

Challenges Citing Preprints and How to Tackle Them

Bruce Rosenblum

Vice President of Content and Workflow Solutions

Inera | An Atypon Company

[@eXtyles](https://twitter.com/eXtyles) | inera.com

Preprint growth

- COVID-19 created
 - an explosion of preprints
 - an explosion in citations to preprints
- “Posting” is moving faster than peer review and publishing

Why are preprint citations challenging?

- Preprint servers don't always identify content as preprints
- Recommended citations may be incomplete
- Author citations are frequently incomplete
- Incomplete metadata makes citation completion and verification challenging

Preprint servers don't always identify content as not peer-reviewed

What we want:

bioRxiv is receiving many new papers on coronavirus SARS-CoV-2. A reminder: these are preliminary reports that have not been peer-reviewed. They should not be regarded as conclusive, guide clinical practice/health-related behavior, or be reported in news media as established information.

What we sometimes see:

MindRxiv: [About](#) | [Support](#) | [Contact](#) |  

MindRxiv is a service provided by the Mind & Life Institute, and is not affiliated with other preprint servers.

Problems when content is not identified as not peer-reviewed

Published on 4/13/21 at <https://www.nationalgeographic.com/science/article/zoom-fatigue-may-be-with-us-for-years-heres-how-well-cope>

New research published today offers some of the first data-backed conclusions about Zoom fatigue and provides a broad look at the causes. It also shows that the weight of Zoom fatigue is not equally distributed. In a survey of more than 10,000 people, described online today on the research-sharing platform known as SSRN, women reported experiencing about 13.8 percent more Zoom fatigue than men, on average.

Is it a preprint?
How can we
tell?

Links to:

Fauville, Geraldine and Luo, Mufan and Queiroz, Anna C. M. and Bailenson, Jeremy N. and Hancock, Jeff, Nonverbal Mechanisms Predict Zoom Fatigue and Explain Why Women Experience Higher Levels than Men (April 5, 2021). Available at SSRN: <https://ssrn.com/abstract=3820035> or <http://dx.doi.org/10.2139/ssrn.3820035>

Recommended citations on preprint servers don't indicate preprint status

Top 3 recommended citations from OSF Preprints:

APA

Beck, J., Ferguson, C. A., Funk, K., Hanson, B., Harrison, M., Ide-Smith, M. B., ... Swaminathan, S. (2020, July 21). Building trust in preprints: recommendations for servers and other stakeholders. <https://doi.org/10.31219/osf.io/8dn4w>

MLA

Beck, Jeffrey, et al. "Building Trust in Preprints: Recommendations for Servers and Other Stakeholders." OSF Preprints, 21 July 2020. Web.

Chicago

Beck, Jeffrey, Christine A. Ferguson, Kathryn Funk, Brooks Hanson, Melissa Harrison, Michele B. Ide-Smith, Rachael Lammey, et al. 2020. "Building Trust in Preprints: Recommendations for Servers and Other Stakeholders." OSF Preprints. July 21. doi:10.31219/osf.io/8dn4w.

Is it a preprint?
How can we
tell?

Recommended citations on preprint servers may not include a DOI

Suggested citation from OSF Preprints:

MLA

Beck, Jeffrey, et al. "Building Trust in Preprints: Recommendations for Servers and Other Stakeholders." OSF Preprints, 21 July 2020. Web.

Authors' preprint citations often don't include a DOI

Examples of real-world author-submitted preprint citations:

80. → Madder, H., I. Moir, R. Moroney, L. Butcher, R. Newnham, M. Sunderland, T. Clarke, D. Foster, P. Hoffman, G. Moore, C.S. Brown, and K.J.M. Jeffery, *Multiuse patient monitoring equipment as a risk factor for acquisition of Candida auris*. [bioRxiv](#), 2017.¶

[12] A. E. [Hassanien](#), L. N. [Mahdy](#), K. A. Ezzat, H. H. [Elmousalami](#), H. A. Ella, Automatic X-ray COVID-19 Lung Image Classification System based on Multi-Level Thresholding and Support Vector Machine, [medRxiv](#).¶

Recommendations for preprint citations

- Update recommended citations, journal style guides, reference managers, and author instructions for preprint citations to always include these elements:
 - Author/s
 - Date of posting
 - Preprint title
 - Preprint server name
 - “Preprint” indicator
 - DOI

Example of a “good” preprint citation

➤ From the *AMA Manual of Style*:

1. Bloss CS, Wineinger NE, Peters M, et al. A prospective randomized trial examining health care utilization in individuals using multiple smartphone-enabled biosensors. bioRxiv. Preprint posted online October 28, 2015. doi:10.1101/029983

Automated link formation between preprint and article doesn't always work

The preprint title:

A simple model to assess Wuhan lock-down effect and region efforts during COVID-19 epidemic in China Mainland

- Zheming Yuan, Yi Xiao, Zhijun Dai, Jianjun Huang & Yuan Chen

Posted: 2 March 2020

<http://dx.doi.org/10.2471/BLT.20.254045> 

The article title:

Modelling the effects of Wuhan's lockdown during COVID-19, China

Zheming Yuan,^a Yi Xiao,^a Zhijun Dai,^a Jianjun Huang,^a Zhenhai Zhang^b & Yuan Chen^b

Source: *Bulletin of the World Health Organization*

Crossref returns inconsistent results for the same query

Same reference entry, two different Crossref query results:

77. Fan H, Zhang L, et al. Retrospective analysis of clinical features in 101 death cases with COVID-19. medRxiv 2020

77. Fan H, Zhang L, et al. Retrospective analysis of clinical features in 101 death cases with COVID-19. medRxiv 2020 <https://doi.org/10.1101/2020.03.09.20033068>.

Crossref reports this preprint has been published as 10.1017/s0950268820002010 (Ref. 77 "Fan, Zhang, et al., 2020")

Authors may post preprints to two or more preprint servers

Lancet preprint on SSRN:

Rodriguez Llanes, Jose Manuel and Castro Delgado, Rafael and Pedersen, Morten Gram and Arcos Gonzalez, Pedro and Meneghini, Matteo, Confronting COVID-19: Surging Critical Care Capacity in Italy (3/25/2020). Available at SSRN: <https://ssrn.com/abstract=3564386> or <http://dx.doi.org/10.2139/ssrn.3564386>

Bulletin of the World Health Organization preprint:

Rodriguez-Llanes JM, Castro Delgado R, Pedersen MG, Arcos González P & Meneghini M. Confronting COVID-19: Surging critical care capacity in Italy. [Preprint]. *Bull World Health Organ*. E-pub: 6 April 2020. doi: <http://dx.doi.org/10.2471/BLT.20.257766>

Preprint on medRxiv:

Confronting COVID-19: Surging critical care capacity in Italy
Jose M Rodriguez-Llanes, Rafael Castro Delgado, Morten Gram Pedersen, Pedro Arcos González, Matteo Meneghini
medRxiv 2020.04.01.20050237; doi: <https://doi.org/10.1101/2020.04.01.20050237>

Preprint sites may replace preprint with published journal article

Wyatt, K., Griffin, R., Guerry, A., Ruckelshaus, M., Fogarty, M., & Arkema, K. K. (2018, January 26). Habitat risk assessment for regional ocean planning in the U.S. Northeast and Mid-Atlantic [preprint]. MarXiv. <https://doi.org/10.31230/osf.io/b3y2w>



Now resolves to...

The image shows two overlapping screenshots. The top screenshot is the MarXiv preprint page for the article 'Habitat risk assessment for regional ocean planning in the U.S. Northeast and Mid-Atlantic'. It features the MarXiv logo, search and sign-in options, and the authors' names: Katherine Wyatt, Robert Griffin, Anne D. Guerry, Mary Ruckelshaus, Michael Fogarty, and Katie K. Arkema. The bottom screenshot is the PLOS ONE journal article page for the same title. It includes the PLOS ONE logo, a 'Check for updates' button, and the full citation: 'Katherine H. Wyatt^{1,2*}, Robert Griffin^{1,2*}, Anne D. Guerry^{1,2,3}, Mary Ruckelshaus^{1,2*}, Michael Fogarty¹, Katie K. Arkema^{1,2}'. The article is marked as 'OPEN ACCESS' and provides a citation: 'Wyatt KH, Griffin R, Guerry AD, Ruckelshaus M, Fogarty M, Arkema KK (2017) Habitat risk assessment for regional ocean planning in the U.S. Northeast and Mid-Atlantic. PLOS ONE 12(12): e0188776. <https://doi.org/10.1371/journal.pone.0188776>'. The abstract is visible, starting with 'Coastal habitats provide important benefits to people, including habitat for species targeted by fisheries and opportunities for tourism and recreation. Yet, such human activities also can impair these habitats and undermine the ecosystem services they provide to people. Cumulative risk assessment provides an analytical framework for synthesizing the influence of multiple stressors across habitats and decision-support for balancing human uses and ecosystem health. To explore cumulative risk to habitats in the U.S. Northeast and Mid-Atlantic Ocean Planning regions, we apply the open-source ioVEST Habitat Risk Assess-'. The PLOS ONE page also features a 'Download paper' button, a 'plaudit' endorsement section, and social media sharing options.

Preprint metadata may not indicate withdrawn status

Preprint on bioRxiv:

WITHDRAWN Comments (131)

Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag

Prashant Pradhan, Ashutosh Kumar Pandey, Akhilesh Mishra, Parul Gupta, Praveen Kumar Tripathi, Manoj Balakrishnan Menon, James Gomes, Perumal Vivekanandan, Bishwajit Kundu
doi: <https://doi.org/10.1101/2020.01.30.927871>

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

Abstract Info/History Metrics Preview PDF

Abstract

This paper has been withdrawn by its authors. They intend to revise it in response to comments received from the research community on their technical approach and their interpretation of the results. If you have any questions, please contact the corresponding author.

But ... Crossref provides no system to indicate withdrawn status in metadata.

Our recommendations for preprints

- Clearly identify non-peer-reviewed content *at all stages*
 - NISO: Create PIE-J-type recommended practice for preprints
- Update citation styles to include all recommended elements
- Implement metadata integrations & common workflows to ensure preprints are:
 - treated as a unique citation type
 - promptly updated with publication status
 - never replaced with the published version
 - never deposited to Crossref as journal articles
- Create systems to automatically
 - link published articles with their preprints
 - identify withdrawn preprints

Recommended reading

- Izzo Hunter, S., I. Kleshchevich, and B. Rosenblum. “What’s wrong with preprint citations?” *The Scholarly Kitchen*, 18 September 2020.
<https://scholarlykitchen.sspnet.org/2020/09/18/guest-post-whats-wrong-with-preprint-citations/>
- Beck, J., et al. “Building trust in preprints: Recommendations for servers and other stakeholders.” OSF Preprints, 21 July 2020.
<https://doi.org/10.31219/osf.io/8dn4w>
- Patel, J. “Opinion: Preprints in the Public Eye.” *The Scientist*, 18 March 2021.
<https://www.the-scientist.com/news-opinion/opinion-preprints-in-the-public-eye--68563>

Thank you!