

프리프린트의 피인용과 사회적 주목에 관한 연구

- medRxiv를 중심으로 -

A Study on Social Attention and Citations of Preprints: Focusing on medRxiv

조 재 인 (Jane Cho)*

목 차

- | | |
|------------|-------------------|
| 1. 연구의 필요성 | 4. 데이터 수집 및 연구 방법 |
| 2. 문헌고찰 | 5. 분석결과 |
| 3. 선행연구 | 6. 결 론 |

초 록

프리프린트로 기탁된 미검증 논문에 대해 이용자는 연구 내용의 신뢰성과 가치를 스스로 판단해야 한다. 본 연구는 프리프린트의 이용 지표로 추후 이 논문의 공식 출판본이 가질 영향력을 미리 예측할 수 있는지 알아보았다. 더불어 미검증 프리프린트가 언론보도 등을 통해 사회적으로 확산되고, 학술논문에서도 인용되고 있는지 분석하였다. 이를 위해 medRxiv health informatics 분야에 프리프린트로 최초 아카이빙된 논문 288건과 해당 논문의 공식 출판본 172건, 총 460건을 수집하여 분석을 수행하였다. 그 결과, 첫 번째로 프리프린트로 발표된 연구 결과에 대한 초기 관심이 출판 후 버전의 피인용 횟수와 관련성을 보이는 것으로 확인되었다. 회귀 분석 결과 초기 6개월간 누적된 이용 지표를 사용하면 향후 공식 출판본에서 나타날 수 있는 영향력을 40% 선($\beta = 0.409$ $p < 0.01$)에서 예측할 수 있을 것으로 분석되었다. 두 번째, 미검증 프리프린트도 언론매체와 정책문서 등에 언급되면서 사회적으로 확산되고 있으며, 영향력 지수(SJR)가 높은 우수 학술지를 통해 인용되고 있는 것으로 분석되었다.

ABSTRACT

For unverified papers deposited as preprints, users must independently judge the reliability and accuracy of the research content. Therefore, this study aims to determine whether the influence of the officially published version of the paper can be predicted in advance using various usage indicators of preprints. In addition, we analyzed whether, despite being an unverified, unpublished preprint, it was being spread socially through media reports and cited in academic papers. Total of 460 papers were collected and analyzed, including 288 papers initially archived as preprints in the medRxiv health informatics field and 172 post-publication versions of the papers. As a result, first, it was confirmed that the initial interest in research results published as a preprint was related to the number of citations in the post-publication version. It was analyzed that using the usage indicators accumulated over the first six months could predict the influence that may appear in future post-publication versions at around 40% ($\beta = 0.409$ $p < 0.01$), helping users to select research results. Second, it was analyzed that unpublished preprints were also spreading socially and being cited in academic papers. In addition, it was analyzed that it is being actively cited not only in other preprints but also in excellent academic journals with a high impact factor.

키워드: 프리프린트, 학술지, 공식 출판본, 피인용, 알트메트릭스, medRxiv
Preprint, Journal, Post-Print, Citation, Altmetrics, medRxiv

* 인천대학교 문헌정보학과 교수(chojane@inu.ac.kr / ISNI 0000 0004 6016 5878)

논문접수일자: 2025년 4월 16일 최초심사일자: 2025년 5월 3일 게재확정일자: 2025년 5월 7일

한국문헌정보학회지, 59(2): 25-42, 2025. <http://dx.doi.org/10.4275/KSLIS.2025.59.2.025>

© Copyright © 2025 Korean Society for Library and Information Science

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 연구의 필요성

피어리뷰 저널 게재에는 많은 시간이 소요되기 때문에 연구 결과의 신속한 배포를 위해 프리프린트 서버에 논문을 기탁하는 사례가 최근 증가하고 있다. 특히 COVID-19 팬데믹 상황에서 급속하게 증가했는데, 최초 감염 사례로부터 10개월 이내에 125,000건 이상의 관련 논문이 발표되었으며, 이 중 30,000건 이상이 프리프린트 서버에 호스팅되었다고 알려져 있다 (Fraser et al., 2021). COVID-19 관련 지식의 보급과 보건, 의료 정책개발을 위해 프리프린트의 활용 범위가 넓어지면서 학술 출판계에서 프리프린트의 역할과 위상은 새롭게 재정의되고 있다.

저자는 논문의 윤리 및 무결성에 대한 책임을 스스로 지며 모든 사람이 공개적으로 사용할 수 있는 프리프린트 서버에 논문을 기탁할 수 있다. 동시에 다른 곳에서의 출판을 위해 논문을 투고할 수 있으며, 수정, 보완할 수 있는 권리를 보유한다. 연구자들이 프리프린트 서버에 논문을 기탁하는 동기는 시간이 많이 소요되는 피어리뷰 프로세스를 우회하여 좀 더 조기에 학술 커뮤니티에 연구 결과를 발표할 수 있으며, 오픈 액세스를 통해 누구나 무료로 이용할 수 있도록 한다는 점이다. 그 뿐 아니라 논문에 대한 관심을 초기부터 집중시켜 알트메트릭스(altmetrics) 지표와 피인용 횟수를 증대시킬 수 있다 (Fraser et al., 2020).

널리 알려진 프리프린트 서버로는 arXiv와 SSRN이 있으며 의생명 분야에는 bioRxiv 및 medRxiv가 급성장하고 있다 (Smart, 2022). 그동안 저자의 프리프린트 기탁에 있어 가장 큰

걸림돌은 학술지측이 프리프린트 서버를 통해 미리 공개된 논문의 출판을 수락할 것인가 이었다. 그러나 학계의 프리프린트에 대한 관심이 증가하면서 이를 수용하는 학술지가 증가하게 되었다. 최근에는 프리프린트 플랫폼과 파트너십을 채택하거나 자체 프리프린트 서버를 출시하여 투고 및 편집 처리 워크플로우에 프리프린트 공개를 통합하는 사례도 증가하고 있다 (Schonfeld & Rieger, 2020).

그러나 프리프린트는 검증 과정에서 수정과 삭제가 이루어지기도 하며 종종 다른 내용으로 출판되거나 아예 출판이 거절되어 버리기도 한다. 더구나 피어리뷰를 거치지 않은 논문이기 때문에 이용자들은 스스로 연구 내용의 신뢰성과 가치를 판단해야 하는데, 이 과정에서 저자의 소속 국가나 기관 (Nishioka et al., 2022), 또는 이용 통계에 의존할 수 밖에 없다.

본 연구는 이와 같은 배경에서 프리프린트의 각종 지표가 추후 공식 출판될 논문의 영향력을 예측할 수 있는지, 그 중 어떠한 지표가 의미 있는 지표로 활용 가능한지 medRxiv의 health informatics 분야에 호스팅된 프리프린트를 활용해 분석해 보고자 한다. 또한 프리프린트 서버에는 피어리뷰가 완료되어 공식 출판된 논문과 피어리뷰 중이거나 출판이 거절되어 미검증 상태로 남겨진 논문이 혼재하는데, 그 중 미검증 프리프린트도 언론보도, 소셜 미디어 등을 통해 사회적으로 확산되고 있는지, 후속 논문을 통해 피인용되고 있는지 추적해 보고자 한다. 본 연구는 위와 같은 목표를 위해 다음과 같은 연구 질문을 설정해 본다. 첫 번째, 프리프린트에 집중된 관심으로 공식 출판본의 피인용을 예측할 수 있는가? 어떠한 지표가 프리프린

트 선별에 도움이 될 것인가? 두 번째, 검증되지 않은 프리프린트도 사회적으로 주목되거나 학술논문에서 피인용되고 있는가?

2. 문헌고찰

프리프린트는 공식 출판 이전에 프리프린트 서버에 공개되는 학술 논문의 출판 전 버전이다. 피어리뷰 절차에서는 논문의 가정, 방법, 결론의 약점 등이 식별되는데, 보통 심사위원의 지적과 우려 사항이 해결된 경우에만 공식 출판될 수 있다. 이 과정이 오래 걸릴 수 있기 때문에 많은 연구자는 피어리뷰 전에 논문을 프리프린트 형태로 공개하여 빠르게 연구 결과가 공유될 수 있도록 하고 있다. 그러나 완성본이 아니므로 오류가 포함될 수 있으며, 공식 출판본과 내용적 차이가 있을 수 있다.

가장 잘 알려진 프리프린트 서버로는 arXiv(1991년)와 SSRN(1994년)이 있으며, 의생명분야에서는 bioRxiv(2013년) 및 medRxiv(2019년)가 대표적이다. 그 밖에 OSF Preprints, HAL, EconStor, RePEc, PhilArchive, Research Square, viXra, Cryptology ePrint Archive, Preprints.org, ChinaXiv, JMIR Preprints, Authorea, ChemRxiv, engrXiv, e-LiS, SciELO, PsyArXiv, F1000 Research 및 Zenodo 등이 있다(Smart, 2022). 프리프린트 플랫폼은 비영리 학술단체, 학회, 연구비 지원 기관 등이 운영하거나 출판사 등에서 운영하고 있다. 다양한 분야의 콘텐츠를 수용하지만 일부는 연구 기금, 지리적 또는 제휴 저널과 관련하여 기탁에 제한이 있기도 하다.

대표적인 의생명 분야 프리프린트 서버인

medRxiv(<https://www.medrxiv.org>)를 통해 프리프린트 기탁 및 운영 방식을 살펴보면 다음과 같다. 첫 번째, 프리프린트가 기탁되면 표절, 비과학적 내용 여부 및 윤리적 문제에 대한 사전 검토가 이루어진다. 또한 medRxiv에 호스팅된 모든 프리프린트에는 해당 내용이 동료심사에 의해 검증되지 않았으며 임상 결정을 하는데 사용되어서는 안된다는 진술이 수반된다. 두 번째, 저자는 저작권을 보유하며 공개된 콘텐츠에 대한 책임을 가진다. 그리고 저자가 선택한 CC(creative commons) 라이선스 조건에 따라 재사용 여부가 결정된다. 세 번째, 저널의 프리프린트 수용 여부는 SHERPA/RoMEO에서 확인 가능하며 특정 저널은 투고와 동시에 프리프린트 서버로 원고를 넘겨주기도 한다. 네 번째, 호스팅된 프리프린트는 DOI가 부여되며 인용 가능하다. 그러나 삭제할 수 없으며 철회할 경우 사유가 제시되어야 한다. 마지막으로 medRxiv는 프리프린트가 저널에 공식 출판된 후 2주 이내 출판된 버전에 대한 링크를 자동 추가한다. 그리고 교신저자가 이메일로 확인하는 과정을 거친다.

한편 연구자들이 프리프린트를 기탁하는 동기는 다음과 같이 정리할 수 있다(Fraser et al., 2020). 첫 번째, 시간이 많이 소요되는 피어리뷰 프로세스를 우회하여 논문이 더 일찍 학계에 공개됨으로써 과학적 발견에 대한 잠재적 우선권 확보가 가능하다. 두 번째, 연구 결과에 대한 오픈 액세스가 가능해 이용자들이 더 쉽게 무료로 접근할 수 있다. 세 번째, 연구 결과의 조기 발표로 인해 피인용 횟수 및 알트메트릭스 지수를 증가시킬 수 있다. 네 번째, 과학 커뮤니티로부터 토론, 논평과 같은 즉각적인

피드백을 받아 논문을 발전적으로 개선하는 기회를 가질 수 있다.

그러나 검증되지 않은 프리프린트는 무결성과 신뢰성을 담보하지 못한 채 무분별하게 유통되면서 다음과 같은 부작용을 야기하기도 한다. 첫 번째, 피어리뷰를 거치지 않았기 때문에 부적절한 품질이나 연구 결과 조작 여부를 확인할 수 없다. 그렇지만 이용자는 프리프린트를 검증이 완료된 피어리뷰 논문과 혼동할 수 있으며 이는 과학의 건전성과 성실성을 해칠 수 있다. 두 번째, 아이디어 도용 및 스쿠핑, 콘텐츠의 우선 순위에 대한 분쟁이 생길 수 있다. 저자가 구체적 증거 없이 프리프린트를 기탁한 후 아이디어의 우선 순위를 주장하면서 동일한 아이디어나 콘텐츠를 가진 논문의 출판을 방해할 수 있다(Yi & Hur, 2021). 세 번째, 언론에 보도되거나 소셜미디어에 언급되면서 부정확한 연구 결과가 사회적으로 확산될 수 있다. 마지막으로 동일한 과학적 산출물에 대해 두 개의 서로 다른 고유 식별자가 생성되기 때문에 인용이 분산될 수 있다. 이로 인해 연구 성과 평가 과정에서 혼란이 야기될 수 있다(Chaleplioglou & Koulouris, 2021).

그러나 이러한 문제점이 존재함에도 불구하고 최근 COVID-19의 영향으로 프리프린트의 급성장이 목도되고 있다. 팬데믹 상황 속에서 질병의 임상적 측면, 전염 방식, 역학 진행 상황 등 관련 지식의 빠른 보급과 전 세계적 예방 조치 수립, 백신 개발, 치료 방식 공유에 지체할 시간이 없었기 때문이다. 따라서 질병 발생 후 첫 4개월 이내에 40%의 연구 결과가 프리프린트 서버에서 호스팅되었고, 다른 분야보다 더 빨리 동료심사 저널에 공식 출판되어(Bertin &

Atanassova, 2022; Kodvanj et al., 2022), 학술 커뮤니케이션에 급격한 변화가 초래되었다고 논평되고 있다(Fraser et al., 2020).

3. 선행연구

프리프린트의 피인용을 다룬 연구는 공식 출판본을 대상으로 하거나 프리프린트 자체를 대상으로 하는 연구로 구분해 볼 수 있다. 전자는 특정 학술지에 출판된 논문을 대상으로 프리프린트로 기탁되었던 논문과 그렇지 않은 논문을 비교하는 방식이며, 후자는 심사 지연이나 거절 등의 사유로 미검증 상태로 남아 있는 프리프린트와 공식 출판된 프리프린트를 구분하여 분석하는 방식이다. 본 연구는 후자에 가깝지만 대다수의 선행연구는 전자의 방식으로 수행되었다.

첫 번째로 전자와 같이 공식 출판본을 대상으로 프리프린트로 기탁된 경우와 그렇지 않은 경우의 피인용을 비교 분석한 연구를 살펴보도록 한다. Davis & Fromerth(2007)는 1997년부터 2005년까지 4개의 수학 저널에 출판된 2,765개의 논문을 arXiv에 기탁되었던 경우와 그렇지 않은 경우로 구분하여 분석하였다. 그 결과 arXiv에 기탁된 논문은 기탁되지 않은 논문보다 평균적으로 35% 더 많이 인용되었으며, 이러한 차이는 극상위 인용 논문에서 더욱 두드러졌다고 하였다. Wang et al.(2020)도 문헌정보학의 3개 주요 저널을 대상으로 508개의 arXiv 기탁 논문과 5,536개의 비기탁 논문의 피인용을 비교하였다. 그 결과 arXiv 기탁 후 공식 출판된 논문에서 상당한 수준의 인용 우위를 발견했다고

밝혔다. 또한 arXiv 기탁 후 출판된 논문은 알트메트릭스 특히 Mendeley와 Twitter를 통해 더 많이 주목되었다고 언급하였다. 또한 Fu와 Hughey(2019)는 bioRxiv 기탁 논문이 비기탁 논문보다 평균적으로 36% 더 많은 피인용과 49% 더 높은 알트메트릭스를 보였다고 밝혔으며, Serghiou와 Ioannidis(2018)도 유사한 결과를 보고한 바 있다. 이렇게 프리프린트 기탁 논문이 그렇지 않은 논문보다 더 많은 주목을 받을 수 있는 이유는 조기 공개로 인해 이용자가 더 빨리 사용할 수 있으며, 오픈 액세스로 무료로 이용할 수 있기 때문이다(Brody et al., 2004; Lawrence, 2001; Moed, 2007). 더우기 저자는 본인의 논문 중 더 우수하다고 생각되는 논문 즉, 더 많은 피인용이 가능하다고 판단되는 논문을 기탁하는 성향이 있어, 프리프린트로 기탁된 논문의 공식 출판본에서 더 많은 피인용이 나타난다고 설명되고 있다(Kurtz et al., 2005).

두 번째로 본 연구와 같이 프리프린트 자체를 대상으로 분석을 수행한 연구를 살펴보도록 한다. Xu et al.(2021)은 프리프린트를 공식 출판된 경우와 그렇지 않은 경우로 구분하여 비교 분석한 결과, 공식 출판본이 존재하는 프리프린트는 그렇지 않은 경우보다 더 높은 알트메트릭스를 보였으며, 알트메트릭스가 높을수록 공식 출판되었을 때 피인용 지수가 높아진다고 설명하였다. 또한 Lariviere et al.(2014)은 arXiv에 공개된 프리프린트 버전은 초기부터 신속하게 인용되지만 시간이 지나면서 빠르게 감소되는 특징이 있다고 하였다. Henneken et al.(2007)도 프리프린트가 유일본으로 존재하는 경우에는 프리프린트 자체에 이용이 집중되지만, 출판이 이루어지면 프리프린트에 대한

이용도는 떨어지고 공식 출판본으로 대체된다고 보고하였다.

마지막으로 프리프린트의 품질과 인용의 분산과 같은 문제점에 대하여 탐구한 연구를 살펴보도록 한다. Gehanno et al.(2022)은 COVID-19 관련 프리프린트 중 절반 가량이 공식 출판이 이루어지기 전에 인용되었으며, 상당수가 공식 출판본의 제목, 데이터, 결론에 있어 차이가 존재한다고 분석하였다. 따라서 저자와 편집자는 프리프린트를 인용한 논문을 출판하기 전에 반드시 인용문과 참고문헌의 정확성을 확인해야 한다고 지적하였다. 또한 Ravinetto et al.(2021)은 아직까지 많은 이용자가 여전히 피어리뷰 논문과 프리프린트를 혼동할 수 있다고 지적하면서 다음과 같은 권장 사항을 제시하였다. ‘미검토 원고’, ‘심사를 기다리는 원고’ 또는 ‘비검토 원고’와 같이 프리프린트보다 더 명확한 용어에 대한 합의를 찾아야 하며, 눈에 잘 띄는 경고 메시지와 ‘주의-피어 리뷰 아님’이라는 빨간색 워터마크가 각 페이지에 포함되어야 한다고 하였다. 더불어 원칙과 정책을 공식화하기 위한 협의체의 설립, 모범 사례를 준수하는 서버 인증을 위한 국제 이니셔티브의 필요성에 대해서도 논의하였다. 한편 프리프린트가 학술 커뮤니케이션에서 점점 더 중요한 역할을 하고 있지만 프리프린트 인용에는 저자의 소속 기관과 국가에 대한 편향이 존재한다는 지적도 존재한다(Nishioka et al., 2022). 그들은 피인용된 프리프린트 저자의 소속기관과 국가의 집중도를 지니계수(Gini coefficient)를 사용하여 조사하였다. 그 결과 공식 출판본보다 프리프린트 인용에서 더 높은 지니계수가 나타난다고 보고했는데, 이는 내용의 신뢰성을 판단할 수

없는 상황에서 저자의 소속 국가와 기관이 후광효과로 작용하기 때문이라고 설명하였다. 또한 Chaleplioglou와 Koulouris(2021)는 프리프린트와 공식 출판본에 두개의 식별자가 생성되어 인용이 분산되는 문제를 지적하였다. 그들은 동일한 과학적 산출물에 대해 두 개의 서로 다른 고유 식별자가 공존하는 것은 학술 성과 평가에 혼란을 초래할 수 있다고 논평하였다(Ferrer-Sapena, 2018).

4. 데이터 수집 및 연구 방법

4.1 데이터 수집 방법

본 연구의 데이터 수집 방법을 기술하면 다음과 같다.

첫 번째, 본 연구에서는 21년도를 기준으로 medRxiv에 프리프린트로 최초 아카이빙된 논문 288건과 해당 논문의 출판 후 버전 172건, 총 460건을 수집하였다. <표 1>을 보면 수집된 데이터는 'A 프리프린트'와 'B 공식출판본'으로 구분되며, 'A 프리프린트' 상태의 논문은 다시 'A-a 검증된프리프린트'와 'A-b 미검증프리프린트'로 구분된다. A-a는 동료심사가 이루어져 공식 출판본이 존재하는 프리프린트이며, A-b는 공식 출판본이 존재하지 않는 미검증프리프린트이다. 'A-a 검증된프리프린트'의 출판후 버전이 B에 해당되므로 이 둘은 172건의 동수이며, A-b는 전체 프리프린트 수집량 288건에서 172건을 제외한 116건에 해당된다.

두 번째, 데이터의 수집 기간은 2023년 8월에서 10월 사이이다. 데이터 수집 시점에서 2년

전인 2021년도에 최초 아카이빙된 프리프린트를 분석 대상으로 선택한 이유는 다음과 같다. 공식 출판이 이루어진 프리프린트의 평균 출판 소요 기간을 대략 1년으로 보면(Sebo et al., 2019; Sommer et al., 2023), 2021년 프리프린트된 논문은 22년도에 출판될 가능성이 크고, 공식 출판본은 데이터 수집 시점인 23년 8-10월까지 피인용 누적 시간을 가질 수 있기 때문이다. 공식 출판본에 피인용이 누적되어야 영향력을 분석할 수 있기 때문에, 수집 시점에서 2년 전에 최초 아카이빙된 프리프린트가 분석 대상으로 결정되었다. <표 2> A-a를 보면 샘플링된 데이터의 공식 출판 시점이 표기되어 있다. 프리프린트 호스팅 당해연도인 21년에 출판된 경우가 43.6%, 차년인 22년에 출판된 경우가 51.2%로 조사되었다. 또한 <표 3>을 보면, 6건의 오류 데이터를 제외한 166건의 공식 출판 소요 기간이 집계되어 있는데, 분석 대상 논문은 비교적 빠르게 출판 프로세스가 이루어져, 평균 7.7개월, 최대 27개월이 소요되었음을 알 수 있다. 팬데믹 상황에서 COVID-19를 다룬 논문은 다른 논문보다 더 신속하게 리뷰되어 빠르게 출판되었는데(Putman et al., 2020), 이러한 영향으로 평균 출판 소요 기간이 일반적인 경우보다 다소 짧게 집계된 것으로 해석된다.

세 번째, 분석 대상이 된 'A 프리프린트' 논문 288건에서 추출된 데이터는 제목, DOI, 최초업로드시점, 누적 PDF view, 초기 6개월 누적 PDF view, 공식 출판본 존재 여부이다. 이들 요소는 모두 medRxiv에서 제공하는 메트릭스에서 추출하였다.

네 번째, 공식 출판이 완료된 'B 공식출판본'에 접근할 수 있는 링크는 'A-a 검증된프리프

〈표 1〉 수집된 데이터

A 프리프린트		B 공식출판본
A-a 검증된프리프린트 (공식출판본 존재)	A-b 미검증프리프린트 (공식출판본 미존재)	
제목, DOI, 최초개방사이트, 누적이용횟수, 초기 6개월 PDF view 횟수, 출판후버전 존재여부		DOI, 게재저널, 출판년월, AAS, 피인용(Dimension) 횟수
172건	116건	172건
288건		
-	소셜미디어, 언론보도, 정책문서 언급횟수 피인용저널 및 SJR	-

〈표 2〉 A-a, A-b 프리프린트 구성

구분	공식출판년도	빈도		비중(%)
A-a	2021	75	166	43.6
	2022	88		51.2
	2023	3		1.7
	오류	6		3.5
	전체	172		100.0
A-b		116		-
전체		288		-

〈표 3〉 A-a 데이터셋의 공식 출판 소요 개월

구분		N	최소값	최대값	평균
A-a	1년 이내	144	0	12	6.58
	1년 이상	22	13	27	15.09
	계	166	0	27	7.7

* 수집 데이터 172건에서 오류데이터 6건 제외

린트' 메타데이터에 생성되어 있다. 본 연구에
서는 링크가 생성된 논문 172건의 출판사 페이지를
방문해, DOI, 게재저널, 출판년월 정보를
수집하였다. 더불어 23년 10월 시점을 기준으로
Dimensions(dimensions.ai)가 제공하는 피
인용 데이터를 추가로 수집하였다. 따라서 〈표
1〉과 같이 'A-a 검증된프리프린트'의 지표와
'B 공식출판본'의 지표가 모두 수집된 논문은
172건(오류 데이터 제외 166건)이며, 나머지

'A-b 미검증프리프린트' 116건은 프리프린트
자체의 지표만이 수집되었다.

마지막으로 'A-b 미검증프리프린트' 중 1회
이상 피인용되고 2회 이상 AAS(Altmetric
Attention Score)가 누적된 77건을 추출하여
소셜미디어, 언론보도, 정책문서 언급 횟수를
추적하였다. 이는 Altmetric(2023)사의 연구
자 지원 프로그램을 통하여 접근 권한이 부여
된 Altmetric Explorer를 활용해 추출한 데이

터를 활용하였다. 그리고 피인용 정보는 5회 이상 피인용된 16개의 프리프린트를 추려 이 논문을 인용한 저널을 조사하였다. 분석 대상 프리프린트를 인용한 저널은 역시 Dimensions를 활용해 조사하였으며, 빈도수를 기반으로 상위 저널의 타이틀과 저널의 SJR(SCImago Journal Rank)도 함께 수집하였다.

4.2 데이터 분석 방법

이렇게 수집된 데이터는 공식 출판본의 영향력을 예측할 수 있는지를 검증하는 Analysis 1과 미검증프리프린트의 사회적 영향력과 피인용을 확인하는 Analysis 2를 통해 다음과 같이 분석되었다.

4.2.1 [Analysis 1] 프리프린트 지표의 공식 출판본 피인용 예측 가능성

Analysis 1에서는 프리프린트로 발표된 연구 결과에 집중한 관심이 출판 후 나타나는 피인용과 관련성이 있는지 분석하였다. 프리프린트의 PDF view 횟수를 독립변수, 공식 출판본의 피인용 횟수를 종속변수로 설정하여 회귀분석을 실시해, 연구결과에 대한 프리프린트의 관심이 추후 나타난 피인용을 얼마나 잘 예측할 수 있는지 확인하였다. 이 분석은 <표 4>와 같이 'A-a 검증된프리프린트' 중 2021년도와 2022년도에 공식 출판이 이루어진 경우를 대상으로 하였다. 또한 독립변수는 프리프린트의 누적 지표와 초기 6개월 지표로 구분하여 분석을 수행하였다. 누적 지표는 프리프린트가 공개된 이후 수집 시점까지 누적된 PDF view 횟수를 의미한다. 그리고 초기 6개월 누적 지표는 프리프린

트가 공개된 이후 매월 이용된 횟수가 제시되는 히스토리 지표를 통해 초기 6개월까지 누적된 횟수를 집계해 활용하였다.

첫 번째, 누적 지표를 활용한 분석에서는 빠르게 심사가 완료되어 1년 이내 신속하게 공식 출판된 경우(144건)와, 1년 이상 소요된 경우(22건)로 구분하여 분석을 실시하였다. 프리프린트는 이용 지표가 누적되는 동안 빠르게 피어리뷰 된 후 출판되어 버리는 경우도 생기며, 반대로 출판 거절이 반복되거나 심사 지연으로 공식 출판이 늦어지는 경우도 생긴다. 따라서 출판에 소요되는 시간이 짧다면 프리프린트 보다 공식 출판본을 인용할 가능성이 높아지며, 반대로 소요 기간이 길다면 공식 출판본이 부재한 상황에서 유일본인 프리프린트를 인용할 가능성이 높아질 수 있다(Henneken et al., 2007). 따라서 이 두 가지 케이스를 구분하여 분석을 수행하였다.

두 번째, 초기 6개월 지표를 활용한 분석에서는 최초 아카이빙 후 공식 출판에 소요된 시간이 6개월 이상인 논문(112건)만을 선별하여 분석을 수행하였다. 6개월을 기준으로 한 이유는 일반적으로 논문이 심사되어 출판에 이르는 시간을 1년 정도로 볼 때(Sebo et al., 2019; Sommer et al., 2023), 이 시점은 대부분의 연구가 피어리뷰와 출판 프로세스 과정 중에 해당되기 때문이다. 동시에 프리프린트가 공개된 이후 초기의 관심이 이용 통계로 누적될 수 있는 기간이기도 하다. 물론 빠른 프로세스를 통해 6개월 이내에 피어리뷰와 공식 출판이 종료된 경우도 존재할 수 있는데, 언급한 바와 같이 본 분석에서는 이 경우에 해당하는 54건은 제외하고 분석을 수행하였다.

〈표 4〉 Analysis 1의 변수

독립변수				종속변수	
A-a 검증된 프리프린트	누적 지표	출판소요 1년 미만	PDF view 횟수	B 출판후버전	Dimension 피인용수
		출판소요 1년 이상			
	초기 6개월 누적 지표	PDF View 횟수			

4.2.2 [Analysis 2]: 미검증 프리프린트의 사회적 영향력과 피인용 현황

한편 Analysis 2에서는 ‘A-a 미검증프리프린트’ 중에서 피인용되었거나 사회적으로 언급된 프리프린트를 대상으로 분석을 수행하였다. 여기에서는 출판되지 못한 ‘A-b 미검증프리프린트’ 중 피인용이 1회 이상이거나 AAS가 2이상을 보이는 77건이 선별되었다.

먼저 Altmetric Explorer를 활용해 언론보도, 소셜미디어, 정책문서 등에서 어떻게 언급되고 있는지 살펴보았으며, 5회 이상 피인용된 프리프린트 16건을 추린 후 Dimensions를 활용해 해당 프리프린트가 어떠한 학술지에서 피인용되고 있는지 추적하였다. 또한 이들 미검증 프리프린트를 다수 인용한 상위 5종의 저널을 각각 추출하여 영향력 지수(SJR)를 확인하였다.

5. 분석결과

5.1 [Analysis 1] 프리프린트 지표의 공식 출판본 피인용 예측 가능성

본 장에서는 A-a 샘플 172건 중 오류 데이터를 제외한 166건을 대상으로 프리프린트 지표로 공식 출판본의 피인용을 예측할 수 있는지

확인해 보도록 한다. 앞에서 언급한 공개 후 수집 시점까지의 누적 지표와 초기 6개월 누적 지표를 활용하여 출판 후 버전의 피인용 예측 가능성을 분석하고 무엇이 더 적절한 지표로 활용 가능한지 논의해 본다.

5.1.1 공개 후 수집 시점까지 누적된 프리프린트 지표의 피인용 예측 가능성

먼저 프리프린트가 공개된 후 수집 시점까지 누적된 지표의 피인용 예측 가능성을 살펴보도록 한다. 본 연구에서 샘플로 사용한 논문은 〈표 2〉와 같이 1년내 심사가 완료되어 비교적 빨리 공식 출판본이 나온 경우와 1년 이상 소요된 경우로 구분된다.

앞에서 언급한 바와 같이, 출판에 소요되는 기간이 길어지면 공식 출판본이 없는 프리프린트 상태가 길어진다. 이럴 경우 이용자는 공식 출판본을 기다리지 않고 프리프린트를 더 적극적으로 이용하게 된다. 반대로 출판 후 소요 기간이 짧으면 공식 출판본이 없는 프리프린트 상태가 짧아진다. 이럴 경우 프리프린트 이용 빈도는 줄어들고 공식 출판 버전의 이용이 촉진될 수 있다(Henneken et al., 2007). 따라서 두 가지 경우에 있어 피인용에 미치는 영향에 차이가 발생할 수 있다는 가정 하에, 다음과 같이 각 케이스를 구분하여 회귀분석을 수행해 보았다. 구체적으로 프리프린트의 열람 정도를

의미하는 PDF view 횟수가 출판 후 버전의 피인용 횟수에 영향을 미치는지 분석하였다.

그 결과 <표 5>와 같이 1년내 공식 출판 버전이 나온 경우에는 $\beta = 0.171(p < 0.10)$ 로 유의한 영향력이 존재하는 것으로 확인되었다. 즉 1년 내 공식 출판본이 나온 경우 프리프린트에 주목된 관심은 공식 출판된 후에 나타날 수 있는 피인용 횟수와 약한 관련성이 있는 것으로 해석할 수 있겠다. 그러나 두 번째 경우, 즉 출판 소요 시간이 긴 경우는 <표 6>과 같이 $\beta = -0.013(p > 0.05)$ 의 유의하지 않은 결과가 나타났다. 그 이유는 공식 출판본이 발표되는데 소요되는 시간이 길기 때문에 그 기간 동안은 유일본으로 존재하는 프리프린트를 이용할 수 밖에 없기 때문이다. 따라서 이러한 경우 프리프린트에 대한 관심은 공식 출판 버전의 피인용으로 유도되지 않고 도리어 프리프린트의 피인용을 유도할 개연성이 있다. 참고로 이러한 케이스 즉, 공식 출판이 1년 이상 소요된 프리프린트의 관심이 프리프린트 자체의 피인용

을 유도하는지 확인해 본 결과, $\beta = 0.361(p < 0.05)$ 로 유의한 효과가 존재하는 것으로 확인되었다.

정리하자면 공식 출판본과 프리프린트의 피인용은 출판 소요 기간의 길고 짧음에 따라 다른 결과가 나타난다. 공식 출판 소요 기간이 짧은 경우 프리프린트에 집중된 관심은 공식 출판본의 피인용을 약하게 유도하지만, 출판 소요 기간이 길어지면 공식 출판본의 피인용을 유도하지 않고, 도리어 프리프린트 자체의 피인용을 유도하는 것으로 검증되었다. 따라서 프리프린트의 누적 피인용 지표만으로는 공식 출판본의 영향력을 예측하기 어려우며, 출판 소요 시간에 영향을 받지 않는 다른 지표를 고민한 필요가 있을 것이다.

5.1.2 초기 6개월 누적 지표가 피인용에 미치는 영향

본 장에서는 위에서 사용된 누적 지표 대신 히스토리 지표를 활용해 프리프린트 공개 후 초

<표 5> 출판 소요 기간이 1년 이내인 경우의 영향력

	coeff	standardized coeff	se	t	p
상수	8.949		1.650	5.422	0.000
PDF view	0.040	0.171	0.020	1.966	0.051
			R^2	F	P
			.029	3.865	0.051

<표 6> 출판 소요 기간이 1년 이상인 경우의 영향력

	coeff	standardized coeff	se	t	p
상수	2.931		1.476	1.986	0.061
PDF view	-0.002	-0.013	0.031	-0.056	0.956
			R^2	F	P
			.000	.003	0.956

기 6개월간 누적된 관심이 피인용 영향을 예측할 수 있는지 검증해 보았다. 초기 6개월은 대부분 공식 출판본이 나오기 이전 시점으로 출판 버전 존재 여부에 영향을 받지 않고, 이용자들의 관심을 끌었다면 공식 출판본에서도 높은 영향력이 나타날 수 있을 것으로 가정하였다.

분석 결과 <표 7>과 같이 유의한 양의 영향력이 나타나(PDF view $\beta = 0.409$ $p < 0.01$), 프리프린트의 지표가 공식 출판본의 지표와 긍정적 관련성을 보임이 확인되었다. 즉 프리프린트 버전의 최초 6개월 누적 PDF view 횟수는 공식 출판 버전의 피인용을 40% 가까이 예측할 수 있어, 출판 소요 기간에 영향을 받지 않고 이용자가 가치 있는 프리프린트를 선별할 수 있는 지표임이 확인되었다. 그러나 계산된 영향 계수는 절대적 관련성을 나타낸다고 보기는 어렵기 때문에 이 지표는 이용자의 선별을 도울 수 있는 보완적 도구로 활용 가능할 것이다. Nishioka et al.(2022)은 그들의 연구에서 프리프린트는 검증 절차를 거치지 않았기 때문에 이용자 스스로가 내용의 신뢰성과 정확성을 판단해야 하는데, 이 과정에서 어쩔 수 없이 저자의 소속 국가와 기관을 기준으로 판단하는 경향이 존재한다고 하였다. 초기 6개월 누적 지표는 프리프린트의 가치를 판단할 때, 이러한 후광 효과를 상쇄할 수 있는 보조적 도구로 활

용될 수 있을 것이다.

정리하자면 프리프린트 버전의 누적 이용 지표는 출판 소요 기간이 짧은 경우 공식 출판본의 피인용을 유도하지만, 출판 소요 기간이 길어질 경우 공식 출판본의 피인용을 유도하지 않고 프리프린트 버전의 피인용을 유도할 수 있다. 따라서 프리프린트를 선별하고자 하는 이용자에게 이는 적절한 지표로 활용되기 어렵다. 대신 초기 6개월 지표를 사용하면 향후 출판 후 버전에서 나타날 수 있는 영향력을 40% 선에서 예측해, 연구 결과 선별에 도움을 줄 수 있을 것으로 보인다.

5.2 (Analysis 2) 미검증 프리프린트의 사회적 영향력과 피인용 현황

두 번째 분석에서는 샘플 중 사회적으로 주목되었거나 높은 피인용을 보이지만 검증되지 않은 프리프린트가 존재하는지 확인해 보고, 그 논문은 어떠한 경로로 확산되고 있는지 살펴보도록 한다. 또한 어떠한 저널을 통해 후속 연구에 피인용되고 있는지도 분석해 보도록 한다. 여기에 사용된 샘플은 앞에서 사용된 샘플과 달리 'A-b 미검증프리프린트'이다. 21년에 공개된 프리프린트 중 수집일 시점에서 미검증 프리프린트로 남아있는 116건 중 피인용이 1회

<표 7> 초기 6개월 누적 PDF view 횟수가 공식 출판 버전의 피인용에 미치는 영향

	coeff	standardized coeff	se	t	p
상수	7.382		1.072	6.888	.000
초기 6개월 PDF view 횟수	.002	.409	.001	4.453	.000
			R ²	F	P
			.167	19.834	.000

이상이거나 AAS가 2점 이상을 보이는 논문 77건이 식별되었다. 즉 최초 아카이빙 후 약 20개월에서 32개월간 출판되지 못하고 미검증 상태로 남아 있지만 그 자체로 피인용되거나 사회적으로 주목된 논문이라고 말할 수 있겠다. <표 8>을 보면, 77건의 논문 중 8건의 논문은 10회 이상 피인용되었으며, 언론에 보도되거나 소셜미디어에 언급되면서 10점 이상의 AAS를 보인 논문도 10건이나 되는 것을 알 수 있다. 한편 제목에 포함된 키워드의 출현 빈도를 통해 대략적인 주제를 살펴보면, 77건 중 28건에서 COVID-19가 출현해, 펜데믹 이슈를 다루는 프리프린트가 높은 사회적 주목을 받은 것으로 확인되었다. 그렇다면 이러한 논문이 사회적으로 확산된 경로와 피인용된 저널을 자세히 살펴보도록 한다.

첫 번째로 알트메트릭스 분석을 통해 미검증 프리프린트가 사회적으로 어떻게 확산되고 있는지 살펴본 결과는 다음과 같다. AAS 268로 가장 많은 사회적 주목을 받은 프리프린트가 존재하

였는데, 이는 이스라엘의 COVID-19 백신 효과에 관한 연구로 소셜미디어, 뉴스, 정책문서 등에 언급되면서 확산된 것으로 확인된다. 한편 분석 대상 전체 프리프린트는 <표 9>와 같이 트윗에서는 총 939회, 뉴스와 블로그에는 각각 14회와 3회, 정책문서에서는 5회 언급되었다. 트윗은 medrxivpreprint 자체 트윗이 가장 많았지만, 그 밖에 COVID19arXiv, SciscoreReports 및 개인 연구자 등에 의해 언급되고 있었다. 뉴스의 경우 The Medical News MSN, The Conversation 등에 언급되고 있었는데, COVID-19 심각도 예측, 백신, 면역력, 임상 특성을 다룬 기사에서 언급되었다. 또한 정책문서에 언급된 경우도 존재하였다. Centers for Medicare & Medicaid Services의 정책 문서에서 백신의 효과에 대한 프리프린트가 인용되었으며, WHO의 주간 역학 업데이트에서도 변종에 대한 백신의 효과, 이스라엘, 브라질의 코로나 현황에 대한 프리프린트가 인용되었음이 확인되었다.

두 번째로 미검증 프리프린트가 어떠한 학술

<표 8> 분석 대상 미검증 프리프린트의 피인용 및 AAS 분포

피인용 분포		AAS 분포	
구간(횟수)	건	구간(횟수)	건
40~	2	200~	1
20-39	2	50-199	1
10-19	4	10-49	8
5-9	9	2-9	40
1-4	40		

<표 9> 미검증 프리프린트의 사회적 언급

소셜미디어	뉴스, 블로그	정책 특허
트윗 939회	뉴스 14회 블로그 2회	정책 문서 5회

지에서 피인용되고 있는지 살펴보도록 한다. <표 8>과 같이 피인용 횟수 5회 이상을 보인 17개의 프리프린트 중 오류 데이터 1건을 제외한 16건을 대상으로 해당 프리프린트를 인용한 상위 5종의 출처를 Dimensions를 통해 수집하였다. 예를 들어 프리프린트 A는 총 48회 피인용되었는데, 이 논문을 다수 인용한 상위 5종은 medRxiv, Vaccines, The Lancet, SSRN, bioRxiv이다. 이 중 3종은 학술지가 아니라 프리프린트 서버인데, 이는 미검증 프리프린트가 학술지를 통해 인용되기도 하지만 프리프린트 서버에 아카이빙된 또 다른 프리프린트를 통해

서도 인용되고 있음을 보여준다. 이렇게 16개의 프리프린트가 각각 몇 회씩 어떠한 출처에서 인용되었는지 상위 5종까지 정리한 결과, 중복을 제외한 총 54종이 집계되었다. 그중 집계된 값이 2회 이상인 22개의 출처를 <표 10>과 같이 제시하고 해당 출처에 게재된 논문들이 분석 대상 프리프린트를 인용한 횟수를 합산한 값도 제시하였다. 가령 합산값이 9로 나온 Vaccines의 경우, 여기에 게재된 9건의 논문이 분석 대상 프리프린트를 인용했다고 해석할 수 있겠다.

<표 10>을 보면 가장 많은 빈도를 보인 것은 medRxiv로 나타났으며, 그 밖에도 arXiv 등의

<표 10> 미검증 프리프린트를 인용한 상위 출처

인용한 출처	분석 대상 프리프린트 인용 합산 횟수	SJR
medRxiv	48	(프리프린트 서버)
Vaccines	9	1.66
EClinicalMedicine	4	4.55
JMIR Preprints	4	(프리프린트 서버)
bioRxiv	3	(프리프린트 서버)
IEEE Access	3	0.926
The Lancet	3	14.6
arXiv	2	(프리프린트 서버)
Biomedical Signal Processing and Control	2	1.07
Biomolecules	2	1.07
Computational Intelligence and Neuroscience	2	
Diagnostics	2	0.67
Eurosurveillance	2	4.14
Journal of Microbiology Immunology and Infection	2	1.45
Lecture Notes in Computer Science	2	0.32
Nature Communications	2	5.12
Open Forum Infectious Diseases	2	1.5
PLOS ONE	2	0.885
Scientific Data	2	2.41
Seminars in Roentgenology	2	0.156
Soft Computing	2	0.819
SSRN Electronic Journal	2	(프리프린트 서버)

프리프린트 서버 5종이 포함되어 있음을 확인할 수 있다. 분석 대상 프리프린트를 인용한 104건의 논문 중 59건인 57%가 프리프린트 서버를 출처로 하고 있다는 것이다. 이는 미검증 프리프린트는 또 다른 프리프린트를 통해 가장 자주 인용되고 있다고 해석할 수 있겠다. 프리프린트를 통한 연구 결과의 선 공개는 전 지구적 위기 상황에 대처하기 위한 학계의 신속한 움직임을 반영하고 있는데, 프리프린트를 인용하는 또 다른 프리프린트의 증가는 이 분야 지식의 빠른 확산 속도를 투영한다고 말할 수 있다. 한편 프리프린트 이외에도 높은 영향을 보이는 저널에 출판된 논문에서 인용된 경우도 존재하였다. The Lancet, Nature Communications, EClinicalMedicine 등 SJR 4.55에서 14.6에 이르는 높은 영향력을 가진 저널에서 미검증 프리프린트를 인용하고 있으며, 그 밖에도 다양한 영향력을 가진 저널이 포함되어 있는 것으로 확인되었다.

6. 결 론

본 연구에서는 공식 출판본의 영향력을 예측하는데 활용될 수 있는 프리프린트 지표를 선별하고 피인용되고 언론에 보도되면서 학술적, 사회적으로 주목되고 있는 미검증 프리프린트의 존재에 대해 살펴보았다. 본 연구를 통해 발견된 시사점을 다음과 같이 정리할 수 있겠다.

첫 번째, 공식 출판이 이루어지면 프리프린트의 이용도는 급감할 수 있으므로 프리프린트 서버에 누적된 이용 지표는 큰 의미가 없을 수 있다. 따라서 본 연구에서는 피인용을 40% 선

에서 예측할 수 있는 초기 6개월 누적 PDF view 횟수를 프리프린트 선별 시 활용할 수 있는 지표로 제안하였다. 최근 OPR(Open Peer Review), 프리프린트와 같이 피어리뷰 전 논문의 유통이 증가하게 되면서 대체적 영향 지표에 대한 중요성이 커지게 되었는데(조재인, 2022; 조재인, 박종도, 2023), 본 연구에서 제안된 지표 이외에도 적절한 지표에 대한 추가적인 고민이 필요할 것이다.

두 번째, 피어리뷰 중이거나 출판이 거절된 미검증 프리프린트도 다수에 의해 열람 및 피인용되고 있으며, 언론보도와 정책문서에 언급되면서 사회적 영향력을 발휘하고 있다는 것을 확인하였다. 그러나 미검증 결과의 무분별한 확산은 위험을 초래할 수 있기 때문에, 배포 시 검증되지 않았음에 대한 경고를 명확하게 할 필요가 있을 것이다. 학술지 편집위원회에서도 프리프린트 인용에 대한 명확한 가이드라인을 설정하여 공식 출판본이 존재하는 경우 프리프린트 인용을 최대한 자제할 수 있도록 유도해야 할 것이다.

세 번째, 공식 출판에 소요되는 시간적 지체를 기다릴 수 없는 팬데믹 상황에서는 프리프린트 출판량이 많아지면서 프리프린트를 인용하여 또 다른 프리프린트를 생산·공개하는 경향이 나타났다. 이는 새로운 지식의 생산과 빠른 전파 속도를 투영하고 있지만, 그 만큼 철회되는 논문의 양도 증가하는 상황(Syed et al., 2023)에서 무분별한 인용과 사회적 확산 문제에 대한 주의가 필요할 것이다. 또한 프리프린트의 증가와 또 다른 프리프린트에서의 인용 현상이 다른 분야에서도 나타나고 있는 일반적인 현상인지는 추가적인 조사가 필요할 것이다.

학술 커뮤니케이션에서 프리프린트가 차지하는 위상은 변화되고 있으나 과학적 발견의 무결성이나 신뢰성에 대한 우려가 뒤따르기도 한다. 프리프린트는 새로운 학술출판 패러다임이므로 연구자들은 장단점과 문제점을 인식하고 해결하면서 받아드릴 수 밖에 없을 것이다. 앞으로 국내외 학술 커뮤니티를 중심으로 미검

증 연구 결과 확산에 대한 경고, 프리프린트 인용에 대한 윤리적 가이드라인, 모범사례 발굴과 원칙의 공식화를 위한 논의가 이루어져야 할 것이다. 본 연구는 아직까지 국내 학계에서 적극적으로 논의되고 있지 않은 프리프린트 관련 주제를 확산시키고 그 중요성을 공유하는 첫 걸음이 될 수 있을 것이다.

참 고 문 헌

- 조재인 (2022). 학술논문 알트메트릭스의 피인용 영향과 오픈액세스의 조절효과에 관한 연구. 한국문헌정보학회지, 56(2), 35-55. <http://10.4275/KSLIS.2022.56.2.035>
- 조재인, 박종도 (2023). 오픈 피어 리뷰 환경에서 학술 논문 심사 결과와 영향력 지표 간의 관련성에 관한 연구. 한국문헌정보학회지, 57(2), 79-96. <http://10.4275/KSLIS.2023.57.2.079>
- Altmetric (2023). Altmetric database in CSV format. Unpublished raw data.
- Bertin, M. & Atanassova, I. (2022). Preprint citation practice in PLOS. *Scientometrics*, 127, 6896-6912. <https://doi.org/10.1007/s11192-022-04388-5>
- Brody, T., Stamerjohanns, H., Harnad, S., Gingras, Y., Vallieres, F., & Oppenheim, C. (2004). The effect of open access on citation impact. National Policies on Open Access (OA) Provision for University Research Output: an International meeting, Southampton, Southampton University. Available: <http://opcit.eprints.org/feb19oa/brody-impact.pdf>
- Chalepioglou, A. & Koulouris, A. (2021). Preprint paper platforms in the academic scholarly communication environment. *Journal of Librarianship and Information Science*, 55(1), 43-56. <https://doi.org/10.1177/09610006211058908>
- Davis, P. M. & Fromerth, M. J. (2007). Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles? *Scientometrics*, 71(2), 203-215. <https://doi.org/10.1007/s11192-007-1661-8>
- Ferrer-Sapena, A., Aleixandre-Benavent, R., Peset, F., & Sánchez-Pérez, E. A. (2018). Citations to arXiv preprints by indexed journals and their impact on research evaluation. *Journal of Information Science Theory and Practice*, 6(4), 6-16. <https://doi.org/10.1633/JISTaP.2018.6.4.1>

- Fraser, N., Brierley, L., Dey, G., Polka, JK., 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 Biol*, 19(4), e3000959.
<https://doi.org/10.1371/journal.pbio.3000959>
- Fraser, N., Momeni, F., Mayr, P., & Peters, I. (2020). The relationship between bioRxiv preprints, citations and altmetrics. *Quantitative Science Studies*, 1(2), 618-638.
https://doi.org/10.1162/qss_a_00043
- Fu, Y. D. & Hughey, J. J. (2019). Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *eLife*, 8, e52646. <https://doi.org/10.7554/eLife.52646>
- Gehanno J-F., Grosjean J., Darmoni SJ., & Rollin L. (2022). Reliability of citations of medRxiv preprints in articles published on COVID-19 in the world leading medical journals. *PLoS ONE*, 17(8), e0264661. <https://doi.org/10.1371/journal.pone.0264661>
- Henneken, E. A., Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Thompson, D., Bohlen, E., Murray, S., Ginsparg, P., & Warner, S. (2007). E-prints and journal articles in astronomy: a productive co-existence. *Learned Publishing*, 20, 16-22.
<https://doi.org/10.1087/095315107779490661>
- Kodvanj, I., Homolak, J., Virag, D., & Trkulja, V. (2022). Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues. *Scientometrics*, 127, 1339-1352. <https://doi-org-ssl.access.inu.ac.kr/10.1007/s11192-021-04249-7>
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. *Information Processing & Management*, 41(6), 1395-1402. <https://doi.org/10.1016/j.ipm.2005.03.010>
- Lariviere, V., Sugimoto, C., Macaluso, B., Milojević, S., Cronin, B., & Thelwall, M. (2014). ArXiv e-prints and the journal of record: an analysis of roles and relationships. *Journal of the Association for Information Science and Technology*, 65(6), 1157-1169.
<https://doi.org/10.1002/asi.23044>
- Lawrence, S. (2001). Online or invisible? *Nature*, 411, 521-523.
- Moed, H. F. (2007). The effect of “open access” on citation impact: an analysis of ArXiv’s condensed matter section. *Journal of the American Society for Information Science and Technology*, 58(13), 2047-2054. <https://doi.org/10.1002/asi.20663>
- Nishioka, C., Farber, M., & Saier, T. (2022). How does author affiliation affect preprint citation count? analyzing citation bias at the institution and country level. In *Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries*, Article 28.

- <https://doi.org/10.1145/3529372.3530953>
- Putman M. S., Ruderman E. M., & Niforatos J. D. (2020). Publication rate and journal review time of COVID-19 related research. *Mayo Clinic Proceedings*, 95(10), 2290-2291.
<http://doi.org/10.1016/j.mayocp.2020.08.017>
- Ravinetto, R., Caillet, C., Zaman, M. H., Singh, J. A., Guerin, P. J., Ahmad, A., Durán, C., Jesani, A., Palmero, A., Merson, L., Horby, P., Bottieau, E., Hoffmann, T., & Newton, P. (2021). Preprints in times of Covid19: the time is ripe for agreeing on terminology and good practices. *BMC Medical Ethics*, 22(1), 1-5. <https://doi.org/10.1186/s12910-021-00667-7>.
- Schonfeld, R. & Rieger, O. (2020) Publishers invest in preprints. The Scholarly Kitchen. Available: <https://scholarlykitchen.sspnet.org/2020/05/27/publishers-invest-in-preprints/>
- Sebo, P., Fournier J., Ragot, C., Gorioux, P., Herrmann, F., & Maisonneuve, H. (2019). Factors associated with publication speed in general medical journals: a retrospective study of bibliometric data. *Scientometrics*, 119, 1037-1058.
<https://doi.org/10.1007/s11192-019-03061-8>
- Serghiou, S. & Ioannidis, J. P. A. (2018). Altmetric scores, citations, and publication of studies posted as preprints. *Journal of the American Medical Association*, 319(4), 402-404.
<https://doi.org/10.1001/jama.2017.21168>
- Smart, P. (2022). The evolution, benefits, and challenges of preprints and their interaction with journals. *Science Editing*, 9(1), 79-84. <https://doi.org/110.6087/KCSE.269>
- Sommer, I., Sunder-Plassmann, V., Ratajczak, P., Emprechtinger, R., Dobrescu, A., Griebler, U., & Gartlehner, G. (2023). Full publication of preprint articles in prevention research: an analysis of publication proportions and results consistency. *Scientific Reports* 13, 17034. <https://doi.org/10.1038/s41598-023-44291-4>
- Syed, Z., Syed, F., Thabane, L., & Rodrigues, M. (2023), COVID-19 retracted publications on retraction watch: a systematic survey of their pre-prints and citations. *Heliyon*, 9(4), e15184. <https://doi-org-ssl.access.inu.ac.kr/10.1016/j.heliyon.2023.e15184>
- Wang, Z., Glänzel, W., & Chen, Y. (2020). The impact of preprints in Library and Information Science: an analysis of citations, usage and social attention indicators. *Scientometrics*, 125, 1403-1423. <https://doi.org/10.1007/s11192-020-03612-4>
- Xu, F., Ou, G., Ma, T., & Wang, X. (2021). The consistency of impact of preprints and their journal publications. *Journal of Informetrics*, 15(2), 101153,
<https://doi.org/10.1016/j.joi.2021.101153>
- Yi, H. & Huh, S. (2021). Korean editors' and researchers' experiences with preprints and attitudes

towards preprint policies. Science Editing, 8(1), 4-9. <https://doi.org/10.6087/kcse.223>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

Cho, Jane (2022). A study on the effect of altmetrics about academic papers on citations and moderating effect of open access. Journal of the Korean Society for Library and Information Science, 56(2), 35-55. <http://10.4275/KSLIS.2022.56.2.035>

Cho, Jane & Park, Jong-do (2023). A study on the relationship between the review results of articles and impact metrics in an open peer review platform. Journal of the Korean Society for Library and Information Science, 57(2), 79-96. <http://10.4275/KSLIS.2023.57.2.079>