

令和 5 年 6 月 9 日現在

機関番号：32710

研究種目：基盤研究(C) (一般)

研究期間：2019～2022

課題番号：19K12707

研究課題名(和文) オープンサイエンスの視点から見たプレプリントの影響の研究 - 新学術情報流通の解明 -

研究課題名(英文) A study on influence of preprint from the point of open science - Clarification of new academic information distribution -

研究代表者

角田 裕之 (Tsunoda, Hiroyuki)

鶴見大学・文学部・教授

研究者番号：30454961

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本研究は世界の研究者が積極的に利用しているプレプリントを活用し、オープンサイエンスの推進に貢献するために計画した。2010年代以降に多数のアーカイブが創設された。本研究は生物学を主題としたプレプリントのbioRxivに2013年11月から2020年2月に投稿されたプレプリントを出版した119の学術雑誌を調査した。多くの雑誌において、プレプリントを経由して学術雑誌に掲載された論文は、その学術雑誌に掲載されたすべての論文の査読期間より短かった。結論として、本研究はプレプリントが研究成果の公開時間を短縮させ、学術情報流通に影響を与えていることを明らかにした。

研究成果の学術的意義や社会的意義

研究者は公的な資金を得て研究活動に従事している。研究成果は即座に広く社会に公開され、多くの研究に活用されることが望まれる。査読は論文の質を保証する重要な審査であり、複数のピアが審査を担当する。よって、多くの雑誌では査読に数か月を要する傾向がある。雑誌に掲載された論文の査読期間を短縮できれば、研究成果の公開が促進につながる。本研究は、プレプリントのbioRxivとジャーナル論文との査読期間の比較分析を通して、プレプリントが査読期間に与える影響を明らかにすることに学術的意義と社会的意義がある。

研究成果の概要(英文)：This study was planned to contribute to the promotion of open science by utilizing preprints, which are actively used by researchers around the world. Several subject archives have been created since the 2010s. The study examined 119 journals that published preprints submitted to bioRxiv in the field of biology between November 2013 and February 2020. In many journals, the peer review period of journal articles via preprints have significantly shorter than peer review period of the journals articles as a whole. In conclusion, this study reveals that preprints reduce the publication time of research results and have an impact on scholarly communication.

研究分野：図書館情報学

キーワード：プレプリント 学術情報流通 査読期間 計量情報学 図書館情報学 科学計量学 オープンサイエンス 学術雑誌

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

生命科学分野ではプレプリントの増加が見られ、ジャーナルや資金提供者の方針転換につながっている (Abdill, 2019)。2017 年以前に投稿されたプレプリントの 3 分の 2 は、後に査読付きジャーナルに掲載され、プレプリントのダウンロード回数は、プレプリントが掲載されたジャーナルのインパクトファクターと関連している (Tsunoda, 2019)。Fraser ら (2021) は、COVID-19 の世界的な大流行を背景に、科学を迅速かつオープンに共有することの重要性と、この取り組みにおけるプレプリントの重要な役割を示した。世界的な COVID-19 パンデミック時のように、後発の研究結果が特別な関心を集める場合、プレプリントはさらに大きな役割を果たした (Vlasschaert, Giles, Hiremath & Lanktree, 2020)。COVID-19 の流行が進むにつれ、プレプリントアーカイブサービスはリアルタイムの学術動向を評価するための有用な指標となるであろう (Wehner, Li, & Nead, 2020)。COVID-19 の流行期には、bioRxiv プラットフォームは研究成果の迅速な共有という特徴を十分に反映し、bioRxiv は再び学者から高い注目を浴びた。しかし、本研究では、調査期間が 2020 年 2 月迄であり、COVID-19 の大流行時に投稿された資料は対象外である。

2. 研究の目的

現在、bioRxiv は 242 のジャーナルと提携しており、原稿ファイルやメタデータを bioRxiv から直接送信することで、ジャーナルや査読サービスに論文を投稿する際の著者の時間を短縮できるように取り組んでいる。プレプリントが出版プロセスに影響を与えるかどうかを確認するために、特に bioRxiv とジャーナル論文の間の査読プロセスの期間に焦点を当てた。Tsunoda et al. (2020) は、プレプリントを通じて発表された論文の多くがジャーナルに掲載され、ジャーナルに投稿する前にプレプリントに投稿された論文の方が査読期間が短いことを明らかにした。本研究では、2013 年 11 月から 2020 年 2 月までの間に、119 の学術雑誌における bioRxiv からジャーナルへの投稿 (B2J) と著者からジャーナルへの投稿 (A2J) の論文の投稿時間と掲載時間を調べ、B2J と A2J 論文の期間時間の違いを比較し、プレプリントが出版プロセスにどのように影響するかを明らかにした。

3. 研究の方法

(1) プレプリントと雑誌論文のメタデータの収集

本処理では、2013 年 11 月から 2019 年 2 月までに bioRxiv に投稿された論文 (以下、プレプリント) を、サーバーのすべてのページを訪問してメタデータ (プレプリントのタイトル、著者、投稿日、バージョン、DOI) を収集する Web クローラーを用いて調査した。さらに、プレプリントがジャーナルに掲載されている場合は、ジャーナルタイトル、論文タイトル、ジャーナル論文の DOI も収集した。ジャーナル論文は、プレプリントと DOI の一致を確認し、PubMed の Publication History Status Date から査読受付日、出版受理日を取得した。

(2) 査読期間の定義

論文の査読期間を論文の査読の受け付けから出版が受理されるまでの期間と定義した。つまり、論文の査読期間は、出版の受理日から査読の受付日を引いたものに等しい。よって、査読期間は 0 日よりも長くなるが、マイナスにはならない。

(3) 投稿種別の定義

プレプリントへの投稿がジャーナルへの投稿より 1 週間以上早い場合、投稿タイプはプレプリントファーストと呼ぶ。順序が逆の場合は、ジャーナルファーストと呼ぶ。この 2 つの行為が 1 週間以内に行われ場合、同時進行と呼ぶ。これらの論文は、3 つの投稿タイプのいずれかに分類される。著者が直接ジャーナルに論文を投稿する場合は、A2J (author to journal) と呼ぶ。一方、著者がまずプレプリントで論文を投稿し、その後ジャーナルに投稿する場合は、B2J (bioRxiv to journal) と呼ぶ。A2J はすべてのジャーナル論文を指し、B2J は A2J のサブセットである。

(4) 仮説の検証

査読期間は右肩上がりの分布であり、その分布の正規性を大きく逸脱していることが明らかになった。一方、A2J はジャーナルの調査年に出版されたすべての論文、B2J はプレプリントによって得られた論文で構成されているため、両者の大きさは異なり分散の均質性の仮定は成り立たない。よって、本研究ではノンパラメトリック検定であるブルンナー・ムンツェル検定 (Brunner-Munzel test) を用いて、査読期間に関する B2J と A2J の中央値を比較した。

4. 研究成果

図1は、119誌の査読期間の中央値の散布図である。各バブルは雑誌を示す。X軸にA2J、Y軸にB2Jの査読期間の中央値をとり、B2Jの論文数はバブルの大きさで表している。赤いバブルは0.05以下の水準で有意差があることを示し、青いバブルは有意差がないことを示す。対角線より下のジャーナルでは、B2Jの査読期間の中央値がA2Jのそれよりも短いことが分かる。一方、対角線上のジャーナルはその逆を示し、対角線上のジャーナルは両者が同等であることを示す。

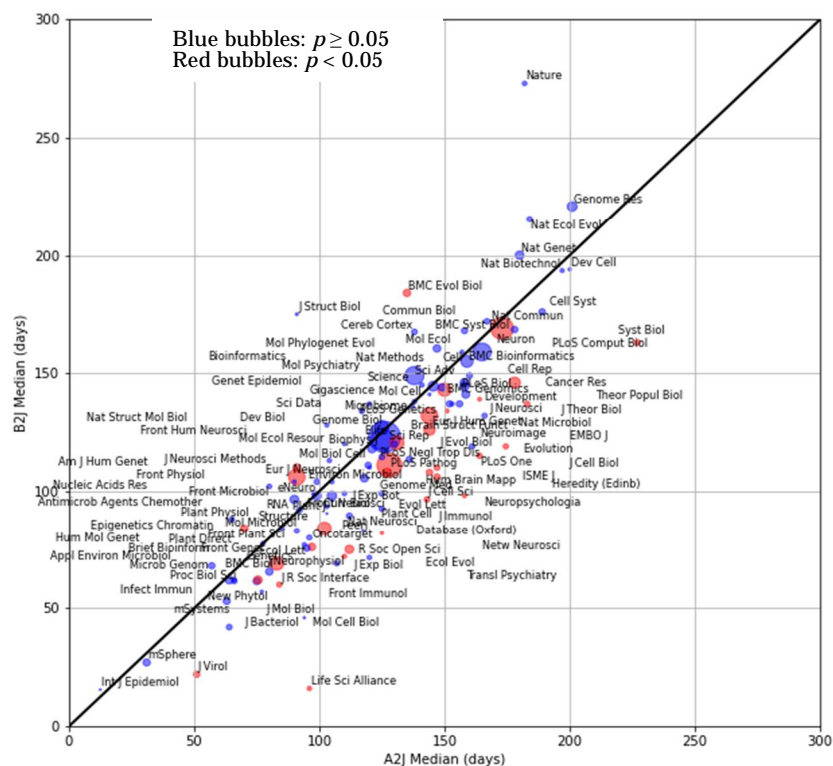


図1 査読期間の中央値（2014年から2019年全体）

表1は、査読期間の有意差に基づく雑誌数を2014年から2019年全体で示した。表1から、調査対象となった119誌のうち、85誌がA2JよりB2Jが短く、30誌がその逆、4誌が同等、33誌が0.05以下の有意差、29誌がA2JよりB2Jが短い査読期間、4誌がその逆であることが分かる。

表1 査読期間の有意差に基づく雑誌数（2014年から2019年全体）

Median of Peer Review	A2J > B2J		A2J = B2J		A2J < B2J	
Number of Journals	85		4		30	
Brunner–Munzel test	$p < 0.05$	$p \geq 0.05$	$p < 0.05$	$p \geq 0.05$	$p < 0.05$	$p \geq 0.05$
Number of Journals	29	56	0	4	4	26
Proportion	0.244	0.471	0.00	0.034	0.034	0.218

図2は、査読期間の中央値は2014年から2019年まで各年で示した。6つの図のバブル(雑誌)の総数は463であり、年が経過するにつれて増加する傾向にあることが分かる。2016年からバブルが増え、2018年が最も多かった。しかし、本調査は2020年の早い時期に行われたため、2019年は十分な数を得ることができなかった。多くのバブルは対角線の下方に描かれ、特に赤バブルは大部分が対角線の下側であることが分かる。

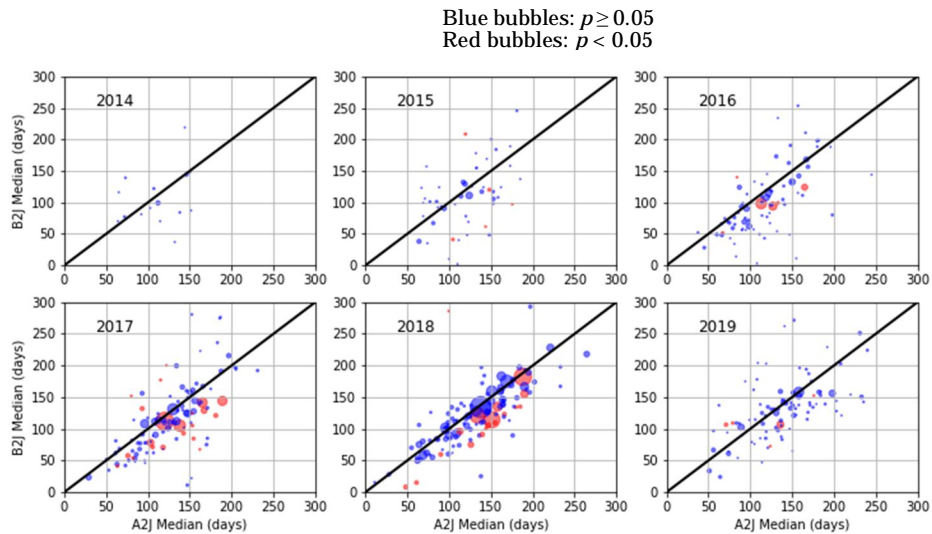


図 2 査読期間の中央値 (2014 年から 2019 年各年)

表 2 は、査読期間の有意差に基づく雑誌数を 2014 年から 2019 年まで各年で示した。A2J > B2J の左列 (背景色グレーのセル) はプレプリントを経由した論文の査読期間が有意差 (0.05) で短い割合を示した。サンプル数が比較的多い 2016 年では、0.063、2017 年では 0.155、2018 年では 0.147 であった。(太字部分)

表 2 査読期間の有意差に基づく雑誌数 (2014 年から 2019 年各年)

Median of Peer Review Period	A2J > B2J		A2J = B2J		A2J < B2J	
	$p < 0.05$	$p \geq 0.05$	$p < 0.05$	$p \geq 0.05$	$p < 0.05$	$p \geq 0.05$
2014		9				6
Proportion		0.600				0.400
2015	4	24			1	20
Proportion	0.082	0.490			0.020	0.408
2016	5	51			1	22
Proportion	0.063	0.646			0.013	0.278
2017	17	56		1	5	31
Proportion	0.155	0.509		0.009	0.045	0.282
2018	17	62		1	1	35
Proportion	0.147	0.534		0.009	0.009	0.302
2019	3	52		2	2	35
Proportion	0.032	0.553		0.021	0.021	0.372

表 3 は B2J の論文が多い 10 誌の査読期間の中央値である。例えば、*PLoS One* の査読期間の中央値が A2J では 128 日、B2J では 111 日であることを示す。つまり *PLoS One* は、プレプリントを経由した論文の査読期間が 17 日短かった ($p < 0.01$)。逆に、*Nucleic Acids Res* は、プレプリントを経由した論文の査読期間が 15 日長かった ($p < 0.05$)。

表3 B2Jの論文数の多い10誌(2014年から2019年全体)

Journal	A2J		B2J		Difference	
	Number of Articles	Median of Peer Review Period	Number of Articles	Median of Peer Review Period	Days	
<i>Sci Rep</i>	96,973	126	573	123.0	3	
<i>PLoS One</i>	133,823	128	343	111.0	17	**
<i>Nat Commun</i>	24,447	173	286	169.0	4	*
<i>Elife</i>	6,402	125	273	122.0	3	
<i>Bioinformatics</i>	3,719	138	198	149.0	- 11	
<i>PLoS Comput Biol</i>	3,363	165	188	159.0	6	
<i>PLoS Genetics</i>	3,762	144	153	132.0	12	**
<i>Nucleic Acids Res</i>	6,030	91	150	106.0	- 15	*
<i>Genome Biol</i>	1,127	125	111	127.0	- 2	
<i>Neuroimage</i>	5,455	131	110	120.5	10.5	*

注記: ** $p < 0.01$, * $p < 0.05$ (A2JとB2J間のブルンナー・ムンツェル検定による)

結論と考察

本研究では、2014年から2019年の33ジャーナルで、B2JとA2Jの査読期間に0.05レベルで有意差があることが示された。うち29ジャーナルはB2JがA2Jよりも早く出版されていると判断した。さらに、各年の分析において、2017年は110誌中17誌、2018年は116誌中17誌が同様の傾向を示し、逆の傾向を示したのは各年1誌のみだった。このことは、プレプリントプラットフォームでの共有が、査読に要する時間の短縮に役立つ可能性を示唆している。プレプリントは学術コミュニケーションを促進するための有効なツールとなり得る。プレプリント数はジャーナル論文の総数に比べればまだ少ない。学術ジャーナルはプレプリントと公開結果を受け入れ、ジャーナル論文の査読の参考とすることがわかった。この現象は出版プロセスにも影響を与え、学術雑誌における査読期間や査読方法の変更を促し、論文の出版を加速させる可能性がある。

引用文献

- Abdill RJ, Blekman R. (2019). Meta-Research: tracking the popularity and outcomes of all bioRxiv preprints. *ELife*, 8, e45133. <https://doi.org/10.7554/eLife.45133.001>.
- Fraser, N., Brierley, L., Dey, G., Polka, J.K., Pálffy, M., Nanni, F., & Coates, J.A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS biology*, 19(4), e3000959. <https://doi.org/10.1371/journal.pbio.3000959>.
- Tsunoda, H., Sun, Y., Nishizawa, M., Liu, X. & Amano, K. (2019). An analysis of published journals for papers posted on bioRxiv. *Proceedings of the Association for Information Science and Technology*, 56, 783-784. <https://doi.org/10.1002/pr2.175>.
- Tsunoda, H., Sun, Y., Nishizawa, M., Liu, X., & Amano, K. (2020). The influence of bioRxiv on PLOS ONE's peer-review and acceptance time. *Proceedings of the Association for Information Science and Technology*, 57, e398. <https://doi.org/10.1002/pr2.398>.
- Vlasschaert, C., Giles, C., Hiremath, S., & Lanktree, MB. (2021). *Journal of the American Society of Nephrology*2021, 16(3), 479-486. <https://doi.org/10.2215/CJN.03800320>.
- Wehner, M.R., Li, Y., & Nead, K.T. (2020). Comparison of the Proportions of Female and Male Corresponding Authors in Preprint Research Repositories Before and During the COVID-19 Pandemic. *JAMA Network Open* 2020, 3(9), e2020335. <https://doi.org/10.1001/jamanetworkopen.2020.20335>.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 4件/うちオープンアクセス 1件）

1. 著者名 Tsunoda Hiroyuki, Sun Yuan, Nishizawa Masaki, Liu Xiaomin, Amano Kou	4. 巻 56
2. 論文標題 An analysis of published journals for papers posted on bioRxiv	5. 発行年 2019年
3. 雑誌名 Proceedings of the Association for Information Science and Technology	6. 最初と最後の頁 783 ~ 784
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/pr2.175	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Tsunoda Hiroyuki, Sun Yuan, Nishizawa Masaki, Liu Xiaomin, Amano Kou	4. 巻 57
2. 論文標題 The influence of bioRxiv on PLOS ONE's peer review and acceptance time	5. 発行年 2020年
3. 雑誌名 Proceedings of the Association for Information Science and Technology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/pr2.398	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Tsunoda Hiroyuki, Sun Yuan, Nishizawa Masaki, Liu Xiaomin, Amano Kou	4. 巻 -
2. 論文標題 Study of the Influence of Preprint in bioRxiv for Peer Review and Acceptance Time of PLOS ONE	5. 発行年 2020年
3. 雑誌名 bioRxiv	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/2020.01.28.920546	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Tsunoda Hiroyuki, Sun Yuan, Nishizawa Masaki, Liu Xiaomin, Amano Kou	4. 巻 59
2. 論文標題 How Preprint Affects the Publishing Process: Duration of the Peer Review Process between bioRxiv and Journal Papers	5. 発行年 2022年
3. 雑誌名 Proceedings of the Association for Information Science and Technology	6. 最初と最後の頁 505 ~ 509
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/pr2.660	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

[学会発表] 計4件(うち招待講演 0件/うち国際学会 4件)

1. 発表者名 Hiroyuki Tsunoda , Yuan Sun , Masaki Nishizawa , Xiaomin Liu , Kou Amano
2. 発表標題 Current status and flow analysis of posted papers in Biorxiv
3. 学会等名 IFLA 's 85th World Library and Information Congress (国際学会)
4. 発表年 2019年

1. 発表者名 Tsunoda Hiroyuki、Sun Yuan、Nishizawa Masaki、Liu Xiaomin、Amano Kou
2. 発表標題 An Analysis of Published Journals for Papers Posted on bioR iv
3. 学会等名 82nd Annual Meeting of the Association for Information Science and Technology (国際学会)
4. 発表年 2019年

1. 発表者名 Tsunoda Hiroyuki、Sun Yuan、Nishizawa Masaki、Liu Xiaomin、Amano Kou
2. 発表標題 The influence of bioR iv on PLOS ONE's peer review and acceptance time
3. 学会等名 83rd Annual Meeting of the Association for Information Science and Technology (国際学会)
4. 発表年 2020年

1. 発表者名 Tsunoda Hiroyuki、Sun Yuan、Nishizawa Masaki、Liu Xiaomin、Amano Kou
2. 発表標題 How Preprint Affects the Publishing Process: The Duration of the Peer Review Process between bioRxiv and Journal Papers
3. 学会等名 85rd Annual Meeting of the Association for Information Science and Technology (国際学会)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	孫 媛 (Sun Yuan) (00249939)	国立情報学研究所・情報社会相関研究系・准教授 (62615)	
研究分担者	西澤 正己 (Nishizawa Masaki) (00281585)	国立情報学研究所・情報社会相関研究系・准教授 (62615)	
研究分担者	天野 晃 (Amano Kou) (20622012)	国立情報学研究所・オープンサイエンス基盤研究センター・ 特任研究員 (62615)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関		
中国	National Science Library	Chinese Academy of Sciences	