

Lemke, Steffen; Brede, Max; Rotgeri, Sophie; Peters, Isabella

**Article — Published Version**

## Research articles promoted in embargo e-mails receive higher citations and altmetrics

Scientometrics

*Suggested Citation:* Lemke, Steffen; Brede, Max; Rotgeri, Sophie; Peters, Isabella (2022) : Research articles promoted in embargo e-mails receive higher citations and altmetrics, *Scientometrics*, ISSN 1588-2861, Springer, Berlin, Vol. 127, pp. 75–97, <https://doi.org/10.1007/s11192-021-04217-1>

This Version is available at:

<http://hdl.handle.net/11108/507>

**Kontakt/Contact**

ZBW – Leibniz-Informationzentrum Wirtschaft/Leibniz Information Centre for Economics  
Düsternbrooker Weg 120  
24105 Kiel (Germany)  
E-Mail: [info@zbw.eu](mailto:info@zbw.eu)  
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

**Standard-Nutzungsbedingungen:**

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.*



<http://creativecommons.org/licenses/by/4.0/>



# Research articles promoted in embargo e-mails receive higher citations and altmetrics

Steffen Lemke<sup>1</sup> · Max Brede<sup>2</sup> · Sophie Rotgeri<sup>3</sup> · Isabella Peters<sup>1,2</sup>

Received: 15 June 2020 / Accepted: 19 November 2021 / Published online: 18 December 2021  
© The Author(s) 2021

## Abstract

In order to be able to provide thorough and timely coverage on the most recent scientific research, science journalists frequently rely on embargoed information sent to them by publishers of scientific journals. In such embargo e-mails, publishers purposefully bring selected upcoming releases to the journalists' attention a few days in advance of their publication. Little is known on how this early highlighting of certain research articles affects their later citations or altmetrics. We present an exploratory case study with the aim of assessing the effects of such promotion activities on scientific articles' bibliometric and altmetric indicators. In a treatment–control design, we analyze citation counts and eight types of altmetrics of 715 articles published between 2016 and 2017 whose DOIs have been mentioned in embargo e-mails and compare these to articles from the same journal issues that have not been highlighted in embargo e-mails. Descriptive statistics and Mann–Whitney-U tests reveal significant advantages for promoted articles across all regarded metrics three to four years after their publication. Particularly large differences can be seen regarding numbers of mentions in mainstream media, in blogs, on Twitter, and on Facebook. Our findings suggest that scholarly publishers exert significant influence over which research articles will receive attention and visibility in various (social) media. Also, regarding utilizations of metrics for evaluative purposes, the observed effects of promotional activities on indicators might constitute a factor of undesirable influence that currently does not receive the amount of consideration in scientometric assessments that it should receive.

**Keywords** Embargo e-mails · Science journalism · External science communication · Scholarly communication · Altmetrics · Citations

## Introduction

Staying informed about new developments in their field is an integral part of most researchers' everyday work. For them, the most common media from which to learn about new research findings will usually be academic journals, books, or conferences, depending on the field at hand. However, just like the general public, researchers also consume mass

---

✉ Steffen Lemke  
s.lemke@zbw.eu

Extended author information available on the last page of the article

media, in which mediators from outside of the scientific community—typically journalists—communicate recent research findings (Kiernan, 2003a; Phillips et al., 1991).

In addition to universities and research institutions, which, amongst others, distribute new research findings via press releases, scholarly publishers occupy a key role in the dissemination of science via mass media. They regularly provide journalists with prepared summaries of selected new research articles in advance of their publication, so that the journalists have some days to prepare their coverage on said articles (Kiernan, 1997). These in advance-summaries are sent to the journalists *under embargo*, meaning their contents are only allowed to be published further after a date specified by the scholarly publisher has passed. This way, releases of research articles and their journalistic coverage are synchronized. So while journalists benefit from the embargo system by being enabled to publish well-prepared reports on recent science at the earliest possible date, publishers benefit from increased control over the timing of their articles' publicity. Moreover, through the selective provision of embargo information, publishers exercise a strong influence on *which* scientific findings can be covered timely and comprehensively in newspapers, television, radio broadcasts, and other forms of mass media (Kiernan, 1997, 2003a). Kiernan (1997) provides a historical overview over the embargo system's origins and, from reviewing past literature, concludes that it gives editors of scientific journals considerable power over what is regarded as scientific news and when mass media can report on it. Although in recent years, new platforms for the early publication of research (e.g., preprint repositories) might subvert this power by providing journalists with additional ways of obtaining early insights on recent findings, the embargo system is still in place as one of the essential channels for scholarly publishers to distribute new research to journalists (see also Franzen (2012), Kiernan (2003b)). In a survey sent to the editors of 120 medical journals, Kiernan (2014) found 67% of the respondents to report that their journal would regularly offer journalists some kind of press material on new publications under embargo. Asked for reasons that would justify the embargo of journal articles, surveyed editors most commonly stated that embargos would be helpful to ensure that media coverage coincides with respective articles' publication and that they would help to ensure high-quality press coverage.

Several past studies examined the relationship between external science communication (i.e., the communication of research via channels not primarily aimed at other researchers, e.g., in news media) and how it affects research articles' later scientific impact, i.e. in terms of the number of citations. Phillips et al. (1991) for instance examined whether articles from the *New England Journal of Medicine* that had been featured in the *New York Times* (NYT) have received higher citation numbers than similar articles that had not been featured. They found that the former group did receive significantly higher citations, particularly in the first year after publication. Interestingly, this was not the case for research articles featured in NYT stories published during a period of strike, in which respective NYT issues were not distributed. For Phillips et al. (1991) these results support the 'publicity hypothesis', which assumes that media coverage genuinely increases a scientific article's visibility and thereby likelihood of being cited (as opposed to the 'earmark hypothesis', which hypothesizes that media coverage merely 'earmarks' outstanding articles which would have received many citations anyway). Kiernan (2003a, p. 4) suspected that Phillips et al. might have "ascribed an elite status to *Times* coverage of scholarly research that may not exist", by not addressing that coverage by other media outlets might also influence scientists' reliance on certain research articles. To address this, Kiernan (2003a) did an analysis additionally taking into account the effects of coverage in twenty-four daily newspapers and the evening broadcasts of three major television networks from the United States. The author

found NYT coverage to not correlate significantly with citation rates once coverage by television and other newspapers is taken into account, suggesting that the NYT does not have unique influence as a disseminator of news about research to scholars. Even more recently, Fanelli (2013) tried to verify the publicity hypothesis for journal articles featured in British and Italian newspapers and found that the publicity effect is much stronger for English media, while publicity effects from Italian media are primarily local, i.e. mainly affect Italian authors.

Similarly, some studies specifically analyzed press releases as an instrument used by research institutions and scholarly publishers to communicate new research. In a quantitative content analysis of press releases, related journal articles, and news items on biomedical and health sciences, Sumner et al. (2014) found exaggerations in news to be strongly associated with exaggerations in press releases, concluding that improving accuracy of press releases might be a promising approach to reducing misleading news on health sciences. In a later study, Sumner et al. (2016) further investigated the relationship between mentions of caveats and exaggerations in journal articles and press releases and journalistic uptake. The study's findings suggest that press releases are frequently the source of both stated caveats and exaggerations concerning the reported research, but neither of the two seems to significantly affect the likelihood of news coverage. Stryker (2002) coded a sample of 95 journal articles from medical sciences for characteristics related to newsworthiness, examining their relationship to the articles' later amount of newspaper coverage, additionally considering whether respective articles were featured in press releases. The author found both newsworthiness and press release coverage to predict later newspaper coverage. In a similar vein, two previous studies had shown that about 80% and 84% respectively of research articles that get newspaper coverage had appeared in a press release beforehand (de Semir et al., 1998; Entwistle, 1995). Stryker (2002) however notes that the apparent effect of press releases on later newspaper coverage is reduced substantially when controlling for factors of newsworthiness. Complementing previous studies on the relationship between media coverage and citation rates, Chapman et al. (2007) did an analysis of the association between receiving a press release and the later citations and usage metrics of research articles from the journal *Tobacco Control*, finding press-released articles to receive more web hits, pdf downloads and citations than their counterparts without press releases.

On another note, concerning the point of intersection of science press releases and altmetrics, Bowman and Hassan (2019) examined *EurekAlert!*—an online science news service maintained by the American Association for the Advancement of Science—regarding the way scientific research is referenced in its news releases and regarding its presence across social media platforms. They found *EurekAlert!* to be the second most active source of Altmetric.com data for news releases, while showing only minimal activity on social media platforms. Moreover, Bowman and Hassan (2019) found that only a small share (18%) of *EurekAlert!* news releases referenced scientific research by DOI.

To briefly summarize, our review of past research on the relationship between external science communication and promoted articles' later metrics has shown that several case studies indicate an association between being selected for a press release or covered by news media and higher citations (Chapman et al., 2007; Fanelli, 2013; Kiernan, 2003a; Phillips et al., 1991) as well as usage metrics (Chapman et al., 2007) for respective research articles. In this exploratory case study, we aim to add to this body of research by (1) describing a previously under-analyzed, usually hard to obtain format of research promotion in the form of scholarly publishers' embargo e-mails to journalists, as well as by (2) also examining the relationship between such promotion and articles' later altmetrics.

Thus, in this exploratory case study we examine whether research articles mentioned in publishers' embargo e-mails sent to journalists differ regarding the bibliometric and altmetric indicators they receive, compared to articles without said mention. To gain insights as to which fields are typically represented in embargo e-mails, we additionally analyze mentioned articles' distribution across journals. We call this study exploratory, as it also represents the first steps to a higher-level goal of our ongoing research project, which is to arrive at a better understanding of the nature and extent of external science communication and its impact on research assessment practices.

Both citation-based indicators and altmetrics are often used as proxies for scientific productivity or relevance (Adie, 2016; Aksnes et al., 2019; Waltman, 2016). Oftentimes respective evaluations—at least if they are not conducted in a remarkably knowledgeable and careful way—start on the premise that as long as two articles originate from similar fields of research, are published in outlets of comparable renown and are of similar scientific quality, they will have roughly the same probability of getting cited. Thus, as long as the most decisive factors like field of research, publication type, and publication venue are controlled for, citations would be a useful proxy for scientific quality or relevance. From numerous past studies we do know however that factors affecting numbers of citations are manifold and diverse (see Tahamtan et al. (2016) for a review). We argue that, under the assumption that the publicity hypothesis as described by Phillips et al. (1991) holds true, research's promotion in external science communication would be such a factor that so far did not receive the scrutiny it would deserve. If articles' individual metrics can be shaped by publishers' promotional operations regardless of their scientific quality or merit, it seems of high importance to be able to inform users of such metrics about the degrees to which this might be the case.

The inclusion of altmetrics in this kind of study is of particular interest for two reasons. First, an argument frequently made in favor of altmetrics is that they might reflect a different form of impact than citations (Bornmann, 2014; Priem, Piwowar, and Hemminger 2012; Weller, Dröge, and Puschmann 2011; Wouters & Costas, 2012). Therefore, it is of interest to see whether altmetric indicators behave differently than citations in relation to press promotion, as this knowledge could help to further delineate altmetrics' potential benefit as complements to bibliometric indicators. Second, compared to citations, altmetrics more strongly blur the line between being the result and being part of promotional activities. Altmetrics that are frequently considered as indicators for received attention in many cases also directly include the amount of promotion undertaken to increase exactly this attention. For instance, measuring the tweets that mentioned a certain scientific paper will oftentimes mean to also measure tweets by publishers or involved authors, which were solely sent to advertise said paper (see also Haustein et al. (2014)). To be able to achieve a better understanding of this particular limitation of altmetrics, it therefore seems necessary to keep close watch over the relationship between manifestations of different altmetric indicators and promotional activities in external science communication, like for instance the embargo e-mail promotion considered in this study. While this study is not designed to conclusively settle the question regarding the degree to which altmetrics are comprised of advertising, it shall provide first numerical evidence to better describe how the promotion of articles to journalistic channels and the activity surrounding respective articles on platforms used as altmetric sources are connected.

## Methods and data

We compare two groups of articles regarding the metrics they received since their publication in 2016 to 2017—a treatment and a control group. Both groups coincide in regard to their individual articles’ journals and dates of publication, but the treatment group consists of articles that have been mentioned in embargo e-mails, while the articles from the control group have not. We measure attention as citations as measured by Web of Science and in the form of eight types of altmetrics provided by Altmetric.com.

Publishers’ embargo e-mails are no openly accessible data source—there is no such thing as a public archive gathering them, and building up a dataset of embargo e-mails is not trivial as usually accredited journalists are their only recipients. We rely on data from the *Science Media Center Germany (SMC)* (see also Rödder (2015)) to identify research articles with mentions in embargo e-mails. The SMC is a non-profit and editorially independent institution that supports journalists in reporting on science-related topics. One of its most important services is to send out statements by scientific experts commenting on new scientific findings while they are still under embargo. The SMC started its work in 2016 and since then has accumulated an archive of over 90,000 e-mails containing press materials, 2,638 of whom were identified as embargo e-mails. These embargo e-mails contain information about one or more upcoming articles in either one journal or several journals belonging to the same publisher. The SMC was founded by experienced science journalists, who specifically aim to monitor as many journals as possible that publish articles under embargo. Therefore, we can be confident that the SMC’s archive contains a large part of all embargo e-mails sent to journalists. However, the SMC does have a focus on topics deemed “public issues” by its staff, e.g. topics affecting particularly large parts of society like medicine, climate change, or artificial intelligence. Hence, the archive and the algorithms extracting the embargo e-mails might be biased towards these topics. Also, the archive and the software doing the extraction were not constructed with a scientific analysis in mind, but rather as a tool to support the resident journalists. Due to these characteristics of the SMC’s archive, we consider our undertaking to be a case study, as we cannot guarantee our sample’s representativeness for the entirety of scholarly publishers’ embargo e-mails. We hope, however, that the case study will shed light on the general characteristics of embargo e-mails and their relationship to citations and altmetrics.

We identified mentions of research articles in the extracted embargo e-mails by searching the archive for Crossref’s recommended regular expression for DOIs.<sup>1</sup> This is a fairly strict criterion for determining whether an article appeared in an embargo e-mail, as in many cases in embargo e-mails articles are not referred to by identifiers. However, a less strict search based on articles’ metadata seemed hardly to be an option—typically, titles used in embargo e-mails are preliminary or shortened, the amount of detail with which other metadata like author names are included varies substantially between publishers. Furthermore, for our study high precision seemed to be of particular importance—after all, the total population of articles that get promoted in an embargo e-mail will surely be much smaller than the population of articles that do not get this specific kind of promotion. Thus, using less restrictive criteria for determining if an article has been referenced in an embargo e-mail comes with a high risk of adding false-positives to our treatment group. The Figs. 1 and 2 show exemplary excerpts of two embargo e-mails sent by different publishers to

<sup>1</sup> <https://www.crossref.org/blog/doi-and-matching-regular-expressions/>.

**\*\*Embargo: 23.30hrs [UK time] Wednesday 3rd July 2019\*\***

Peer-reviewed / Observational and modelling study / People

**The Lancet Public Health: Incarceration and economic hardship strongly associated with drug-related deaths in the USA**

- *Unique analysis of US county-level data finds a strong association between incarceration and drug-mortality, and economic hardship and drug-mortality, independent of opioid prescription rates*
- *County-level incarceration may provide a further, plausible explanation to the underlying geographic variations in US drug-mortality, with the highest incarceration rates linked with a more than 50% increase in drug-mortality compared to counties with lowest incarceration*

Growing rates of incarceration in the USA since the mid-1970s may be linked with a rise in drug-related mortality, and may exacerbate the harmful health effects of economic hardship, according to an observational study involving 2,640 US counties between 1983 and 2014, published in *The Lancet Public Health* journal.

Major increases in admissions rates to local jails (with average rate of 7,018 per 100,000 population) and state prisons (averaging 255 per 100,000 population) were associated with a 1.5% and 2.6% increase in death rates from drug overdoses respectively, over and above the effects of household income and other county-level factors, such as violent crime, ethnicity, and education. Even after taking into account the role of opioid prescription rates, the association

Fig. 1 Excerpt of an embargo e-mail by a scholarly publisher—example A

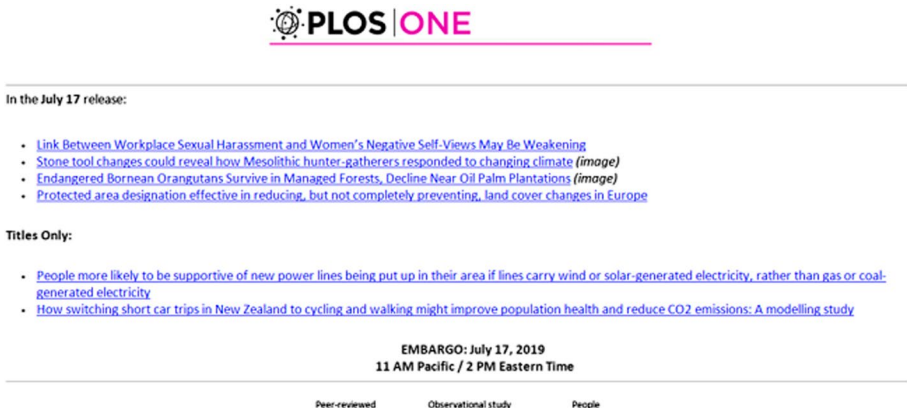
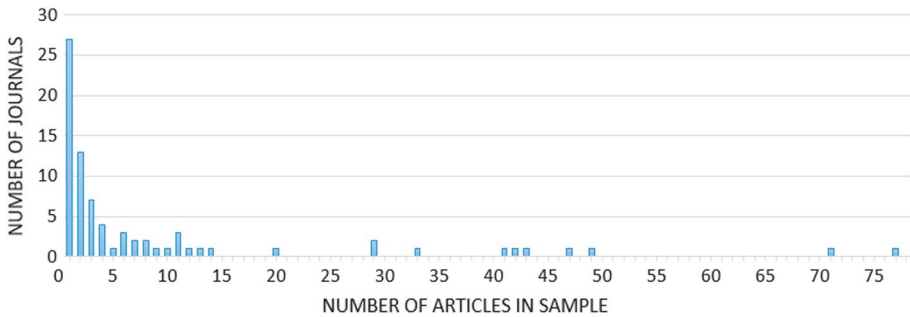


Fig. 2 Excerpt of an embargo e-mail by a scholarly publisher—example B

illustrate the e-mails' heterogeneity regarding structure and level of detail of the information included.

As we wanted to perform citation analysis of articles mentioned in embargo e-mails, to achieve appropriate citation windows we only considered articles that according to Crossref were published either in 2016 or 2017, the first two years since the SMC started archiving embargo e-mails. In addition to this treatment group of articles published in 2016 or 2017 that were mentioned in embargo e-mails from the SMC's archive, we constructed a control group of comparable articles that had not been mentioned in embargo e-mails. For every article in the treatment group, this control group consisted of one randomly chosen article published in the same month, year, and journal and that itself was not already part of neither treatment nor control group. Publication dates, journal affiliations, and DOIs of control group articles were obtained from Crossref's REST API.<sup>2</sup> For further analysis,

<sup>2</sup> <https://github.com/CrossRef/rest-api-doc>.



**Fig. 3** Distribution of treatment group articles across journals

journals' field categories were retrieved manually via web search, with Web of Science's Master Journal List<sup>3</sup> as a primary source, and then matched to Web of Science research areas.<sup>4</sup> Quantitative data on journal-level was obtained from Incites Journal citation reports<sup>5</sup> (JCR).

We obtained citation data for all articles from both groups from the Competence Centre for Bibliometrics,<sup>6</sup> which hosts bibliometric databases (which we from here on will call 'CCB databases') built on data from Web of Science. Altmetric data for all articles was obtained from Altmetric.com.<sup>7</sup> All queries were made in July 2021. To maximize recall, DOI-based queries were also performed for all-uppercase and all-lowercase transformations of the DOIs. It should be noted that the CCB databases are updated once per year in summer to Web of Science's state of calendar week 17, which means that the citation data retrieved in this study reflects the state of April 2020. Bibliometric and altmetric data was subsequently analyzed in R (R Core Team, 2020).

The regular expression search for DOIs in the SMC's e-mail archive initially retrieved 1,638 unique DOIs. After both automatic and manual cleansing of the data, 1,456 unique DOIs remained (the removed records for the most part consisted of false-positives returned by the extraction script, i.e. strings that did not include a real DOI known to Crossref). Of these, 715 referred to articles which had been first published in 2016 or 2017, according to Crossref. These DOIs form our treatment group.

## Results

To get an impression of the disciplines represented in embargo e-mails of the sample used, we have a look at the journals the 715 articles from our treatment group were published in (for the control group these numbers would be the same, as it contains exactly one counterpart from the same journal for every article from the treatment group). In total the articles were published across 78 different journals, with the numbers of articles per journal ranging from 1 to 77 articles (*PLoS Neglected Tropical Diseases* being

<sup>3</sup> <https://mjl.clarivate.com/>.

<sup>4</sup> [https://images.webofknowledge.com/WOKRS535R76/help/WOS/hp\\_research\\_areas\\_easca.html](https://images.webofknowledge.com/WOKRS535R76/help/WOS/hp_research_areas_easca.html).

<sup>5</sup> <https://jcr.clarivate.com/>.

<sup>6</sup> <http://www.bibliometrie.info/>.

<sup>7</sup> <https://www.altmetric.com/>.



**Table 1** Descriptive data on journals that published the articles in our sample of embargo e-mails

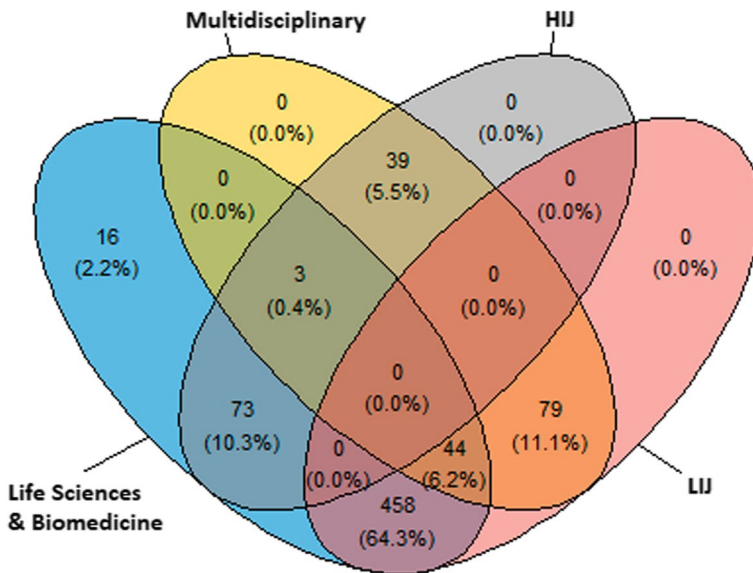
Indicator	Life sciences & biomedicine (n = 71)				Multidisciplinary (n = 9)			
	Min	Mean	Max	SD	Min	Mean	Max	SD
JIF '16	1.48	9.01	72.41	12.07	2.81	12.80	40.14	11.13
JIF'20	2.15	11.80	91.25	16.36	2.52	13.38	49.96	13.84
Total cites '16	957	26,584	315,143	50,146.39	9,495	260,400	671,254	268,743.23
Total cites '20	140	38,260	464,351	73,702.64	1,017	407,735	915,925	368,105.76
Citable it. '16	54	293	1,998	332.57	172	6,351	22,077	8,726.17
Citable it. '20	15	330	3,266	481.69	121	5,432	21,222	7,405.05

the most heavily represented journal). The most prevalent publishers in the sample were *PLoS* (357 articles), *BMJ* (198 articles) and *Springer Science and Business Media LLC* (102 articles). Figure 3 shows how many journals are represented in the sample by how many individual articles each. These first findings indicate a heavily skewed distribution regarding the representation of individual publishers and journals in the embargo e-mails, with few high-profile outlets accounting for a large share of the communication captured by our sample.

Examining the 78 journals' field categories as reported on the Web of Science Master Journal List reveals a strong representation of life sciences & biomedicine—71 of the 78 journals primarily cover topics from this research area. Three of these cases are additionally categorized as multidisciplinary, another three as also explicitly covering fields from social sciences. Six of the remaining 7 journals without explicit affiliation to life sciences & biomedicine were identified as multidisciplinary and most notably refer to certain prominent mega-journals (e.g., *Nature*, *PLoS ONE*, or *Proceedings of the National Academy of Sciences*). The last remaining journal primarily publishes articles from the field of psychology and is therefore categorized as primarily covering social sciences, when applying the Web of Science research areas. Out of the 78 journals in our sample, 9 were not found in the Incites JCR database (*BMC Hematology*, *BMC Psychology*, *BMC Research Notes*, *BMJ Case Reports*, *Drugs and Therapeutics Bulletin*, *Heart Asia*, *Injury Epidemiology*, *Marine Biodiversity Records*, and *Trauma Surgery & Acute Care Open*). Hence, we see that most but not all journals that send embargo e-mails are also indexed by the Web of Science.

The observed dominance of journals from life sciences and biomedical domains is in line with previous studies on the prevalence of certain scientific fields and topics in journalistic media like newspapers (see e.g., Elmer et al., 2008) or in scientific press releases (Hahn & Lemke, 2020). Due to the heavy representations of life science- and multidisciplinary journals in our sample, we will regard these two categories in particular detail for the rest of our analysis.

Table 1 shows, for the journals in our sample with JCR records, journal impact factors (JIF), total numbers of cites and numbers of citable items, in 2016, our first year of observation, and in 2020, the most recent year with available JCR data. The table shows large ranges across all parameters for both life science and multidisciplinary journals, indicating that the sample contains journals that implement highly diverse publication patterns. It should be noted that regarding numbers of citable items especially the native open access outlets of prominent publishers (e.g., *Scientific Reports*, *Nature Communications*, *PLoS ONE*) stand out as featuring particularly high numbers, which are often also tied to high numbers of total cites. However, regarding total cites also some high-profile (non-open



**Fig. 4** Venn diagram of relations between articles of the treatment group published in life sciences-, multi-disciplinary-, lower impact-, and higher impact journals

access) journals like *Nature* or the *New England Journal of Medicine* constitute outliers to the top.

To account for the substantial heterogeneity regarding expected average citations of the journals in our sample, we will also perform selected subsequent steps of analysis separately for articles published in “higher impact journals” (HIJ) and “lower impact journals” (LIJ). To define these two groups we apply the method of characteristic scores and scales as introduced by Glänzel and Schubert (1988) with  $k=2$ , based on the journals’ impact factors as reported in 2020. This procedure classifies 17 journals from our sample as HIJ, 51 journals as LIJ, while 10 journals remain unclassified due to them not having a journal impact factor in 2020. On an article-level, our treatment and control group each consist of 115 (16.08%) articles from HIJ, 581 (81.26%) articles from LIJ, and 19 (2.66%) articles from journals without journal impact factor in 2020.

If we measure the distribution of research areas across the 715 individual articles of our treatment group, we find that 594 (83.08%) of the articles were published in journals explicitly associated with life sciences & biomedicine, while 165 (23.08%) articles were published in multidisciplinary journals. Only 17 (2.38%) articles were published in one of the nine journals not covered by the Incites JCR.

Figure 4 shows the relations between articles from the two most prevalent discipline categories in our sample, life sciences & biomedicine and multidisciplinary, and their respective representation in higher and lower impact journals as a Venn diagram.

For 683 (95.52%) of the 715 articles from the treatment group Web of Science citation counts could be retrieved from the CCB databases, as was the case for 680 (95.10%) control group DOIs (the majority of DOIs for which no citation count could be retrieved belongs to articles from journals that were either not indexed in Web of Science or had only been indexed after the CCB databases’ latest update). On Altmetric.com, records were found for 714 (99.86%) treatment group articles, i.e. the article was mentioned at

**Table 2** Numbers (shares) of articles in both groups that received zero mentions on respective altmetric sources

Altmetric source	Treatment ( $n = 715$ )	Control ( $n = 715$ )
No Twitter mentions	7 (0.98%)	45 (6.29%)
No Facebook mentions	195 (27.27%)	447 (62.52%)
No Wikipedia mentions	615 (86.01%)	683 (95.52%)
No mainstream media mentions	57 (7.97%)	536 (74.97%)
No blog mentions	208 (29.09%)	566 (79.16%)
No Reddit mentions	555 (77.62%)	645 (90.21%)
No Youtube mentions	635 (88.81%)	704 (98.46%)
No Mendeley readers	1 (0.14%)	38 (5.31%)

least once on one of the platforms tracked by Altmetric.com, and for 677 (94.69%) of the articles in the control group. In analyses of altmetric counts, articles without records on Altmetric.com were assumed to have values of 0 across all altmetrics, as this should be the only regular circumstance under which an article with a valid DOI does not have a record on Altmetric.com.

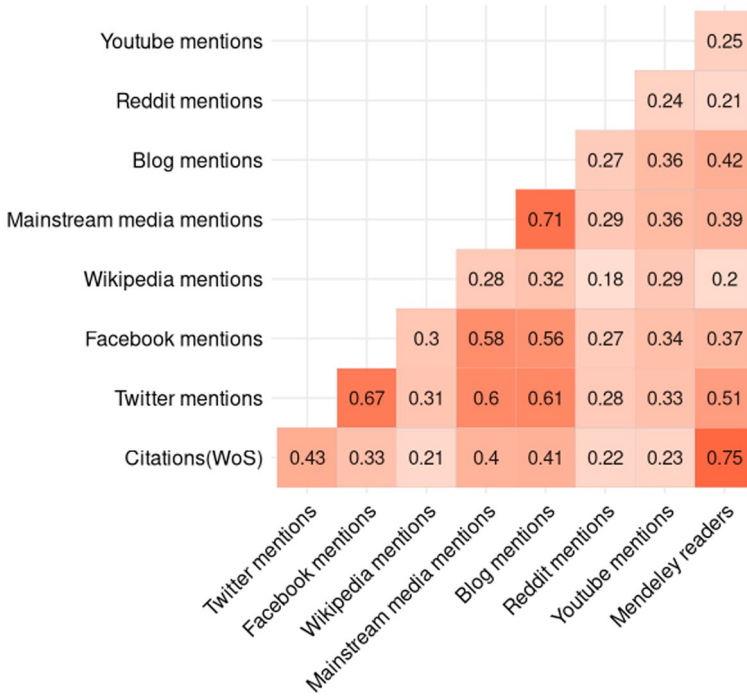
It has been shown that on various types of altmetric sources only very few articles receive any mentions at all, which can complicate their usage (Thelwall, 2018). To get a first rough idea of individual altmetric indicators' suitability for comparisons between the two groups of articles, the shares of articles in both groups that did not receive any mention in respective altmetric sources as provided by Altmetric.com's API are presented in Table 2.

The shares of articles that received no attention at all are on all platforms higher in the control group. Particularly high differences (percentage-wise) between both groups can be found for mentions in mainstream media and on blogs. On Wikipedia, Reddit, and Youtube, the majority of articles from both groups did not receive any mentions. Group-wise comparisons based on these indicators should therefore be interpreted with caution, as only few articles are responsible for all measured differences.

To arrive at a more precise picture of the data's structure, we next inspect correlations between the metrics in our sample. Figure 5 shows the Spearman rank correlations (pairwise complete observations) for the metrics data of the combined set of 1,430 articles from treatment and control group.

Among mostly weak to moderate correlations (ranging from 0.18 to 0.43), some strong correlations stand out: first, mentions on Twitter, Facebook, mainstream media, and blogs all correlate strongly with each other; second, Mendeley readers correlate strongly both with Twitter mentions and citations from Web of Science for the articles in our sample.

Table 3 shows the means and standard deviations of citations and altmetric counts that articles from both groups received between their publication in 2016 or 2017 respectively and our data collection. Comparing the means between treatment and control group reveals higher average counts per article from the treatment group in every single metric. Regarding most metrics the counts for treatment group articles are on average between 2 to 6 times higher than for control group articles. The most extreme relative difference is measured for mainstream media mentions, where average values among the treatment group are more than 10 times as high as among the control group.



**Fig. 5** Spearman rank correlations of sample articles’ metrics ( $n = 1,430$ ); all correlations are statistically significant with  $p < 0.001$

We cannot infer from Table 3 if both groups are to different extents characterized by extreme outliers, although high standard deviations in comparison to respective means indicate generally wide spreads of individual values. To get a better understanding of the data’s distributions, the 5% trimmed mean and median of all indicators per group have been calculated, as shown in Table 4.

Using outlier-robust measures overall does not reduce the previously observed advantages of treatment group articles substantially. On the contrary, for mentions on mainstream media and blogs the relative advantages of the treatment group rise even further. The medians of zero for Wikipedia, Reddit and Youtube mentions confirm once more what we have already seen in Table 2: for articles from both groups, mentions on these platforms are the exception rather than the rule.

Figure 6 depicts the two groups’ distributions explored in Table 3 and Table 4 graphically, as boxplots.

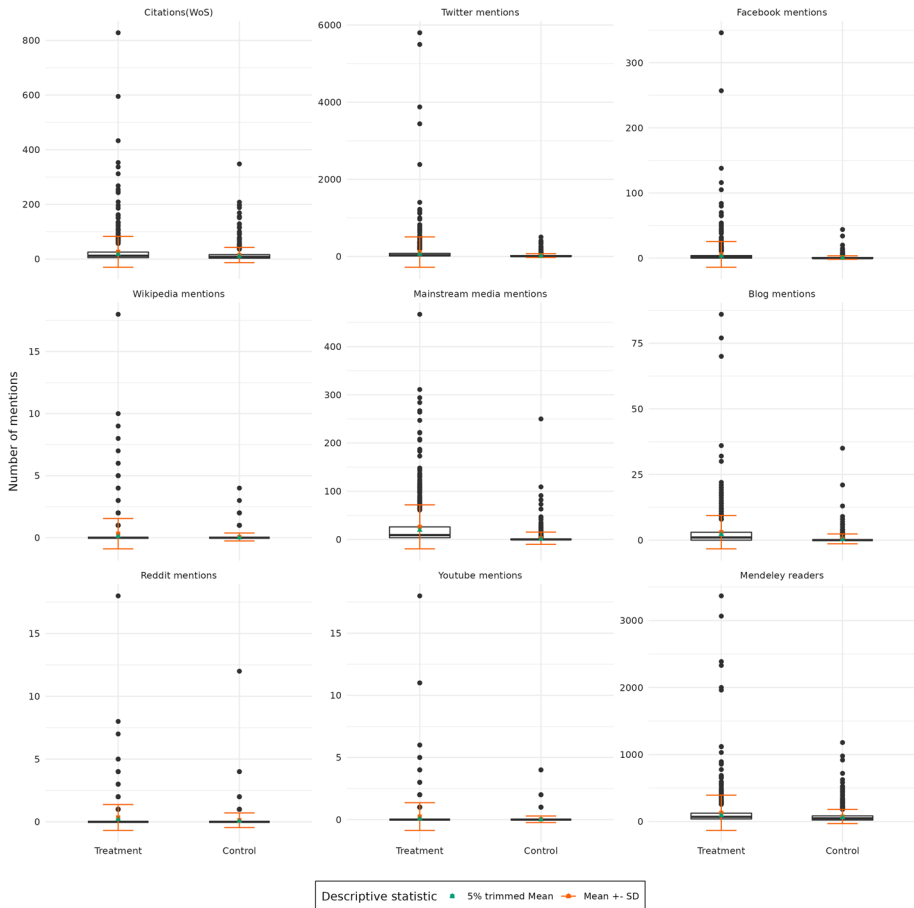
Next, we calculate means, 5% trimmed means, and standard deviations individually for the two discipline categories prevalent in our sample, ‘life sciences & biomedicine’ and ‘multidisciplinary’, to explore whether articles from the discipline-wise diverse journals behave differently from the articles published in more mono-thematic journals. Table 5 shows the three measures for articles published in life sciences- and biomedicine-journals ( $n = 594$  per group), Table 6 shows the respective measures for articles published in multidisciplinary journals ( $n = 165$  per group).

**Table 3** Means and standard deviations of bibliometric and altmetric indicators across both groups

Indicator	Mean		Standard deviation	
	Treatment		Control	
	Treatment	Control	Treatment	Control
Citations (WoS)	26.37	14.63	56.50	27.80
Twitter mentions	114.82	24.23	393.37	49.59
Facebook mentions	5.78	0.92	19.76	2.68
Wikipedia mentions	0.33	0.06	1.22	0.32
Mainstream media mentions	26.12	2.52	45.62	12.67
Blog mentions	3.03	0.48	6.34	1.87
Reddit mentions	0.35	0.13	1.03	0.58
Youtube mentions	0.25	0.03	1.11	0.26
Mendeley readers	130.44	74.44	263.11	105.62

**Table 4** Outlier-robust measures of bibliometric and altmetric indicators across both groups

Indicator	5% trimmed mean		Median	
	Treatment	Control	Treatment	Control
Citations (WoS)	17.74	10.08	12	8
Twitter mentions	61.65	16.02	33	9
Facebook mentions	3.07	0.55	2	0
Wikipedia mentions	0.13	0.00	0	0
Mainstream media mentions	19.20	0.74	9	0
Blog mentions	2.15	0.22	1	0
Reddit mentions	0.20	0.05	0	0
Youtube mentions	0.07	0.00	0	0
Mendeley readers	92.29	59.13	70	47



**Fig. 6** Boxplots of the two groups’ metrics

**Table 5** Descriptive statistics for articles from life sciences & biomedicine-journals

Indicator	Mean		5% trimmed mean		Standard deviation	
	Treatment	Control	Treatment	Control	Treatment	Control
Citations (WoS)	21.36	12.38	16.11	9.21	37.90	23.14
Twitter mentions	104.14	22.87	61.52	15.54	318.82	45.70
Facebook mentions	5.21	0.83	2.85	0.53	19.42	2.31
Wikipedia mentions	0.23	0.05	0.07	0.00	1.03	0.29
Mainstream media mentions	25.98	1.69	19.01	0.47	45.80	7.70
Blog mentions	2.48	0.30	1.86	0.16	4.57	0.78
Reddit mentions	0.28	0.10	0.17	0.04	0.68	0.57
Youtube mentions	0.19	0.02	0.05	0.00	0.97	0.20
Mendeley readers	106.84	67.67	84.47	56.64	157.47	80.68

**Table 6** Descriptive statistics for articles from multidisciplinary journals

Indicator	Mean		5% trimmed mean		Standard deviation	
	Treatment	Control	Treatment	Control	Treatment	Control
Citations (WoS)	43.66	22.20	28.67	15.92	90.09	36.78
Twitter mentions	163.93	37.41	84.91	27.56	557.58	61.57
Facebook mentions	7.26	1.31	3.86	0.75	18.46	3.54
Wikipedia mentions	0.71	0.14	0.41	0.05	1.70	0.52
Mainstream media mentions	25.25	5.36	18.72	2.33	43.23	21.83
Blog mentions	5.24	1.16	3.74	0.58	10.20	3.57
Reddit mentions	0.65	0.20	0.39	0.11	1.75	0.58
Youtube mentions	0.42	0.07	0.17	0.00	1.41	0.40
Mendeley readers	219.86	105.62	139.83	80.53	453.82	159.88

In both treatment and control group, our sample of articles from multidisciplinary journals on average performs better across almost all considered metrics—which, considering the larger relative share of articles from high impact journals among the group of multidisciplinary articles seen in Fig. 4, may not come as too much of a surprise. Higher average counts for multidisciplinary articles might to a considerable extent be explained by the fact that our small selection of multidisciplinary journals is mostly comprised of fairly high-profile journals like *Nature* or *PLoS ONE*, while our larger sample of journals covering life sciences & biomedicine represents a more mixed range. An interesting exception are the mainstream media mentions of treatment group articles from life sciences & biomedicine, which are slightly higher than those of the articles with embargo e-mail promotion that were published in multidisciplinary journals. This might indicate that, regarding coverage in mainstream media, articles from life sciences & biomedicine derive an even larger relative benefit from the treatment of being promoted in an embargo e-mail, than the articles from multidisciplinary journals.

**Table 7** Descriptive statistics for articles from higher impact journals

Indicator	Mean		5% trimmed mean		Standard deviation	
	Treatment	Control	Treatment	Control	Treatment	Control
Citations (WoS)	67.01	32,44	50.80	24.22	112.41	58.54
Twitter mentions	299.98	46.73	183.39	32.85	711.22	85.29
Facebook mentions	17.77	2.15	10.81	1.24	43.42	5.58
Wikipedia mentions	0.50	0.17	0.30	0.09	1.19	0.49
Mainstream media mentions	40.94	8.98	32.83	4.26	62.43	28.12
Blog mentions	5.80	1.48	4.27	0.81	10.42	4.16
Reddit mentions	0.60	0.18	0.42	0.08	1.23	0.63
Youtube mentions	0.68	0.12	0.31	0.01	2.19	0.59
Mendeley readers	297.37	118.57	210.77	87.77	511.15	209.80

**Table 8** Descriptive statistics for articles from lower impact journals

Indicator	Mean		5% trimmed mean		Standard deviation	
	Treatment	Control	Treatment	Control	Treatment	Control
Citations (WoS)	18.31	11.24	14.06	9.47	30.75	14.11
Twitter mentions	79.88	19.70	45.25	13.95	286.88	36.25
Facebook mentions	3.47	0.69	2.24	0.45	8.67	1.52
Wikipedia mentions	0.30	0.04	0.10	0.00	1.25	0.28
Mainstream media mentions	22.95	1.31	16.29	0.40	41.20	5.67
Blog mentions	2.52	0.29	1.80	0.15	5.12	0.80
Reddit mentions	0.31	0.12	0.18	0.05	1.00	0.58
Youtube mentions	0.17	0.01	0.05	0.00	0.73	0.12
Mendeley readers	100.12	67.40	79.27	59.17	164.71	67.39

**Table 9** Significances and effect sizes of mean differences between treatment and control group, based on all articles in the sample. All *p*-values are Bonferroni-adjusted

Indicator	<i>U</i>	<i>p</i>	<i>r</i>
Citations (WoS)	286,952	<0.001	0.199
Twitter mentions	384,232	<0.001	0.436
Facebook mentions	372,105	<0.001	0.415
Wikipedia mentions	280,352	<0.001	0.167
Mainstream media mentions	453,776	<0.001	0.695
Blog mentions	395,562	<0.001	0.518
Reddit mentions	288,520	<0.001	0.175
Youtube mentions	280,270	<0.001	0.197
Mendeley readers	321,304	<0.001	0.222

In the next step, we once more calculate means, 5% trimmed means and standard deviations, this time separately for the group of articles published in higher impact journals (*n* = 115; Table 7) and articles published in lower impact journals (*n* = 581; Table 8).



**Table 10** Significances and effect sizes of mean differences between treatment and control group, based on publishing journals' disciplines. All *p*-values are Bonferroni-adjusted

Indicator	Life Sciences & Biomedicine			Multidisciplinary		
	<i>U</i>	<i>p</i>	<i>r</i>	<i>U</i>	<i>p</i>	<i>r</i>
Citations (WoS)	199,262	<0.001	0.212	16,262	0.031	0.187
Twitter mentions	263,026	<0.001	0.426	19,650	<0.001	0.384
Facebook mentions	253,555	<0.001	0.400	20,132	<0.001	0.432
Wikipedia mentions	189,060	<0.001	0.137	16,092	<0.001	0.236
Mainstream media mentions	319,713	<0.001	0.731	22,376	<0.001	0.567
Blog mentions	273,427	<0.001	0.528	21,066	<0.001	0.496
Reddit mentions	198,811	<0.001	0.181	16,124	0.005	0.214
Youtube mentions	190,684	<0.001	0.185	15,588	0.001	0.232
Mendeley readers	221,190	<0.001	0.220	17,337	<0.001	0.237

**Table 11** Significances and effect sizes of mean differences between treatment and control group, based on publishing journals' impact factors. All *p*-values are Bonferroni-adjusted

Indicator	Higher Impact Journals			Lower Impact Journals		
	<i>U</i>	<i>p</i>	<i>r</i>	<i>U</i>	<i>p</i>	<i>r</i>
Citations (WoS)	8,490.5	<0.001	0.322	193,312	<0.001	0.161
Twitter mentions	11,304	<0.001	0.613	249,472	<0.001	0.414
Facebook mentions	10,838	<0.001	0.560	240,629	<0.001	0.393
Wikipedia mentions	7,456	0.585	0.164	184,373	<0.001	0.173
Mainstream media mentions	11,044	<0.001	0.586	304,403	<0.001	0.728
Blog mentions	10,506	<0.001	0.525	260,371	<0.001	0.522
Reddit mentions	7,907	0.013	0.239	189,012	<0.001	0.166
Youtube mentions	7,754.5	0.008	0.248	182,752	<0.001	0.190
Mendeley readers	9,491.5	<0.001	0.377	204,550	<0.001	0.183

Mean values for both treatment and control group articles are across all metrics substantially higher if we only consider articles from higher impact journals, compared to articles from the lower impact journals. Also, just like in the previous comparisons, mean values are across all metrics higher for treatment than for control groups.

As a last step, Mann–Whitney-*U* tests are conducted to check for statistical significance of the observed mean differences. Effect sizes are reported in form of *r* values (Fritz et al., 2012) and calculated following the implementation from R-package *rcompanion*.<sup>8</sup> We start with calculating effect sizes for the treatment and the control group in their entirety (Table 9) and then repeat the process for article groups distinguished by discipline (Table 10) and impact factor of their journal (Table 11), analogous to our previous steps. Table 9 shows the respective *U* statistics, Bonferroni-adjusted *p*-values and effect sizes *r* for the differences regarding the entire treatment and control group.

The observed differences between treatment and control group in their entireties are statistically significant regarding all nine examined metrics (Table 9). According to

<sup>8</sup> <https://www.rdocumentation.org/packages/rcompanion/versions/2.3.25/topics/wilcoxonR>.

guidelines reported by Fritz et al. (2012),  $r$  values of 0.1 indicate a small effect, values of 0.3 a medium effect, and values of 0.5 a large effect. Large effects were measured for mainstream media mentions and blog mentions, a large to medium effect was measured for Twitter mentions and Facebook mentions. Regarding the remaining metrics, measured effect sizes were medium to small.

The discipline-specific effect sizes seen in Table 10 suggest that across most indicators, the association between embargo e-mail promotion and an article's later metrics does not differ substantially between the two discipline categories under observation, with most differences between respective effect sizes only ranging between 0.02 and 0.05. However, one difference stands out particularly in this regard: for articles from life sciences & biomedicine, the positive association with mainstream media coverage seems even higher ( $r=0.731$ ) than for articles from multidisciplinary journals ( $r=0.567$ ). This backs up a similar observation we made during the comparison of Tables 5, 6.

More evident differences between effect sizes can be seen in Table 11 when comparing articles from higher impact journals to those from lower impact journals. Regarding most metrics (citations, Mendeley readers, mentions on Twitter, Facebook, Reddit, and Youtube), for articles from higher impact journals being promoted in an embargo e-mail is associated with stronger positive effects than for articles from lower impact journals. Regarding coverage in mainstream media, however, articles from lower impact journals seem to profit even more from the embargo e-mail promotion.

## Discussion

We compared the bibliometric and altmetric counts of articles which had been mentioned in publishers' embargo e-mails to journalists (treatment group) with those of articles from the same journals and publication months without such mentions (control group). We observed statistically significant advantages for treatment group articles across all nine examined metrics. Particularly strong effects were measured regarding mentions in mainstream media and on blogs, followed by Twitter mentions, Facebook mentions, Mendeley reader counts, and citations. Still significant small to medium effects were measured for mentions on Youtube, Wikipedia, and Reddit, although only small shares of both article groups received any mentions on these platforms at all. Our observations when differentiating between articles based on their journals' disciplines and impact factors suggested that the observed effects for the most part do not depend substantially on discipline, while across most metrics, articles from higher impact journals do seem to derive even larger comparative advantages from embargo e-mail promotion than those published in lower impact journals. A remarkable exception to these observations concerns mainstream media mentions, where measured effects are stronger for articles from life sciences & biomedical and lower impact journals than for multidisciplinary and higher impact journals. A possible explanation for this might be the fact that articles from higher impact journals tend to get a fairly high chance of being represented in mainstream media in any case (see also Table 7), regardless of them receiving additional promotion in embargo e-mails. Presumably, many science journalists regularly scan the prominent high impact journals for interesting articles anyway, while articles in lower impact journals might be more reliant on specific promotion to catch the attention of a significant number of journalists.

Multiple possible effects could explain our findings. First, one could assume embargo e-mails to evoke a publicity effect, in line with findings by Phillips et al. (1991) and Fanelli (2013) for newspaper articles. Our results indicate that articles mentioned in embargo e-mails do in fact get significantly more attention on mainstream media and social media platforms. Under the assumption of the publicity hypothesis, these findings would imply that scholarly publishers do indeed exert substantial influence on which articles (and thereby topics) get presence in the mass media, which would be in accordance with suppositions by Kiernan (1997). Moreover, the attention-increasing effect of embargo e-mail promotion apparently goes beyond the media outlets addressed by the e-mails themselves, as is evidenced by promoted articles' higher counts of scientific citations and Mendeley readers. In other words, the attention attracted by embargo e-mails in the public sphere of news media seems to also radiate into the scholarly sphere, which is represented by citations and Mendeley.

The validity of the publicity hypothesis for embargo e-mails would have further implications with regard to impact metrics as a means of research assessment. The use of citations as indicators for academic influence is underpinned by the theory that scholarly publishing is—at least for the most part—its own system, separate for instance from public media. If the publicity hypothesis holds true, it curtails the validity of this assumption. If promotional tools such as embargo e-mails could be used to push the metrics for publications independently from their scientific qualities, this would pose a limitation for the usefulness of such indicators. The promotion of selected articles by scholarly publishers would constitute a substantial interference with the scientific reward system and would add another bias to those already known in the context of citation-based evaluations (Hicks et al., 2015; Tahamtan et al., 2016). And unlike in peer review, the scholarly publishers are not compelled to apply established criteria of good scholarly practice when selecting which articles to promote in embargo e-mails, nor to make their criteria transparent.

However, our approach does not allow us to make definite statements about causalities between events behind different metrics, e.g., whether an article's mainstream media mentions typically motivate a substantial share of its scientific citations, or whether by and large its inherent qualities determine both academic and media uptake. Therefore, assuming embargo e-mails to be the decisive entry point into a chain of different media that an article then passes through with its publicity increasing along the way is not the only possible explanation of our findings. Instead, we could also assume a variation of the earmark hypothesis (Phillips et al., 1991) to be the determining factor behind our observations. In this case it would not be the science journalists, but the scholarly publishers and/or editors before them who on their own identify particularly noteworthy articles as being worthy to be included in embargo e-mails. And because journalists, researchers, and other readers later independently reach similar conclusions about which articles are noteworthy, the same articles that get chosen for embargo e-mails also get referenced more in mainstream media, on social media, in other scientific articles, and so on. Surely, although the degree of this hypothesis' validity may be unknown, it is to be expected that the selection of articles for embargo e-mail promotion does not happen at random but follows a variety of criteria (e.g., newsworthiness, prominence of the authors and institutions involved, scientific quality, etc.), some of which might also explain higher expected citations and/or altmetrics. To which degrees the promoted articles' metric advantages can be explained by which of the two hypotheses discussed in this article—publicity or earmark—remains an open question that will require further research. What this study could show, however, is the existence and magnitude of the advantage in metrics that embargo e-mail promoted articles can be expected to have.

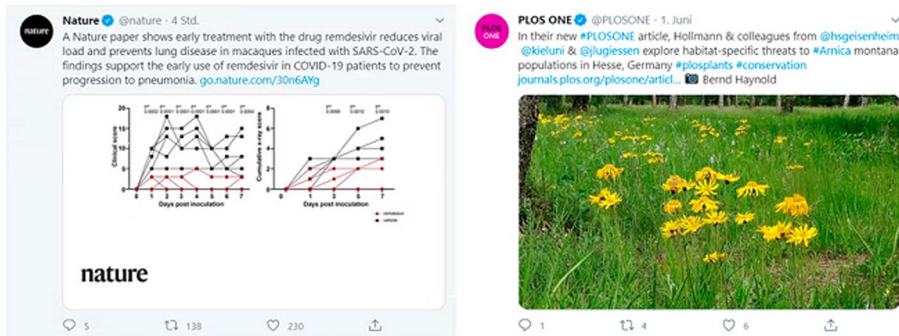


Fig. 7 Publishers ‘ tweets promoting research articles

For the scholarly community and its use of citations (and sometimes altmetrics) as indicators of scientific relevance, validity of the earmark hypothesis would surely be less problematic than validity of the publicity hypothesis. In any case, metrics’ susceptibility to numerous intrinsic and extrinsic factors (Tahamtan et al., 2016) will continue to be an issue that needs to be communicated with utmost transparency to their users, especially in light of increasing commitments to Open science. Furthermore, the various interests and opportunities to exert influence of the diverse stakeholders involved in science communication should continue to be discussed and made more transparent, as should the relationships between authors, publishers, and scientific reputation system.

As existing research on embargo e-mails’ role in the scientific communication system is relatively sparse, many different research streams could proceed from our results. One important aspect that should be investigated are the aforementioned possible causalities between metrics events. For instance, how common is the scenario of researchers first reading about scientific articles in mainstream- or social media, and because of that deciding to read (and later cite) the scientific article itself? Or the other way around: which role do scientific citations play for science journalists when they choose sources to back up their articles?

Furthermore, it should be examined whether mentions in embargo e-mails are merely a specific facet of broader promotional activities by the scholarly publishers. Such promotional activities could also have more direct effects on some of the metrics examined in this study—it should for instance be researched which role publishers’ postings play regarding individual articles’ mentions on social media platforms, like Twitter or Facebook. Examples for publishers directly engaging in social media postings to promote their own publications are manifold; see Fig. 7 for two typical examples of promotional tweets by publishers. The strong correlations we witnessed between mainstream media and various social media mentions (see Fig. 5) could also be an indicator of coordinated marketing efforts spanning multiple channels. Partly these correlations can be explained by the permeability and content overlap between respective platforms—many journalistic news articles also get tweeted and/or shared on Facebook the moment they are released—while the existence of a ‘publicity snowball effect’, where certain creators of social media content just copy stories that already received coverage on other platforms, might also play into these correlations. It remains to be analyzed which share of the total amount of postings mentioning scientific publications publishers account for, whether postings, tweets, and embargo e-mails are all parts of purposefully coordinated promotional activities for selected articles, and

which role cross-platform transitions play in the distribution of embargo e-mail promoted research.

Another worthwhile subject for further studies would be the comparison of the effects of embargo e-mail mentions on metrics to potential similar effects of different forms of external science communication, e.g. in press releases. Helpful insights would be provided by studying whether the promotion of certain articles leads to “spillover effects” of increased attention towards thematically similar articles without explicit promotion. Finally, it would be interesting to analyze the long-term temporal developments of the effects of article promotion examined in this study.

Although only indirectly related to our initial research question (and maybe peculiar to the sample of embargo e-mails used in this study), we have seen that articles from life sciences and biomedicine seem to dominate publishers’ communication in embargo e-mails. Moreover, our observations suggest that articles from life sciences-journals might benefit even stronger from being promoted in embargo e-mails than those from multidisciplinary journals regarding visibility in mainstream media. We cannot safely generalize this particular finding to the larger population of scientific articles though, as with regard to the journals represented in them, our samples were uneven between the two subject categories. This observation does however point towards another interesting area for future research.

Our study comes with some limitations. First, as noted before, we cannot guarantee our sample taken from the SMC’s e-mail archive to be representative of the entirety of all embargo e-mails sent by scholarly publishers to journalists. Besides, while access to the SMC’s e-mail archive allowed us to identify article DOIs that have definitely been mentioned in publishers’ embargo e-mails, it is virtually impossible to prove for control group articles that they have *not* been featured in similar forms of external communication, as not all such communication is recorded somewhere. Also, multidisciplinary journals constituted a considerable share of our sample—due to their multidisciplinary nature, the respective segment of our control group could differ topically from the respective segment of the treatment group. In addition, our choice of databases for citations and altmetrics (CCB and Altmetric.com) does of course introduce its own limitations—for instance, regarding the question what constitutes as a ‘mainstream media platform’, we rely on Altmetric.com’s list of monitored outlets. For a recent analysis of four major altmetric data providers’ methodological peculiarities, see Zahedi and Costas (2018). Finally, our restrictive way of determining articles’ appearances in embargo e-mails by looking for explicit DOIs probably means that our treatment group is not representative of all embargo e-mail activities performed by scientific publishers in 2016 and 2017—certain publishers might be more inclined to state DOIs in their e-mails than others (see also the findings Bowman and Hassan (2019) made regarding news releases on *EurekaAlert!*). How such potential publisher-specific strategies of promoting their articles affect impact measurements in different ways would be yet another interesting subject for future extensions to this research.

The exploratory case study presented in this article provides first insights on the relationship between publishers’ embargo e-mails and the attention that research articles promoted within them receive. Further research is necessary to explain the causes of the observed effects. The difficult accessibility of embargo e-mails will continue to be a major challenge for such endeavors.

**Acknowledgements** We thank Altmetric.com and the Competence Centre for Bibliometrics for providing us with access to their APIs and databases, respectively.

**Authors’ contributions** IP, MB, and SL contributed conception and design of the study; SR conducted the acquisition of data on embargo e-mails; MB and SL performed the statistical analyses and provided the

figures; IP acquired funding for the research project; SL retrieved data from Altmetric.com, Crossref, the *Competence Centre for Bibliometrics* (CCB) database, and Web of Science, performed the coding of journals to research areas and prepared this manuscript's first draft; all authors contributed to manuscript revision, read and approved the submitted version.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This study is part of the research project *MeWiKo* funded by the German Federal Ministry of Education and Research (Grant number 01PU17018).

**Data availability** The non-proprietary data on publications collected and analyzed in this study will be made available upon request to the corresponding author (SL). In accordance with their terms and conditions, altmetrics data obtained from Altmetric.com as well as the data from embargo e-mails collected by the Science Media Center Germany cannot be made available.

**Code availability** The R-scripts used for data analysis will be made available upon request to the corresponding author (SL).

## Declarations

**Conflicts of interests** One of the authors (IP) is member of the editorial board at *Scientometrics*.




**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adie, E. (2016). The rise of altmetrics. In A. Tattersall (Ed.), *Altmetrics: A Practical Guide for Librarians, Researchers and Academics* (pp. 67–82). Facet. <https://doi.org/10.29085/9781783301515.005>
- Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. *SAGE Open*, 9(1), 2158244019829575. <https://doi.org/10.1177/2158244019829575>
- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895–903. <https://doi.org/10.1016/j.joi.2014.09.005>
- Bowman, T. D., & Hassan, S. (2019). Science News and Altmetrics: Looking at EurekAlert!. Presented at *altmetrics19. The 2019 Altmetrics Workshop*. Stirling, Scotland. [http://altmetrics.org/wp-content/uploads/2019/10/Bowman\\_altmetrics19\\_paper\\_6.pdf](http://altmetrics.org/wp-content/uploads/2019/10/Bowman_altmetrics19_paper_6.pdf)
- Chapman, S., Nguyen, T. N., & White, C. (2007). Press-released papers are more downloaded and cited. *Tobacco Control*, 16(1), 71–71. <https://doi.org/10.1136/tc.2006.019034>
- de Semir, V., Ribas, C., & Revuelta, G. (1998). Press Releases of Science Journal Articles and Subsequent Newspaper Stories on the Same Topic. *JAMA*, 280(3), 294. <https://doi.org/10.1001/jama.280.3.294>
- Elmer, C., Badenschier, F., & Wormer, H. (2008). Science for Everybody? How the Coverage of Research Issues in German Newspapers Has Increased Dramatically. *Journalism & Mass Communication Quarterly*, 85(4), 878–893. <https://doi.org/10.1177/107769900808500410>
- Entwistle, V. (1995). Reporting research in medical journals and newspapers. *BMJ : British Medical Journal*, 310(6984), 920. <https://doi.org/10.1136/bmj.310.6984.920>
- Fanelli, D. (2013). Any publicity is better than none: Newspaper coverage increases citations, in the UK more than in Italy. *Scientometrics*, 95(3), 1167–1177. <https://doi.org/10.1007/s11192-012-0925-0>

- Franzen, M. (2012). Making Science News: The Press Relations of Scientific Journals and Implications for Scholarly Communication. In S. Rödder, M. Franzen, & P. Weingart (Eds.), *The Sciences' Media Connection –Public Communication and its Repercussions* (pp. 333–352). Dordrecht: Springer Netherlands. [https://doi.org/10.1007/978-94-007-2085-5\\_17](https://doi.org/10.1007/978-94-007-2085-5_17)
- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology. General*, 141(1), 2–18. <https://doi.org/10.1037/a0024338>
- Glänzel, W., & Schubert, A. (1988). Characteristic scores and scales in assessing citation impact. *Journal of Information Science*, 14(2), 123–127. <https://doi.org/10.1177/016555158801400208>
- Hahn, O., & Lemke, S. (2020). An Exploration of Scientific Press Releases in the Context of Altmetrics. Presented at *altmetrics20. The 2020 Altmetrics Workshop*. <https://doi.org/10.5281/zenodo.4446908>
- Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 65(4), 656–669. <https://doi.org/10.1002/asi.23101>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News*, 520(7548), 429. <https://doi.org/10.1038/520429a>
- Kiernan, V. (1997). Ingelfinger, Embargoes, and Other Controls on the Dissemination of Science News. *Science Communication*, 18(4), 297–319. <https://doi.org/10.1177/1075547097018004002>
- Kiernan, V. (2003a). Diffusion of News about Research. *Science Communication*, 25(1), 3–13. <https://doi.org/10.1177/1075547003255297>
- Kiernan, V. (2003b). Embargoes and Science News. *Journalism & Mass Communication Quarterly*, 80(4), 903–920. <https://doi.org/10.1177/107769900308000410>
- Kiernan, V. (2014). Public relations practices at medical journals. *Learned Publishing*, 27(1), 5–13. <https://doi.org/10.1087/20140102>
- Phillips, D. P., Kanter, E. J., Bednarczyk, B., & Tastad, P. L. (1991). Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community. *New England Journal of Medicine*, 325(16), 1180–1183. <https://doi.org/10.1056/NEJM199110173251620>
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. [ArXiv:1203.4745](https://arxiv.org/abs/1203.4745).
- R Core Team. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rödder, S. (2015). Science Media Centres and public policy. *Science and Public Policy*, 42(3), 387–400. <https://doi.org/10.1093/scipol/scu057>
- Stryker, J. E. (2002). Reporting Medical Information: Effects of Press Releases and Newsworthiness on Medical Journal Articles' Visibility in the News Media. *Preventive Medicine*, 35(5), 519–530. <https://doi.org/10.1006/pmed.2002.1102>
- Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., & Chambers, C. D. (2014). The association between exaggeration in health related science news and academic press releases Retrospective observational study. *BMJ British Medical Journal*. <https://doi.org/10.1136/bmj.g7015>
- Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Bott, L., Adams, R., & Chambers, C. D. (2016). Exaggerations and Caveats in Press Releases and Health-Related Science News. *PLoS ONE*, 11(12), e0168217. <https://doi.org/10.1371/journal.pone.0168217>
- Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, 107(3), 1195–1225. <https://doi.org/10.1007/s11192-016-1889-2>
- Thelwall, M. (2018). Altmetric Prevalence in the Social Sciences, Arts and Humanities: Where are the Online Discussions? *Journal of Altmetrics*. <https://doi.org/10.29024/joa.6>
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Weller, K., Dröge, E., & Puschmann, C. (2011). Citation analysis in Twitter. In: *Approaches for Defining and Measuring Information Flows within Tweets during Scientific Conferences*. CEUR-WS.Org, Tübingen University.
- Wouters, P., & Costas, R. (2012). *Users, narcissism and control – tracking the impact of scholarly publications in the 21st century*. SURFfoundation.
- Zahedi, Z., & Costas, R. (2018). General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PLoS ONE*, 13(5), e0197326. <https://doi.org/10.1371/journal.pone.0197326>

## Authors and Affiliations

Steffen Lemke<sup>1</sup>  · Max Brede<sup>2</sup>  · Sophie Rotgeri<sup>3</sup> · Isabella Peters<sup>1,2</sup> 

Max Brede  
mbre@informatik.uni-kiel.de

Sophie Rotgeri  
sophie.rotgeri@sciencemediacenter.de

Isabella Peters  
i.peters@zbw.eu

<sup>1</sup> ZBW–Leibniz Information Centre for Economics, 24105 Kiel, Germany

<sup>2</sup> Kiel University, 24118 Kiel, Germany

<sup>3</sup> Science Media Center Germany, 50678 Köln, Germany