Like · 2 · April 12 at 12:36am

Matt Williams >The hypothesis you describe in the intro (is the null true, or is there an effect larger than 0) can only be tested with p-values. It is underspecified for Bayesian stats. In Bayes, the alternative is 'is there a true effect between x and y with the distribution like z'. [Daniel]

>More importantly though, the p(Model | D) can only be interpreted in the family of models that you're testing, but I think people interpret it as "probability I'm correct", [Stephen]

I agree given the standard interpretation of Bayes factors (where the prior on effect size is treated as part of the H1 model itself). But if you separate out the H1 "hypothesis" from the statistical model/prior the problem becomes sort-of-resolvable. This is what I was banging on about in my recent blog: ://thepathologicalscience.blogspot.com/.../separating...

PS. Like Stephen Martin I'm also a Bayesian who doesn't really like Bayes factors, but I'm working on a manuscript at the moment where I have been asked to write an introduction to them for a special issue on methods in a particular sub-area of psych. It's been bloody difficult trying to produce a 'balanced' view of Bayes factors (i.e., balancing reasonable views of frequentists, pro-BF people, and Bayesians who prefer estimation). Thanks Daniel Quintana for provoking a discussion that has been helpful to me in making final revisions.

Separating model from hypothesis in the Bayes factor test
Premise When using statistical analyses, we will often test a statistical model that has one or more parts that we regard as forming an hy...
THEPATHOLOGICALSCIENCE.BLOGSPOT.COM

Like · April 12 at 4:09pm

Daniel Quintana That blog post is really handy, thanks for sharing! We're working on an update now based on everyone's great feedback
Like · April 13 at 4:48am

Donald Williams Hi Daniel Quintana. To all providing comments, I think it is important to remember the likely readership of this article. I imagine this paper is targeted to those in more clinical fields who have not been exposed to much Bayesian stuff. That said, I am not sure I see this as an introduction to Bayes factors, and especially not Bayesian statistics. Instead, I think this is more of an introduction into the doctrine of Rouder, Wagenmakers..etc (i.e., the BF crew) in psychology. Now that there approach has become more common, this has also resulted in finding several limitations in their approach and downright rebukes of their use of statistics (e.g., our paper: Uli Schimmack and Rickard Carlsson.) That said, I think the BF crew does a lot great research, but has also oversold BF and feel as though they have sought extreme examples to show how BF and p differ, but always in favor of their method being superior. That said, rather than introduce this approach circa a few years ago, I see this as a unique opportunity to introduce what might be a "new" method to a field, but also include the recent critiques and other ways of using Bayesian statistics. In this way, we have a fair and balanced paper, and not one slanted towards the BF crew's philosophy that has dominated psychology. Not that Dominant means the approach is necessarily good (or bad), just that they were shouting the loudest and often publishing things that were not novel other than computing a Bayes factor. This resulted in a flurry of opportunistic Bayes factor publications. Those days are hopefully winding down, although now the challenge is that more people are using JASP without really understanding what is going on. I cannot blame them, as the ease with which BF can be manipulated is not really described in any amount of detail--e.g., the infamous prior odds of Bem's ESP. As for the paper, I would steer away from critiquing p-values and instead think of ways we can think about using them. For example, p can be considered as a kind of model fit indices, not for the observed data, but to the null sampling distribution. That is, if we set up a null model (or envision a hypothetical null model), p gives us a measure of departure from that model. The question then becomes contexts in which this is useful, or what needs to accompany p to ensure it is valid and allows for rich inferences--there are lots and lots of assumptions that may or may not make sense depending on the situation, but no less sensible than any statistical quantities assumptions. While much attention has been paid to the Bayesian prior, what is less considered is the chosen likelihood, which is a modeling based decision both frequentist's and Bayesian's make, but Bayesian more explicitly so. That said, Bayesian's do not often examine the influence of distributional departures from the chosen likelihood on the resulting posterior (to my knowledge). These are important issues, as they directly affect the density with which Bayes factors are computed. How does non-normality, unequal variances, treating a count variable as continuous influence the resulting Bayes factor, for example? This says nothing about the importance of fully understanding that BF is a model comparison metric. It provides relative evidence. This generally comes with even odds on the null and alternative. This does not makes much sense, but I have also made this assumption in some of my work. I am not sure this is more unreasonable than testing the value of zero in a frequentist framework, so proceeded but with effect size estimates and intervals on those effects (quantities not provided by Bayes factors). These are important issues, and I see that you have a unique opportunity to introduce the current state of Bayesian methods to your field (prior odds, the importance of the prior, and inferences obtained from the posterior...etc.). This also comes with great responsibility, and I think it would be a shame to align yourself so heavily with the BF crew in their use of not only Bayesian statistics, but also their arguments against p-values.
Like · 5 · April 12 at 12:57am · Edited

Donald Williams Let me also say that I too made many of the similar arguments against p-values in the past. Since then, I learned that p is not evil, and that Bayes factors are not great. They simply are what they are, and the problem really arises from misuse or misunderstandings.
Like · 4 · April 12 at 12:38am

Daniel Quintana Thanks for these comments. In earlier versions of the manuscript we went into a lot more depth (including the importance of the chosen likelihood) but were squeezed for space. The tricky thing here is to make this paper approachable to those who are more clinically oriented, while also appropriately covering all the important issues (and keeping within word limits).
Like · 1 · April 12 at 2:26am

Donald Williams One thing I forgot to mention is whether in clinical oriented work we even care about model selection via bayesian null hypothesis testing? For example, for making treatment decisions, what is more informative: d = 0.30, 95% CI

....I reached out to one of the people who wrote some of the critical feedback and asked if he wanted to join as a co-author.

He agreed 🎉 So with his input and re-writes, along with input from others, the paper was updated to its current version.

Now the paper is under review at a top journal. I also mentioned in the cover letter that the preprint had been downloaded 700+ times

8 Revise your Preprint

Website Resources

Reasons to Worry Less About the Explosion of Preprints - Absolutely Maybe

On the value of preprints: An early career researcher perspective.

What are Preprints, and How Do They Benefit Authors? | AJE

Preprints (PLoS)