

Can Twitter data help in spotting problems early with publications? What retracted COVID-19 papers can teach us about science in the public sphere

*The COVID-19 pandemic has brought science into mainstream public and political debates in novel ways, notably through the widespread use of social media to share and discuss new findings. In this post, **Robin Haunschild** and **Lutz Bornmann** discuss their recent findings on how retracted papers were talked about on the social media platform Twitter and how this can be mapped onto the eventual retraction notices of these articles. They suggest that exploring how Twitter and other forms of open post-publication peer review highlight potential errors in the scientific record might serve as an early warning mechanism for spotting research errors.*

Publications that are based on wrong data, methodological mistakes, or contain other types of severe errors can spoil the scientific record if they are not retracted. Retraction of publications is one of the effective ways to correct the scientific record. However, before a problematic publication can be retracted, the problem has to be found and brought to the attention of the people involved (the authors of the publication and editors of the journal). The earlier a problem with a published paper is detected, the earlier the publication can be retracted and the less wasted effort goes into new research that is based on disinformation within the scientific record. Therefore, it would be advantageous to have an early warning system that spots potential problems with published papers, or maybe even before based on a preprint version.

One candidate for such a system is Twitter. In [a recent case study](#), we tested the suitability of Twitter data for being a part in an early warning system to spot problems with publications. We analysed the meta data of tweets that mentioned three retracted publications related to SARS-CoV-2/COVID-19 and their retraction notices. We selected these three publications from the Retraction Watch post about retracted SARS-CoV-2/COVID-19 papers (<https://retractionwatch.com/retracted-coronavirus-covid-19-papers/>) because they all have a DOI for the publication and the retraction notice and the dates of publication and retraction are at least two weeks apart. We looked at the following studies:



Date originally posted: 2021-08-09

Permalink: <https://blogs.lse.ac.uk/impactofsocialsciences/2021/08/09/can-twitter-data-help-in-spotting-problems-early-with-publications-what-retracted-covid-19-papers-can-teach-us-about-science-in-the-public-sphere/>

Blog homepage: <https://blogs.lse.ac.uk/impactofsocialsciences/>

[Bae, et al. \(2020a\)](#) studied the effectiveness of surgical and cotton masks in blocking the virus SARS–CoV-2. The study was published on 6 April 2020 and retracted on 2 June 2020 because they “had not fully recognized the concept of limit of detection (LOD) of the in-house reverse transcriptase polymerase chain reaction used in the study” (Bae, et al., 2020b). In this case, the [retraction](#) was made because of a methodological error that was not detected in the peer-review process.

[Wang, et al. \(2020a\)](#) reported that “SARS-CoV-2 infects T lymphocytes through its spike protein-mediated membrane fusion”. The peer review process was very fast: The paper has been submitted on 21 March 2020 and accepted three days later on 24 March 2020. This paper has been published on 7 April 2020, and retracted on 10 July 2020 (Wang, et al., 2020a) because “[a]fter the publication of this article, it came to the authors attention that in order to support the conclusions of the study, the authors should have used primary T cells instead of T-cell lines. In addition, there are concerns that the flow cytometry methodology applied here was flawed. These points resulted in the conclusions being considered invalid.” In this case, the [retraction](#) was made because of methodological errors that were not discovered during the peer-review process.

Probably the most attention among the three publications was drawn to the study by [Mehra, Desai, Ruschitzka, and Patel \(2020\)](#). They reported that they could not confirm a benefit in COVID-19 treatment with hydroxychloroquine. They even reported that hydroxychloroquine increases the risk of complications during medical treatment against COVID-19. The study was published on 22 May 2020, and [retracted](#) on 05 June 2020, because “several concerns were raised with respect to the veracity of the data and analyses conducted by Surgisphere Corporation and its founder” and co-author of the study. Surgisphere declined to transfer the full dataset to an independent third-party peer reviewer because that would violate client agreements and confidentiality requirements. Potential benefit or risk of hydroxychloroquine for treatment of COVID-19 is still not clear. In this case, the retraction was made because of doubts regarding the validity of the employed data that was not discovered in the peer-review process.

We downloaded the meta data of the tweets in August 2020. All three retracted publications received rather high numbers of tweets (between 3,095 and 42,746). An analysis of the profile descriptions of the Twitter users shows that most tweets originated from personal accounts, faculty members and students, or professionals. Tweets from bots represented a small minority. Therefore, we can expect informative content from the tweet texts.

Our results indicate that some studies are indeed robustly discussed by experts on Twitter.

We produced word clouds (see for example the word cloud regarding [Wang, et al. \(2020a\)](#) before its retraction) for each retracted publication and their retraction notices before and after retraction. We also performed searches in the tweet texts for phrases related to the retraction reasons (e.g., ‘LOD’, ‘limit of detection’, ‘flow cytometry’, ‘Surgisphere’, and ‘data’). Inspection of the word clouds did not provide additional phrases than those taken from the retraction notices. A manual inspection of the tweet texts that contain these phrases revealed that these tweets indeed mentioned the problems of the retracted papers before the retraction date in the case of two of the three publications in our case study. Our results indicate that some studies are indeed robustly discussed by experts on Twitter.

