

Identifying science in the news: An assessment of the precision and recall of Altmetric.com news mention data

Authors and affiliations

Alice Fleerackers*, Interdisciplinary Studies, Simon Fraser University, Canada

Lise Nehring, Centre for Forest Biology, University of Victoria, Canada

Lauren A. Maggio, Department of Medicine, Uniformed Services University, USA

Asura Enkhbayar, Interdisciplinary Studies, Simon Fraser University, Canada

Laura Moorhead, Journalism, College of Liberal and Creative Arts, San Francisco State University, USA

Juan Pablo Alperin*, Publishing Program, Simon Fraser University, Canada

Funding

This research is supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) insight grant, *Sharing health research* (#453-2020-0401). AF is supported by a Social Sciences and Humanities Research Council Joseph Bombardier Doctoral Fellowship (#767-2019-0369).

Disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University of the Health Sciences, the Department of Defense, or the U.S. Government.

Conflicts of Interest

None declared

Acknowledgements

We gratefully acknowledge Altmetric.com for access to data on news media mentions.

Authors' contributions (optional)

The authors confirm contribution to the paper as follows: study conception and design: AF, JPA; data collection: AF, LN, AE, JPA, LLM, LAM; analysis and interpretation of results: JPA, AF; draft manuscript preparation: AF, JPA, LAM, LLM, LN. All authors reviewed the results and approved the final version of the manuscript.

Abstract

The company Altmetric is often used to collect mentions of research in online news stories, yet there have been concerns about the quality of this data. This study investigates these concerns. Using a manual content analysis of 400 news stories as a comparison method, we analyzed the precision and recall with which Altmetric identified mentions of research in 8 news outlets. We also used logistic regression to identify the characteristics of research mentions that influence their likelihood of being successfully identified. We find that, for a predefined set of outlets, Altmetric's news mention data were relatively accurate (F-score = 0.80), with very high precision (0.95) and acceptable recall (0.70), although recall is below 0.50 for some news outlets. Altmetric is more likely to successfully identify mentions of research that include a hyperlink to the research item, an author name, and/or the title of a publication venue. This data source appears to be less reliable for mentions of research that provide little or no bibliometric information, as well as for identifying mentions of scholarly monographs, conference presentations, dissertations, and non-English research articles. Our findings suggest that, with caveats, scholars can use Altmetric news mention data as a relatively reliable source to identify research mentions across a range of outlets with high precision and acceptable recall, offering scholars the potential to conserve resources during data collection. Our study does not, however, offer an assessment of completeness or accuracy of Altmetric news data overall.

Keywords: Altmetric, data quality, news, accuracy, scholarly communication, journalism

Introduction

More than a decade after the release of the Altmetrics Manifesto (Priem et al., 2010), metrics derived from online mentions of research are now commonplace on the webpages of scholarly articles. However, while these data are collected through Altmetric, Elsevier's PlumX, and Crossref Event Data, the scholarly community continues to grapple with understanding their limits and potential uses. Although often extracted from openly available sources (i.e., web pages or public social media posts), the uncontrolled and ever-changing nature of the web makes collecting and analyzing altmetric data especially challenging. Whereas scholarly citation data is extracted from fairly well-structured documents that were created by scholars with the explicit intention of connecting documents using established professional conventions, mentions of research found on the web lack any such common intention or conventions. As a result, they are challenging to analyze and interpret. Further complicating matters, the algorithms used to identify mentions are not generally publicly shared, which casts further doubt on what is an already messy problem.

Despite these challenges, news coverage-related altmetrics have the potential to offer promising insights into the wider societal impact of research. Journalists have long played a gatekeeping role in society, as their decisions on what to cover can determine which information makes its way into public discourse (Shoemaker & Vos, 2009). While social media may be challenging this gatekeeping role to some extent (Bruns, 2018), news coverage remains one of the most common sources citizens use to learn about science (Covens et al., 2018; Funk et al., 2017). In part, this reliance on the news may be linked to journalists' efforts to make science more accessible. By providing relevant, understandable, and contextualized coverage of complex findings, journalists can play a "knowledge broker" role in society, allowing publics to comprehend and use research in their daily lives (Gesualdo et al., 2020; Yanovitzky & Weber, 2019). News-based metrics may thus offer a meaningful way to examine a wider societal impact of research than is possible through other altmetrics (Casino, 2018). Yet, despite the potential of news-based altmetrics, there have been concerns about the quality of this data source (Ortega, 2019a, 2020a), as little is known about what is actually being captured. This lack of empirical evidence into the quality of news mention data is becoming more concerning as the body of scholarship that relies on this data source grows. This study seeks to help fill this knowledge gap by assessing the quality of the news mention data provided by Altmetric.com, using manual content analysis of news stories as a comparison method.

In particular, this study supports individuals who rely on Altmetric to understand journalistic use of research by examining the precision and recall of the company's news mention data. Using a manual content analysis of 400 science news stories as a comparison method, it addresses the following research questions:

- RQ1: What proportion of Altmetric research mentions represent actual mentions (i.e., what is Altmetric's *precision*)?
- RQ2: What proportion of actual mentions does Altmetric successfully identify (i.e., what is Altmetric's *recall*)?
- RQ3: What characteristics of the research mention (e.g., presence of a hyperlink, description of research as "a study," inclusion of a journal title, author name, or publication date) influence its likelihood of being successfully identified?

Literature Review

Assessing the quality of altmetric provider data

A first challenge in determining the quality of altmetric data is that there is no formally agreed-upon definition of what constitutes a research “mention” and that any such definition is bound to be different for every social media platform or website being examined. However, a mention typically includes research information such as author names, journal titles, and study timeframes, or hyperlinks to research items (e.g., journal articles, datasets, images, white papers, reports). Altmetric data providers also vary in the methods they use to collect and process these mentions, meaning that different providers can yield different event data for the same research outputs (Karmakar et al., 2021). While there is evidence that differences between providers are decreasing over time (Bar-Ilan et al., 2019), concerns about data quality remain (Barata, 2019; Torres-Salinas, Castillo-Valdivieso, et al., 2018). Despite these concerns, the use of altmetric data in research studies remains common practice, likely because of the convenience of the data source and because of persistent interest in using metrics for research assessment, despite their potentially pernicious effects (Fitzpatrick, 2021; Hatch & Curry, 2020).

Among altmetric data providers, Altmetric.com (or “Altmetric”) is the most popular, used in more than half of altmetrics studies analyzed in a recent meta-analysis (Ortega, 2020b). Altmetric may be the preferred service because it gathers more mentions of research on Twitter, in blogs, and in news articles, relative to other popular providers (Karmakar et al., 2021; Ortega, 2020b; Zahedi et al., 2014). However, Altmetric’s data are also imperfect. For example, the platform only tracks Mendeley readership for research outputs that have already been mentioned on at least one social network, overlooking a substantial number of readers (Bar-Ilan et al., 2019). Altmetric data also appears to be particularly problematic when it comes to certain types of research items, such as scholarly monographs (Torres-Salinas, Gorraiz, et al., 2018) and non-English research (Barata, 2019).

More recently, and more relevant to this study, a series of studies by Ortega (2019a, 2020b, 2020b, 2021) have raised concerns about the quality of Altmetric’s news mention data in particular. Although Altmetric appears to collect more mentions of research in news stories than other providers, it does so for a smaller proportion of research items (e.g., peer-reviewed papers, preprints, monographs, clinical trials) and sometimes categorizes sources as both blogs and news, resulting in overlaps in event data (Ortega, 2019b). As many as 36% of links to news mentions on Altmetric are broken, a problem that is exacerbated for older mentions that were collected through external parties that have since become defunct (Ortega, 2019a). While Lehmkuhl and Promies (2020) offer some confidence in Altmetric’s ability to identify instances where research is *not* mentioned (i.e., its *recall*), this study is limited by a lack of understanding of the recall of the comparison data source used (i.e., the Nexis database).

In addition to these concerns, Altmetric appears to be biased towards news outlets that are published in English, based in English-speaking countries, and focused on general-interest issues, although its news data are still less biased than that of its competitors, PlumX and CED (Ortega, 2020a, 2021). In fact, news mention data gathered from different providers can differ widely; one study found that the Spearman correlation of Altmetric’s news data and that of its competitor, PlumX, was only 0.11 (Meschede & Siebenlist, 2018). Biases in news data, regardless of the provider, are likely mostly a

product of the selection of news outlets being tracked, as providers like Altmetric do not automatically gather mentions across the entire web, but rather in “a manually curated collection of sources” (Altmetric, 2020a). They may also be exacerbated by the algorithms used to identify mentions, which could themselves work better in English texts that follow certain journalistic conventions. As others have noted, it is unclear how representative this “completely unsystematic selection of media titles” is of the wider online media landscape (Lehmkuhl & Promies, 2020, p. 15), particularly given that many of the sources Altmetric counts as “news” outlets are news agencies or highly specialized media focusing on scientific literature (Robinson-Garcia et al., 2019). It is also likely that the quality of Altmetric news data changes over time. In 2022, Altmetric reported tracking more than 7,000 sources in over 165 countries (C. Williams, personal communication, February 18, 2022), but the list of sources is continually updated and revised. Over time, these updates may help improve the breadth of coverage that Altmetric is able to track and reduce some of the biases Ortega (2020a) identified; yet, it remains unknown when, or whether, they will be eliminated.

Collectively, the studies described above provide important insights into the limitations of Altmetric news mention data in terms of the sources it relies on; however, they do not address other possible limitations, such as those related to the methods Altmetric uses to identify mentions. According to Altmetric, news mentions are collected automatically, through a mixture of *link matching*—identifying URLs in news stories that link to research outputs—and *text mining*—crawling the stories for text-based descriptions of research-related metadata (Altmetric, 2020b). While these methods have strengths, they also have weaknesses. In particular, the second approach likely misses actual mentions of research, as any given news story “must include at least the name of an author, the title of a journal, and a publication date” in order for Altmetric’s text mining technology to classify it as a mention (Altmetric, 2020b). Although journalists often include at least one of these details when mentioning research, including all three is uncommon (Matthias et al., 2020). Instead, journalists often use “general terms” to describe the research they cover, referring to “scientific studies” rather than specific research institutions involved in the work (De Dobbelaer et al., 2018). Altmetric’s text mining technology likely overlooks at least a proportion of these text-based news mentions. To our knowledge, no empirical study has documented the frequency of these missed mentions.

Previous work that relies on Altmetric news mention data

Altmetric data is increasingly used to conduct research and assess broader, societal “impact” (Ortega, 2020b); flawed data compromises the integrity of both of these activities. Although the use of Altmetric data as a measure of social impact can be problematic, the use of the data as a source for research can yield rich insights into where, how often, and among whom research circulates online. Altmetric data has been used to find that research articles first posted as preprints receive more altmetric attention than those that were never preprinted (Fraser et al., 2020; Fu & Hughey, 2019). It has also been used to compare online attention to retracted papers, finding that retracted articles were 1.2–7.4 times more likely than matched, unretracted articles to receive altmetric attention (Serghiou et al., 2021). More recent work relied on altmetric data to document online attention to COVID-19-related preprints, demonstrating that unreviewed COVID-19 studies received far more attention on Twitter, news sites, and blogs than preprints on other topics (Fraser et al., 2021; Sevryugina & Dicks, 2021).

When it comes to *news* mention data specifically, providers such as Altmetric allow scholars to examine important questions at the intersection of journalism and science, such as the newsworthiness of

different types of research items, the “media impact” of particular countries or publications, and more (Casino, 2018). A small but growing body of research has taken advantage of these opportunities. For example, Maggio et al. (2019) relied on Altmetric data to examine online media coverage of US-government funded cancer research, finding a mismatch between the types of cancers that received the most media coverage and those with the highest incidence rates. Schultz (2021) used Altmetric to understand the relationships between open access (OA) status and the number of news mentions of articles in high-impact journals, finding that OA articles tended to receive more mentions than those published as completely closed access. Other news-focused studies have used Altmetric to not only identify the *amount* of coverage that research outputs receive but to *generate corpora of stories* that can later be analyzed to understand how journalists portray those outputs. Using this approach, Moorhead and colleagues (2021) found that traditional, legacy news sources included significantly more mentions of research on common cancers than digital native news sources. Matthias et al. (2020) used Altmetric data to identify and analyze mentions of research on opioid-related disorders in US and Canadian media, finding that this research was most often portrayed as “valid science,” with little discussion of study methods or limitations. More recently, scholars have used Altmetric data to understand how COVID-19 preprints are framed in English-language (Fleerackers et al., 2021) and Brazilian news media (Oliveira et al., 2021), finding that journalists inconsistently disclose the unreviewed nature of the research they cite. Studies such as this highlight the potential value of Altmetric News data but also underscore the importance of assessing the quality of the data source.

Methodology

To conduct the study, we gathered all articles published in the science and health sections of the following eight news media outlets during March-April 2021: *The Guardian (Science Section)*, *HealthDay*, *IFLScience*, *MedPage Today*, *News Medical*, *New York Times (Science Section)*, *Popular Science*, and *Wired*. These news publications were selected for their science and health focus, as well as their representation of the changing media landscape (i.e., *The Guardian* and *New York Times* as traditional, legacy news organizations; *Popular Science* and *Wired* as historically print-only science magazines; *News Medical* and *MedPage Today* as digital native health sites; and *HealthDay* and *IFLScience* as niche science and health blogs). In addition, of all the news sources Altmetric tracks, general-interest and specialized health outlets such as these also appear to cover the most research (Ortega, 2021). Between March and May 2021, we used a custom-built web-crawler to read the RSS and Twitter feeds of these sites to identify 5,172 articles and associated metadata (i.e., URLs, dates of publication, authors, article titles; Enkhbayar, 2022). The eight outlets appeared to vary widely in how frequently they posted content. Most outlets published around 50 articles per week during our collection period. However, both *Popular Science* and *News Medical* stood out with more than 145 stories a week, while *Wired* only published 11 articles a week. To ensure balanced representation of outlets within the dataset, and to make the manual content analysis more manageable, a random sample of 50 articles from each publication (400 articles total) was used to perform the study.

Next, adopting methods utilized in systematic reviews (Page et al., 2021) two independent coders (AF and LN) read and analyzed each of the 400 news stories using a codebook developed for this purpose (see Appendix). For each news story, each coder looked for mentions of research in the form of either hyperlinks to academic publications or text-based descriptions of research (e.g., “a study,” “new evidence,” etc.). They then used either the hyperlink provided or a standardized web search method

developed by the authors (outlined in Appendix) to identify the specific study mentioned and saved the corresponding identifier (e.g., DOI, Pubmed ID, arXiv ID, ISBN); each research mention was then coded to assess the contextual information provided in the news story (e.g., Was an author name mentioned? Did the story include the title of the journal where the research was published?). The two coders' coding sheets were then compared and any discrepancies in the coding were resolved through discussion. In instances in which consensus was unclear, LAM and LLM acted as a tiebreaker. The resulting dataset was thereafter considered the standard for comparison.

We used the Altmetric Explorer on September 9, 2021, to download the 10,021 research mentions that Altmetric identified in the same 8 outlets during March-April 2021. We then matched stories identified by Altmetric with our standard using the story URL and matched mentions using the research identifiers (e.g., DOI, Pubmed ID, arXiv ID, ISBN). We looked at all mentions identified by Altmetric for which there was no match in the standard data to verify that the news story did not have a corresponding mention. We corrected instances where Altmetric had an erroneous identifier if it was clearly caused by a typographical error (e.g., DOIs or arXiv IDs that did not resolve due to brackets or spaces that were out of place) but did not correct instances when the record pointed to an incorrect research article. Certain outputs, such as books, can be associated with more than one identifier (e.g., multiple ISBNs for multiple versions); we reviewed these cases manually and replaced the identifier provided by the coders with the one provided by Altmetric. The final news mention data can be found online at (Fleerackers et al., 2022)

Data were analyzed using Python's pandas¹ package. Binary logistic regressions were calculated using Python's statsmodels² package. All analysis scripts can be found online at (Alperin, 2022).

Results

We identified 502 research mentions among 228 of the 400 stories in our sample (the remaining 172 did not mention research). The number of stories with mentions and the number of mentions per story varied by outlet (Table 1); at one extreme, *The Guardian* only mentioned research in 18 (36%) of the 50 stories in the sample (with an average of 1.7 mentions in every story) and at the other, News Medical mentioned research in 43 (86%) of stories (with an average of 1.1 mentions per story). Importantly, it is possible that these outlets mentioned other research items that were not identifiable using either of our methods (i.e., Altmetric or content analysis). Specifically, the coders noted that many stories included phrases such as “studies have found” or “there’s a wealth of research,” which suggest that a body of evidence has been referenced but does not provide enough detail to identify specific research items. While we did not track such instances systematically, the results represented in Table 1 likely underestimate the true amount of research that is referenced in the media stories.

Outlet	Number of Stories	Number of Stories with Mentions	Percent of Stories with Mentions	Number of Mentions	Average Number of Mentions per Story
Health Day	50	30	60	32	1.1
Ifscience	50	31	62	62	2.0
MedPage Today	50	34	68	108	3.2
New York Times	50	21	42	54	2.6
Popular Science	50	18	36	37	2.1
The Guardian	50	18	36	31	1.7
News Medical	50	43	86	46	1.1
Wired.com	50	33	66	132	4.0
Total	400	228	57	502	2.2

Table 1. Number of stories and mentions across news outlets

Outlets also varied in how they referred to the research outputs they mentioned (Table 2). For example, while outlets such as *Wired*, News Medical, and MedPage Today provided hyperlinks to almost all of the research they mentioned, others did so less consistently, and one outlet (Health Day) never linked to research. In contrast, Health Day was among the most consistent when it came to describing bibliometric details—providing an author name and institution for almost every research mention, often alongside a journal and publication date—while other outlets did so less frequently.

Despite these differences, there were several observable trends in how outlets referred to the research they mentioned. The two most common ways to mention research items were to hyperlink to them and/or describe them with terms such as “research” or “studies.” Providing an author and an institution were the next most common strategies, followed by mentioning the journal or publication venue, and finally, indicating the publication date.

Outlet	Number of Mentions	Describes as research		HasLink		Institution Mentioned		Author Mentioned		Journal Mentioned		Study Date Mentioned	
		N	n	%	n	%	n	%	n	%	n	%	n
Health Day	32	31	97	0	0	31	97	30	94	25	78	25	78
Ifscience	62	45	73	50	81	20	32	21	34	20	32	9	15
MedPage Today	108	98	91	93	86	38	35	35	32	61	56	31	29
New York Times	54	36	67	40	74	26	48	28	52	15	28	16	30
Popular Science	37	29	78	25	68	15	41	14	38	10	27	10	27
The Guardian	31	24	77	21	68	18	58	11	35	6	19	9	29
News Medical	46	42	91	40	87	30	65	39	85	36	78	34	74
Wired.com	132	84	64	116	88	60	45	59	45	29	22	41	31
Total	502	389	77	385	77	238	47	237	47	202	40	175	35

Table 2. How research was mentioned across news outlets

Precision and recall of Altmetric’s research mention data

Altmetric identified at least one mention in 163 (71%) of the 228 news stories with mentions. Within these stories, it correctly identified 349 (70%) of the 502 mentions in the standard dataset, while also identifying 21 incorrect mentions. It also missed identifying 153 mentions. Standard information retrieval measures (e.g., precision, recall, and F-score) of Altmetric’s research mention data are summarized in Table 3; common errors contributing to these scores are described in detail below.

Metric	Description	Value
True Positive	Number of correctly identified mentions	349
False Positive	Number of incorrectly identified mentions	21
False Negative	Number of unidentified mentions	153
Precision	Proportion of identified mentions that were correct	0.95
Recall	Proportion of correctly identified	0.70
F-score	The harmonic mean of precision and recall	0.80

Table 3. Precision and recall of Altmetric research mention data

False Positives

In some instances (n = 21), Altmetric identified mentions to research that the two coders did not. A closer examination of these errors revealed that in most of these cases (n = 15) Altmetric had identified mentions to items that were not research, but rather journalistic stories in sources like *The Conversation* or *Nature News*. In three of the remaining cases, Altmetric had identified an erroneous research item (i.e., research that was not actually mentioned in the story) that had been published in the same journal, by one or more of the same authors, and around the same time as the true research mention (most likely due to Altmetric’s reliance on author names and journals for text mining). The final three cases were duplicates, in which Altmetric had identified the same mention twice for the same research item. This particular error is known to occur when news stories are revised or updated; even if only minor changes have taken place, resulting in duplicate mentions (C. Williams, personal communication, February 18, 2022).

False Negatives

False negatives, or instances in which Altmetric failed to identify research that both coders agreed was mentioned, were far more frequent than false positives (n = 153). While it is not possible to explain why these mentions were missed, we observed common trends in the types of mentions that most often resulted in false negatives. For example, Altmetric often failed to identify mentions of research in news stories that described the *results* of the research but provided little or no bibliographic information (i.e., no author name, journal, publication date) and did not include a hyperlink. Conference presentations, books, dissertations, and non-English research items were also often missed, particularly those that

were not associated with a DOI. Altmetric also failed to identify research when the news story included a link to a press release describing the research, but no link to the research itself. In a few cases, it was unclear why Altmetric had missed the mention, as all of the information needed to identify the research had been provided (i.e., a hyperlink or text-based description of the journal, author name, and publication date).

Accuracy across news outlets

Perhaps because news outlets differed in their use of research (as discussed above), the accuracy of Altmetric’s data varied across outlets (Table 4). In particular, Altmetric was most successful in identifying mentions of research in stories from outlets such as *News Medical* and *Wired*, and less accurate when it came to outlets such as *Popular Science* or *The Guardian*. We investigate some of the factors that could be driving these differences in the following section.

Outlet	Number of mentions	Precision	Recall	F-Score
Health Day	52	0.95	0.56	0.71
Ifscience	81	0.85	0.71	0.77
MedPage Today	124	1.00	0.77	0.87
New York Times	83	0.86	0.56	0.67
Popular Science	69	0.85	0.46	0.60
The Guardian	63	1.00	0.45	0.62
News Medical	53	0.97	0.83	0.89
Wired.com	149	0.97	0.80	0.88

Table 4. Precision, Recall, and Accuracy (F-score) by news outlet

Characteristics of research mentions that influence accuracy

To answer RQ3, we calculated a logistic regression that examined whether the probability of Altmetric identifying a research item depended on how the research was mentioned in the story. More formally, we calculated a model in the form $P(Y=1) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$, where Y is a binary outcome variable coded as 1 if Altmetric identified the correct mention and 0 otherwise, and $X_1...X_6$ are a set of predictor variables corresponding to the six characteristics examined, each coded as 1 if we identified the characteristic and 0 otherwise. We found that Altmetric was significantly more likely to identify the research mention when the it included a link (odds ratio = 53.8, $p < 0.001$), when the journal name was mentioned (odds ratio = 3.8, $p = 0.003$), and when the author was mentioned (odds ratio = 4.2, $p = 0.010$). The coefficients for the other three predictor variables were not significant (see Table 4). That is, Altmetric was no more or less likely to successfully identify a mention if the research in question was described as a study, included the author’s institutional affiliation, or provided a publication date.

	Standard Error	Odds Ratio (95% Confidence Interval)	p-value
Describes as Research	-1.014	0.65 (0.29–1.49)	0.311
Has Link	9.517	53.77 (23.67–122.17)	0.000
Journal Mentioned	3.002	3.85 (1.60–9.28)	0.003
Author Mentioned	2.587	4.23 (1.42–12.61)	0.010
Institution Mentioned	-1.367	0.49 (0.18–1.36)	0.171
Study Date Mentioned	-0.558	0.79 (0.35–1.79)	0.577

Table 5. Results of logistic regression

Discussion

This study assessed the accuracy of Altmetric’s news mention category, a data source increasingly used by scholars in a wide range of fields but about which relatively little is known. Using a manual content analysis of 400 news stories as a comparison method, we analyzed the precision and recall with which Altmetric identified mentions of research for a set of 8 news outlets, as well as which characteristics of how the research was mentioned influenced its likelihood of being successfully identified.

Our findings suggest that, when working from a predefined list of news outlets, and applying some caveats and several limitations, Altmetric can be a useful and relatively reliable source of news mention data. Doing so allows scholars to easily identify mentions of research in those outlets with very high precision and with variable, but in many cases, acceptable recall. That is, given a set of outlets similar to the ones we studied, Altmetric appears to use a conservative approach that ensures accuracy in their data at the expense of completeness. As a result, research mentions found by Altmetric for any given news outlet can be considered to represent a reliable lower bound of all research mentions for that outlet.

This finding is important because Altmetric offers scholars the potential to save considerable time and resources at the data collection stage of their research. The comparison method used for this study—manual content analysis—proved to be complex and resource-intensive. It took two researchers around 44 hours each to identify the 502 mentions in the final data. Of these, 36 hours were needed to identify and code the mentions and the remaining 8 were spent discussing discrepancies to arrive at a consensus. In comparison, identifying mentions via the Altmetric Explorer took about 5 minutes. In addition, our approach required the selection and curation of relevant news sources and a custom set of scripts to collect their published articles. While we relied on the publicly available dissemination

channels (e.g., APIs, RSS feeds, or Twitter feeds) to crawl relevant content, Altmetric might have more reliable access to such data through paid third-party services or direct agreements with publishers.

That said, our analysis offers assurances about the accuracy of Altmetric News data only under a limited set of conditions and should only be relied upon with an understanding of some important limitations. In particular, our analysis does not offer assurances about Altmetric News data as a whole. Scholars seeking to use Altmetric in their research should thus consider our results alongside other known limitations of this data source, including its linguistic, geographic, and disciplinary biases (Ortega, 2020a) and its high incidence of broken links (Ortega, 2019a). We note that there may also be other limitations of Altmetric News data that are not yet known.

In addition to assessing precision and recall, this study provides insight into the characteristics of how research is mentioned that influence its likelihood of being successfully identified by Altmetric. In particular, this data source seems to be most reliable for mentions of research that are accompanied by a hyperlink, followed by mentions of research that include an author name or journal title. In some ways, this finding is unsurprising; Altmetric's website explicitly states that, in cases where no hyperlink is present, a research mention must include an author name, journal title, and study date to be identified (Altmetric, 2020b). However, it is interesting that the presence of a study date made no significant difference in the likelihood that Altmetric would correctly identify a given research output. It is unclear why this is the case, but it could be related to how news stories tend to describe publication dates. In particular, many of the stories we analyzed did not include an explicit year of publication; instead, they referenced publication dates using statements such as "released last week," "published Monday," or, even more frequently, simply "new." Future research is needed to better understand how or whether these more conversational references to study dates influence Altmetric's ability to identify the research. More broadly, our results suggest that Altmetric struggles to identify mentions of research that offer few or no bibliographic details (for example, mentions that describe research studies using only their key findings). Our coders identified this as a relatively uncommon practice; only 30 (6%) of the 502 research mentions in our data set did not include any bibliographic information, but its prevalence could be different among a different set of stories.

In addition, Altmetric news mention data for scholarly books appears to be particularly problematic. Books are a challenge in part because multiple ISBNs may be associated with the same book, as noted by other scholars (Torres-Salinas, Gorraiz, et al., 2018). However, our study suggests an additional challenge of distinguishing between mentions of academic and trade books. In our data, Altmetric had an inflated number of research mentions in the data set from books that both coders agreed were not scholarly (e.g., self-help or health books that were written by authors with academic affiliations but were published by trade publishers rather than scholarly presses). This is an important limitation that will likely be exacerbated when exploring questions related to humanities research, whose papers are less frequently mentioned in altmetric sources (Thelwall, 2018), and where monographs are a more common form of scholarship (Hicks, 2005). As such, extra caution should be taken by those who wish to use altmetric data to measure academic success or "impact."

Finally, we note that in addition to tracking news mentions of research items, such as journal articles or preprints, Altmetric sometimes also tracks news mentions of blog posts or news stories *about* research items—so-called "second-order citations" (Priem & Costello, 2010). While understanding the circulation

of science news or blog posts may be useful in certain contexts (e.g., for examining how science journalists cite each other's work), it may add confusion in studies seeking to understand media attention to scholarly research. Researchers using Altmetric for this purpose should be aware that the data may contain both first- and second-order citations of research and, if appropriate, filter out those outputs marked as “news” before performing any analysis. They should also be aware that Altmetric only tracks certain types of news stories and blog posts mentioning research (e.g., those published by *The Conversation*, *Science*, *IEEE Spectrum News*), while other types of second-order citations, such as press releases about research, are not.

Limitations

These findings must be understood alongside several limitations. First, we report on the quality of Altmetric news data for a predefined set of 8 news outlets, while Altmetric collects research mentions for more than 4,000. Although these 8 sources were selected because they cover a range of outlet types that are likely to mention research, they are not representative of all outlets that might mention research, nor all outlets tracked by Altmetric. For example, it is likely that Altmetric's precision and recall for identifying mentions of research in non-English outlets differ from those found in this study. Additionally, as noted by Lehmkuhl and Promies (2020), Altmetric data includes sources that could arguably not be considered “news,” including content aggregators, such as *Foreign Affairs New Zealand*; press alert services, such as EurekAlert; websites that allow anyone with an account to produce content, such as Medium; and outlets known to promote misinformation, such as Zero Hedge (Cranley, 2020). The data can also include duplicate links or discrete links that highlight duplicate mentions of research (e.g., research briefs or press releases with minor changes in text or publishing timestamps). Future research should extend these findings using a larger or distinct subset of news outlets, including content aggregators, particularly given that precision and recall do appear to differ somewhat across outlets.

Finally, this study examined how characteristics of the news story (e.g., outlet) and news mention (e.g., presence of a hyperlink, author name) influence Altmetric's ability to accurately identify mentions of research, but we did not examine any characteristics of the research items themselves. Although we provide some qualitative evidence that certain kinds of outputs (e.g., books, conference papers) appear to generate more problems for Altmetric than others, future research should more rigorously examine the factors that influence a given output's likelihood of being successfully identified (e.g., type of research output, journal published in, how old it is, language, etc).

Endnotes

1. <https://pandas.pydata.org/>
2. <https://www.statsmodels.org/>

References

- Alperin, J. P. (2022). *Code for publication: Identifying science in the news*. Zenodo. <https://doi.org/10.5281/zenodo.6332953>
- Altmetric. (2020a, September 17). *News and mainstream media*. Altmetric. <https://help.altmetric.com/support/solutions/articles/6000235999-news-and-mainstream-media>
- Altmetric. (2020b, September 17). *Text mining*. Altmetric. <https://help.altmetric.com/support/solutions/articles/6000240263-text-mining>
- Barata, G. (2019). Por métricas alternativas mais relevantes para a América Latina. *Transinformação*, 31. <https://doi.org/10.1590/2318-0889201931e190031>
- Bar-Ilan, J., Halevi, G., & Milojević, S. (2019). Differences between Altmetric Data Sources – A Case Study. *Journal of Altmetrics*, 2(1), 1. <https://doi.org/10.29024/joa.4>
- Bruns, A. (2018). *Gatewatching and News Curation: Journalism, Social Media, and the Public Sphere* (New Edition). Peter Lang Inc., International Academic Publishers.
- Casino, G. (2018). *Press citation: The impact of scientific journals and research articles on news media*. 27(3). <https://doi.org/10.3145/epi.2018.may.22>
- Covens, L., Kennedy, J., & Ryan, P. (2018). *Ontario Science Centre Canadian Science Attitudes Research*. Leger. https://www.ontariosciencecentre.ca/Uploads/AboutUs/documents/Ontario_Science_Centre-Science_Literacy_Report.pdf
- Cranley, E. (2020, February 1). Finance blog Zero Hedge was banned from Twitter for Wuhan coronavirus misinformation. It's not the first time the publication has raised eyebrows. *Business Insider*. <https://www.businessinsider.com/who-is-zero-hedge-finance-blog-that-spread-coronavirus-misinformation-2020-2>
- De Dobbelaer, R., Van Leuven, S., & Raeymaeckers, K. (2018). The Human Face of Health News: A Multi-Method Analysis of Sourcing Practices in Health-Related News in Belgian Magazines. *Health Communication*, 33(5), 611–619. <https://doi.org/10.1080/10410236.2017.1287237>
- Enkhbayar, A. (2022). *ScholCommLab/altmetric-news-quality: Release for published dataset on Dataverse*. Zenodo. <https://doi.org/10.5281/zenodo.5963168>
- Fitzpatrick, K. (2021). *Generous Thinking: A Radical Approach to Saving the University*. Johns Hopkins University Press.
- Fleerackers, A., Nehring, L., Alperin, J. P., Enkhbayar, A., Maggio, L. A., & Moorhead, L. (2022). *Replication data for Identifying science in the news* [Data set]. Harvard Dataverse. <https://doi.org/10.7910/DVN/WNDOFL>
- Fleerackers, A., Riedlinger, M., Moorhead, L., Ahmed, R., & Alperin, J. P. (2021). Communicating scientific uncertainty in an age of COVID-19: An investigation into the use of preprints by digital media outlets. *Health Communication*, 0(0), 1–13. <https://doi.org/10.1080/10410236.2020.1864892>
- Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálffy, M., Nanni, F., & Coates, J. A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLOS Biology*, 19(4), e3000959. <https://doi.org/10.1371/journal.pbio.3000959>

- 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, D. Y., & 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>
- Funk, C., Gottfried, J., & Mitchell, A. (2017, September 20). Science News and Information Today. *Pew Research Center*. <https://www.journalism.org/2017/09/20/science-news-and-information-today/>
- Gesualdo, N., Weber, M. S., & Yanovitzky, I. (2020). Journalists as Knowledge Brokers. *Journalism Studies*, 21(1), 127–143. <https://doi.org/10.1080/1461670X.2019.1632734>
- Hatch, A., & Curry, S. (2020). Changing how we evaluate research is difficult, but not impossible. *ELife*, 9, e58654. <https://doi.org/10.7554/eLife.58654>
- Hicks, D. (2005). The Four Literatures of Social Science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems* (pp. 473–496). Springer Netherlands. https://doi.org/10.1007/1-4020-2755-9_22
- Karmakar, M., Banshal, S. K., & Singh, V. K. (2021). A large-scale comparison of coverage and mentions captured by the two altmetric aggregators: Altmetric.com and PlumX. *Scientometrics*, 126(5), 4465–4489. <https://doi.org/10.1007/s11192-021-03941-y>
- Lehmkuhl, M., & Promies, N. (2020). Frequency distribution of journalistic attention for scientific studies and scientific sources: An input–output analysis. *PLOS ONE*, 15(11), e0241376. <https://doi.org/10.1371/journal.pone.0241376>
- Maggio, L. A., Ratcliff, C. L., Krakow, M., Moorhead, L. L., Enkhbayar, A., & Alperin, J. P. (2019). Making headlines: An analysis of US government-funded cancer research mentioned in online media. *BMJ Open*, 9(2), e025783. <https://doi.org/10.1136/bmjopen-2018-025783>
- Matthias, L., Fleerackers, A., & Alperin, J. P. (2020). Framing Science: How Opioid Research Is Presented in Online News Media. *Frontiers in Communication*, 5(64). <https://doi.org/10.3389/fcomm.2020.00064>
- Meschede, C., & Siebenlist, T. (2018). Cross-metric compatability and inconsistencies of altmetrics. *Scientometrics*, 115(1), 283–297. <https://doi.org/10.1007/s11192-018-2674-1>
- Moorhead, L., Krakow, M., & Maggio, L. (2021). What cancer research makes the news? A quantitative analysis of online news stories that mention cancer studies. *PLOS ONE*, 16(3), e0247553. <https://doi.org/10.1371/journal.pone.0247553>
- Oliveira, T., Araujo, R. F., Cerqueira, R. C., & Pedri, P. (2021). Politização de controvérsias científicas pela mídia brasileira em tempos de pandemia: A circulação de preprints sobre Covid-19 e seus reflexos. *Revista Brasileira de História da Mídia*, 10(1), Article 1. <https://doi.org/10.26664/issn.2238-5126.101202111810>
- Ortega, J. L. (2019a). Availability and audit of links in altmetric data providers: Link checking of blogs and news in altmetric.com, crossref event data and plumx. *Journal of Altmetrics*, 2(1), 4. <https://doi.org/10.29024/joa.14>
- Ortega, J. L. (2019b). The coverage of blogs and news in the three major altmetric data providers. *ISSI 2019 Proceedings*, 1, 75–86.
- Ortega, J. L. (2020a). Blogs and news sources coverage in altmetrics data providers: A comparative analysis by country, language, and subject. *Scientometrics*, 122(1), 555–572. <https://doi.org/10.1007/s11192-019-03299-2>

- Ortega, J. L. (2020b). Altmetrics data providers: A meta-analysis review of the coverage of metrics and publication. *El Profesional de La Información*, 29(1). <https://doi.org/10.3145/epi.2020.ene.07>
- Ortega, J. L. (2021). How do media mention research papers? Structural analysis of blogs and news networks using citation coupling. *Journal of Informetrics*, 15(3), 101175. <https://doi.org/10.1016/j.joi.2021.101175>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine*, 18(3), e1003583. <https://doi.org/10.1371/journal.pmed.1003583>
- Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter: How and why scholars cite on Twitter. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–4. <https://doi.org/10.1002/meet.14504701201>
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010, October 26). *Altmetrics: A manifesto*. <http://altmetrics.org/manifesto>
- Robinson-Garcia, N., Arroyo-Machado, W., & Torres-Salinas, D. (2019). Mapping social media attention in Microbiology: Identifying main topics and actors. *FEMS Microbiology Letters*, 366(7), fnz075. <https://doi.org/10.1093/femsle/fnz075>
- Schultz, T. (2021). All the research that's fit to print: Open access and the news media. *Quantitative Science Studies*, 2(3), 828–844. https://doi.org/10.1162/qss_a_00139
- Serghiou, S., Marton, R. M., & Ioannidis, J. P. A. (2021). Media and social media attention to retracted articles according to Altmetric. *PLOS ONE*, 16(5), e0248625. <https://doi.org/10.1371/journal.pone.0248625>
- Sevryugina, Y. V., & Dicks, A. J. (2021). Publication practices during the COVID-19 pandemic: Biomedical preprints and peer-reviewed literature [preprint]. *BioRxiv*. <https://doi.org/10.1101/2021.01.21.427563>
- Shoemaker, P. J., & Vos, T. P. (2009). *Gatekeeping theory*. Routledge.
- Thelwall, M. (2018). Altmetric Prevalence in the Social Sciences, Arts and Humanities: Where are the Online Discussions? *Journal of Altmetrics*, 1(1), 4. <https://doi.org/10.29024/joa.6>
- Torres-Salinas, D., Castillo-Valdivieso, P.-Á., Pérez-Luque, Á., & Romero-Frías, E. (2018). Altmétricas a nivel institucional: Visibilidad en la Web de la producción científica de las universidades españolas a partir de Altmetric.com. *El Profesional de La Información*, 27(3), 483. <https://doi.org/10.3145/epi.2018.may.03>
- Torres-Salinas, D., Gorraiz, J., & Robinson-Garcia, N. (2018). The insoluble problems of books: What does Altmetric.com have to offer? *Aslib Journal of Information Management*, 70(6), 691–707. <https://doi.org/10.1108/AJIM-06-2018-0152>
- Yanovitzky, I., & Weber, M. S. (2019). News Media as Knowledge Brokers in Public Policymaking Processes. *Communication Theory*, 29(2), 191–212. <https://doi.org/10.1093/ct/qtz023>
- Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, 101(2), 1491–1513. <https://doi.org/10.1007/s11192-014-1264-0>